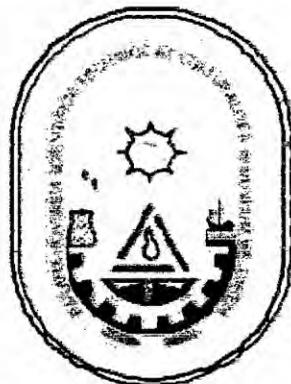


UNIVERSIDAD NACIONAL DEL CALLAO

ESCUELA DE POSGRADO

SECCIÓN POSGRADO DE LA FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA



**“MODELAMIENTO DE LA POTENCIA DE
INYECCIÓN, DE UN PARQUE EÓLICO DE
60MW EN PIURA”**

**TESIS PARA OPTAR EL GRADO ACADÉMICO DE MAESTRO
EN INGENIERÍA ELÉCTRICA. MENCIÓN EN GESTIÓN DE
SISTEMAS DE ENERGÍA ELÉCTRICA**

JESÚS PERFUMO FUERTES ALVARADO

Callao, 2014

PERÚ

HOJA DE REFERENCIA DEL JURADO

MIEMBROS DEL JURADO

DOCTOR	: FERNANDO JOSÉ OYANGUREN RAMÍREZ	PRESIDENTE
MAGÍSTER	: SANTIAGO LINDER RUBIÑOS JIMENEZ	SECRETARIO
MAGÍSTER	: HUGO FLORENCIO LLACZA ROBLES	MIEMBRO
MAGÍSTER	: FRANCO IVÁN VÉLIZ LIZÁRRAGA	MIEMBRO
MAGISTER	: JUAN HERBER GRADOS GAMARRA	ASESOR

Nº DE LIBRO : 01

FOLIO : 025

FECHA DE APROBACIÓN : Diciembre 23, 2014.

RESOLUCIÓN DIRECTORAL : 108-2014-DSPG-FIEE

DEDICATORIA

**A mis padres, hermanos,
esposa e hijos, a mi
amigo el Dr. Ing Alfredo
Novoa Peña y en
especial a mi alma mater
que es la Universidad
Nacional del Callao**

AGRADECIMIENTO

Agradezco a DIOS, a la Virgen del Carmen y a las personas que han marcado mi vida de algún modo u otro. Son muchas, pero aquí va un esbozo de las más importantes.

Empiezo por mis padres quienes me dieron vida, educación, apoyo y consejos y a mis hermanos por su apoyo emocional en cada etapa de mi vida. Especialmente quiero dedicar la tesis a mis hijos porque son los motivos de continuar escalando en cada peldaño de la vida. Otra dedicatoria muy especial es para mi esposa Rossmery por su apoyo constante y su comprensión.

Agradezco al Dr. Ingeniero Juan Grados Gamarra por su asesoramiento en la realización de la presente tesis y al Dr. Ingeniero Alfredo Novoa Peña por la idea inicial y a su apoyo constante en la materialización de la presente tesis.

Quienes me conocen saben que no pueden faltar mis compañeros de trabajo y en especial al Crl Marco Campos Gerónimo y al Crl Augusto Ruiz Carreño, además de ser grandes compañeros quiero agradecerle la paciencia que han demostrado conmigo no sólo durante la realización de la maestría y la tesis sino durante el tiempo que nos tocó trabajar juntos.

INDICE
MODELAMIENTO DE LA POTENCIA DE INYECCION, DE UN PARQUE
EOLICO DE 60MW EN PIURA

CARÁTULA	I
PAGINA DE RESPETO	II
HOJA DE REFERENCIA DEL JURADO Y APROBACIÓN	III
DEDICATORIA	IV
AGRADECIMIENTO	V
ÍNDICE	VI
TABLAS DE CONTENIDO	X
RESUMEN	1
ABSTRACT	2
I. PLANTEAMIENTO DE LA INVESTIGACIÓN	3
1.1 Identificación del problema	3
1.2 Formulación de problemas	4
1.3 Objetivos de la investigación (general y específicos)	6
1.4 Justificación	6
II. MARCO TEÓRICO	8
2.1 Antecedentes del estudio	8
2.1.1 Antecedentes Nacionales	8
2.1.2 Antecedentes Internacionales	13
2.2 Bases teóricas	28
2.2.1 Parque eólico	28
a. Parques eólicos en el mundo	28
b. Parques eólicos en el Perú	31
c. Componentes de un parque eólico	38
d. El viento	39
1) Naturaleza del viento	39
2) Historia de la utilización del Viento	42
3) Características del viento	51
a) Velocidad del viento media anual	51
b) Velocidad del viento diurno anual	53
c) Velocidad del viento diurno mensual	54
d) Variación mensual de la velocidad del viento	54
e) Distribución de frecuencia de la velocidad del viento	56
f) Dirección del viento	59
4) Tipos de viento	59
a) Brisas marinas	59
b) Vientos de montaña	59
c) Variaciones diurnas (noche y día) del viento	60
d) Efecto túnel	60
e) Efecto de la colina	61
f) Variación vertical del viento	63
5) Clasificación del viento	67
e. Temperatura del aire	68

f. Rugosidad	70
1) Clasificación del terreno	70
a) Flujo sobre terreno con obstáculos	72
g. La humedad y su influencia con la temperatura	74
h. Variación de la densidad vertical	74
i. Tipos de modelos para la predicción del tiempo	76
2.2.2 Modelamiento de la potencia de inyección	77
a. Potencia del viento. Máxima potencia extraíble.	80
b. Modelos de previsión de viento	82
c. Necesidad de predicción de potencia eólica.	83
d. Horizonte de predicción	84
e. Incertidumbre en la predicción	85
f. Probabilidad de la potencia	86
g. Redes neuronales	87
1) Neuronas biológicas	87
2) Modelo de una Neurona Artificial	92
a) Modelo general	92
b) Modelo estándar	97
3) Redes neuronales y arquitecturas	98
a) Tipos de arquitectura	100
4) Aprendizaje de una red neuronal	101
a) Fase de aprendizaje. Convergencia	103
b) Fase de recuerdo o ejecución. Estabilidad	106
5) Clasificación de los modelos neuronales	108
6) Computabilidad neuronal	110
7) Realización y aplicaciones de las redes neuronales	111
8) Redes neuronales supervisadas	114
a) Asociador lineal	114
b) Perceptron simple	119
c) Adaline	127
d) El perceptron multicapa (MLP)	129
9) Capacidad de generalización de la red	139
a) Validación cruzada (cross-validation)	140
b) Número de ejemplos de entrenamiento	143
c) Reducción del tamaño de la arquitectura de una red	145
h. Medida de Error	146
2.3 Definición de términos básicos	148
III. VARIABLES E HIPÓTESIS	152
3.1 Definición de las variables	152
3.2 Operacionalización de variables	152
3.3 Hipótesis general e hipótesis específicas	154
IV. METODOLOGÍA	155
4.1 Tipo de investigación	155
4.2 Diseño de la Investigación	155

4.3 Población y muestra	156
4.4 Técnicas e Instrumentos de recolección de datos	156
4.5 Procedimientos de recolección de datos	156
4.6 Procesamiento estadístico y análisis de datos	156
4.7 Metodología de trabajo	157
4.8 Matlab Toolbox	157
4.8.1 Introducción	157
4.8.2 Backpropagation	158
4.8.3 Red feedforward	161
4.8.4 Creando una red (newff).	162
4.8.5 Simulación (sim)	164
4.8.6 Entrenamiento	164
a. Entrenamiento por lote (train)	165
b. Entrenamiento por incremental (<i>traingd</i>)	165
c. Descenso de gradiente por lotes con momentum (<i>traingdm</i>).	167
4.8.7 Entrenamiento rápido	169
a. Taza de aprendizaje (<i>traingda</i> , <i>traingdx</i>)	169
b. Rezago (<i>trainrp</i>)	171
4.8.8 Algoritmos de gradiente conjugado	172
a. Actualización de fletcher-reeves (<i>traincgf</i>)	173
b. Actualización de polak-ribière (<i>traincgp</i>)	175
c. Restablecimiento de powell-beale (<i>traincgb</i>)	176
d. Gradiente conjugado escalado (<i>trainscg</i>)	177
4.8.9 Rutinas de búsqueda lineal	178
a. Búsqueda de sección dorada (<i>srchgol</i>)	178
b. Búsqueda de brent (<i>srchbre</i>)	178
c. Búsqueda bisección-cúbica híbrida (<i>srchhyb</i>)	179
d. Búsqueda de carambola (<i>srchcha</i>)	180
e. Backtracking (<i>srchbac</i>)	180
4.8.10 Algoritmos cuasi-newton	181
a. Algoritmo de bfgs (<i>trainbgf</i>)	181
b. Algoritmo secante de un paso (<i>trainoss</i>)	182
c. Levenberg-marquardt (<i>trainlm</i>)	183
d. Levenberg-marquardt de memoria reducida (<i>trainlm</i>)	185
4.8.11 Regularización	186
a. Función de desempeño modificada	186
b. Regularización automatizada (<i>trainbr</i>)	187
4.8.12 Detención temprana	189
4.8.13 Preproceso y postproceso	191
4.9 Comparación de resultados estadísticos y Matlab	194
4.10 Elaboración de un modelo de predicción de potencia en la zona	195

V.RESULTADOS	196
5.1 Resultados finales	196
5.1.1 Determinación de la potencia de generación eólica con RETScreen	196
5.1.2 Determinación de la potencia de generación eólica con fórmulas matemáticas	197
VI. DISCUSIÓN DE RESULTADOS	198
6.1 Contrastación de hipótesis con los resultados	198
6.2 Contrastación de resultados con otros estudios similares.	199
VII. CONCLUSIONES	200
VIII. RECOMENDACIONES	203
IX. REFERENCIAS BIBLIOGRÁFICAS	204
ANEXOS:	
1. Matriz de Consistencia	207
2. Distribución Weibull	208
3. Atlas eólico de Piura	210
4. Croquis de la zona de estudio	211

TABLAS DE CONTENIDO

• FIGURA N° 2.1 VARIACIÓN DEL VIENTO	40
• FIGURA N° 2.2 VARIACIÓN DEL VIENTO A ESCALA GLOBAL	41
• FIGURA N° 2.3 MOLINO PERSA	43
• FIGURA N°2.4 MOLINO PARA BOMBEO	46
• FIGURA N°2.5 ROSA DE VIENTOS	58
• FIGURA N°2.6. EFECTO TÚNEL	61
• FIGURA N°2.7. EFECTO COLINA	62
• FIGURA N° 2.8 PERFIL VERTICAL DE TEMPERATURA PARA LA ATMOSFERA	69
• FIGURA N°2.9 DETERMINACION DE UN TERRENO LLANO	71
• FIGURA N°2.10 ESTELA DE LA VELOCIDAD DEL VIENTO CAUSADO POR LOS OBSTACULOS	73
• FIGURA N°2.11, DESVIO DEL VIENTO	80
• FIGURA N°2.12 MODELO GENÉRICO NEURONA ARTIFICIAL	92
• FIGURA N°2.13 FUNCIONES DE ACTIVACIÓN MÁS USUALES	96
• FIGURA N°2.14 MODELO DE NEURONA ESTÁNDAR	97
• FIGURA N°2.15 EJEMPLO DE UNA RED NEURONAL TOTALMENTE CONECTADA.	100
• FIGURA N°2.16 EJEMPLOS DE ARQUITECTURAS NEURONALES	100
• FIGURA N°2.17 CLASIFICACIÓN DE LOS ANS POR EL TIPO DE APRENDIZAJE Y ARQUITECTURA	109
• FIGURA N°2.18 ASOCIADOR LINEAL	115
• FIGURA N°2.19 MODELO DEL FOTOPERCEPTRÓN DE ROSENBLATT	120
• FIGURA N°2.20 ARQUITECTURA (IZQUIERDA) Y FUNCIÓN DE TRANSFERENCIA (DERECHA) DE UN PERCEPTRÓN SIMPLE	121
• FIGURA N°2.21 REGIÓN DE DECISIÓN CORRESPONDIENTE A UN PERCEPTRÓN SIMPLE CON DOS NEURONAS DE ENTRADA	123
• FIGURA N°2.22: FUNCIONES LÓGICAS NAND (a) y XOR (b)	124
• FIGURA N°2.23 ESTRUCTURA GENERAL DE LA RED TIPO ADALINE	128
• FIGURA N°2.24 REGIONES DE DECISIÓN OBTENIDAS PARA EL PERCEPTRÓN SIMPLE (ARRIBA), EL PERCEPTRÓN MULTICAPA CON UNA CAPA OCULTA (EN MEDIO) Y EL PERCEPTRÓN MULTICAPA CON DOS CAPAS OCULTAS (ABAJO)	130
• FIGURA N° 2.25: ARQUITECTURA (IZQUIERDA) Y FUNCIÓN DE DESEMPEÑO(DERECHA) PARA EL PERCEPTRÓN	

MULTICAPA	131
• FIGURA N°2.26: ARQUITECTURA DEL PERCEPTRÓN MULTICAPA	132
• FIGURA N°2.27 EVOLUCIÓN DEL ERROR DE APRENDIZAJE Y DEL ERROR DE GENERALIZACIÓN. A LA IZQUIERDA, SITUACIÓN IDEALIZADA, A LA DERECHA SITUACIÓN REAL	141
• FIGURA N° 4.1 NEURONA CON R ENTRADAS	158
• FIGURA N°4.2 FUNCIÓN DE TRANSFERENCIA LOGARÍTMICA	159
• FIGURA N°4.3 FUNCIÓN DE TRANSFERENCIA TANGENCIAL	159
• FIGURA N° 4.4 FUNCION DE TRASLADO LINEAL(PURELIN)	160
• FIGURA N°4.5 RED NEURONAL FEEDFORWARD	161
• FIGURA N°4.6 RED QUE PUEDE USARSE COMO APROXIMADOR	162
• GRAFICO N° 2.1 VARIACIÓN EN EL DÍA DE LA VELOCIDAD MEDIA ANUAL DEL VIENTO	52
• GRAFICO N°2.2 VARIACIÓN MENSUAL DE LA VELOCIDAD DEL VIENTO PARA DIFERENTES ALTURAS DE LAS ESTACIONES	54
• GRAFICO N°2.3 FRECUENCIA DEL VIENTO ANUAL PARA UNA ALTURA DE 24m	57
• GRAFICO N° 2.4PROBABILIDAD DEL VIENTO ANUAL PARA UNA ALTURA DE 24m	58
• GRAFICO N° 2.5 VARIACION EN EL DIA DE LA VELOCIDAD MEDIA ANUAL DEL VIENTO A 80m.	66
• GRAFICO N°2.6 VARIACION DE LA TEMPERATURA PROMEDIO A 24m	69
• GRAFICO N°2.7 VARIACION DE LA TEMPERATURA PROMEDIO A 80m	70
• GRAFICO N°2.8 ESCALA DE RUGOSIDADES	73
• GRAFICO N°2.9 VARIACION DE LA DENSIDAD DEL AIRE MEDIA ANUAL	76
• GRAFICO N°2.10 CURVA P-V DE UN AEROGENERADOR DE 2MW	79
• GRAFICO N° 2.11 COEFICIENTE DE POTENCIA (C_p) Vs k	81
• GRAFICO N°2.12 VARIACION DE LA POTENCIA MEDIA ANUAL	81
• GRAFICA N°2.13 PROBABILIDAD DE LA POTENCIA PARA 80m	87
• GRAFICO N° 4.1 MUESTRA LA SALIDA DE LA RED ESPECIALIZADA	189
• GRAFICO N° 4.2 RESPUESTA DE LA RED	191
• GRAFICO N°4.3 COMPARACION DE LA VELOCIDAD DE	

PREDICCIÓN ESTADÍSTICA Y VELOCIDAD REAL	195
• GRAFICO N° 4.4 COMPARACION DE LA VELOCIDAD DE PREDICCIÓN VIENTO POR REDES NEURONALES Y LA VELOCIDAD REAL	195
• GRAFICO N° 5.1 PREDICCIÓN DE LA POTENCIA DE SALIDA PARA EL 30 DE ABRIL. UTILIZANDO RETScreen.	197
• GRAFICO N° 5.2 PREDICCIÓN DE LA POTENCIA DE SALIDA PARA EL 30 DE ABRIL; UTILIZANDO FORMULAS MATEMATICAS	198
• TABLA N° 1.1 DESMEMBRACIÓN MENTAL DE LA INGENIERÍA ELÉCTRICA	4
• TABLA N° 2.1 POTENCIAL EÓLICO DEL PERÚ	10
• TABLA N°2.2 RESUMEN DE PARÁMETROS METEOROLÓGICOS	33
• TABLA N°2.3 CARACTERÍSTICA DE SENSORES METEOROLÓGICOS	42
• TABLA N° 2.4 VELOCIDAD DEL VIENTO Y LA FRECUENCIA	57
• TABLA N° 2.5 CLASIFICACION DEL VIENTO	67
• TABLA N° 3.1 VARIABLES DIMENSIONES E INDICADORES	153

RESUMEN

El aprovechamiento de la energía del viento es antiguo pero en los últimos años se ha desarrollado mucho, aprovechando los progresos en aerodinámica y electrónica de los reguladores. La energía eólica en la actualidad se puede considerar como una tecnología fructífera y con sus raíces que le dan el sustento, en la generación de electricidad, la cual crece en todos los continentes de manera vertiginosa. El problema principal de esta energía es contar con un predictor de potencia eólica, en un rango estimado donde exista una probabilidad pre asignada del valor de la salida del predictor, en un intervalo de confianza para el Parque eólico de Piura; siendo el propósito del presente trabajo de contar con un plan para predecir la **potencia generada por el parque eólico de Piura** de manera que el COES programen sus operaciones de despacho con un día de anticipación.

Los métodos que se emplearon fueron Weibull que determino la probabilidad de buenos vientos. Para la predicción se utilizó métodos estadístico y RNA, donde se obtuvo mejores resultados con RNA, siendo muy importante los datos de ingreso para el entrenamiento.

Se concluyó que el método utilizado en la predicción es eficiente a la vez que es económico, también es rentable por presentar vientos elevados en la hora punta del servicio eléctrico.

ABSTRACT

Harnessing wind energy is old but in recent years there has been much progress in taking advantage of aerodynamics and electronics regulators. Wind energy today can be considered as a technology bears fruit and roots that give sustenance in electricity generation, which grows on every continent dramatically. The main problem of this energy is to have a predictor of wind power in an estimated range where pre assigned a probability value of the output of the predictor, a confidence interval for the wind farm there Piura; being the purpose of this work to have a plan to predict the power generated by the wind farm so Piura COES scheduled operations office one day in advance.

The methods used were determined Weibull probability of good winds. Statistical methods and RNA, where best results were obtained with RNA, the input data to be important for training was used for prediction.

It was concluded that the method used in the prediction is efficient while it is economical, it is also profitable for its high winds rush hour electrical service

CAPÍTULO I

PLANTEAMIENTO DE LA INVESTIGACIÓN

1.1 Identificación del problema

Cuándo las turbinas eólicas o parque eólico son conectados a la red eléctrica local, hay ciertas consideraciones a tener en cuenta en relación a la potencia y al funcionamiento del sistema durante desórdenes en el mismo, por el viento que es variable.

Los Parques Eólicos en proyecto no cuentan aún con un sistema predictor de potencia excepto las ya operativas como el de Marcona. Todos los parques eólicos integrados al SEIN², cuentan con sistema de este tipo como exigencia de parte del COES¹; organismo encargado de la coordinación de la operación del SEIN²; con el objetivo de realizar una predicción a corto plazo de la potencia de generación eólica a inyectar con mayor precisión.

El problema que se plantea en este trabajo es el desarrollo de un predictor de potencia para este parque, en un rango estimado donde exista una probabilidad pre asignada del valor de la salida del predictor, en adelante intervalo de confianza, de las predicciones entregadas por el modelo.

El problema está enmarcado el centrales de generación eléctricas, no convencionales, como lo especifica el Tabla N° 1.1 (véase la tabla N°1.1, en la página "4").

¹Comité de Operación Económica del Sistema

²Sistema Eléctrico Interconectado Nacional

TABLA N° 1.1
DESMEMBRACIÓN MENTAL DE LA INGENIERÍA ELÉCTRICA

I. GENERACIÓN DE ENERGÍA ELÉCTRICA.	A. CONVENCIONALES	<ul style="list-style-type: none"> a. Hidráulicas b. Térmicas c. Nucleares d. Geotérmica e. A Gas y Diessel f. Biomasa g. Microcentrales
	B. NO CONVENCIONALES	<ul style="list-style-type: none"> a. Paneles solares b. Centrales Eólicas c. Mareomotrices d. Parques Fotovoltaicos

1.2 Formulación del problema

1.2.1 Problema general:

El presente proyecto se encuentra enmarcado en la línea de investigación N° 03 sugerida por el Ministerio de Energía y Minas (CARELEC¹): Energías Renovables y el tema de interés: Fluctuaciones de potencia en la generación eólica.

La energía eólica está en pleno crecimiento y va en camino de liderar con respecto a las otras energías por no contaminar el ambiente y ser una producción de energía limpia, que es considerada los suministros de energía del mañana juntamente con la energía solar; el Perú está siendo uno de los países que menos está aprovechando de este tipo de energía por consiguiente uno de los problemas o temores es la Predicción de la Potencia debido a las Fluctuaciones de Potencia de Parques Eólicos,

¹Consejo de Administración de Recursos para la Capacitación en Electricidad.

esto es debido a que las turbinas eólicas se basan en la acción del viento sobre palas y la variación de la velocidad del viento hace variar la potencia.

Uno de los problemas que presentan estos generadores es su inconstancia de la velocidad, lo que repercute la desconfianza para ingresar al SEIN. Sin embargo, se han desarrollado equipos de naturaleza electrónica, que permiten varias formas de control de este parámetro.

La potencia producida por una turbina eólica no es constante, especialmente no lo es si el viento está por debajo de la potencia nominal.

Por lo que se plantea la siguiente pregunta:

¿Cómo se modela la potencia de inyección de un parque eólico de 60MW ubicado en la ciudad de Piura?

a. Problema específica:

Para la selección específica del problema se plantea las siguientes preguntas:

- 1) ¿Cómo se realiza la predicción de la potencia en países que lideran la producción de energía eólica?
- 2) ¿Cuáles son los conceptos fundamentales para el modelamiento de la potencia de inyección de un parque eólico?
- 3) ¿Cuáles son los programas que se emplean en la predicción (el cálculo) de la potencia de un parque eólico?

- 4) ¿Qué modelos lineales y no lineales se pueden emplear para la predicción de la potencia de un parque eólico con intervalos de confianza?

1.3 Objetivos de la investigación

a. General:

Modelar la potencia de inyección de un parque eólico de 60MW ubicado en la ciudad de Piura.

b. Específicos:

- 1) Revisar los modelos de predicción de la potencia en países que lideran la producción de energía eólica.
- 2) Identificar los conceptos fundamentales para el modelamiento de la potencia de inyección de un parque eólico.
- 3) Identificar los programas que se emplean para simular la potencia de salida de un parque eólico.
- 4) Revisar los modelos lineales y no lineales que se emplean para la predicción de la potencia de un parque eólico con intervalos de confianza.

1.4 Justificación de la investigación

1.4.1 Tecnológica.

Es factible aplicar un modelo de predicción de potencia a corto plazo, de un parque eólico de 60MW en Piura, con aerogeneradores de 2.0 MW cada uno. Las investigaciones existentes en otros países y se refieren a

parques eólicos de diferentes magnitudes, las mismas que tienen características diferentes.

1.4.2 Económica.

Los resultados de la investigación serán aplicados en beneficio de las empresas generadoras de energía eólica. Asimismo, servirán a las facultades de ingeniería eléctrica y electrónica de las universidades y otras entidades que tienen que ver con el problema de la interconexión de un parque eólico con la red del SEIN.

Así mismo, La energía eólica es una de las tecnologías más económicas para obtener energías renovables.

1.4.3 Social.

Este tipo de energía, por su tecnología al principio el costo de Kw-h es elevado, comparado con las centrales convencionales, pero con el transcurrir de los años dicho precio se está reduciendo considerablemente, por consiguiente los consumidores se beneficiarán al reducir dichos costos. Adicionalmente va de la mano con el desarrollo sostenible.

CAPITULO II MARCO TEÓRICO

2.1 Antecedentes del estudio

2.1.1 Antecedentes Nacionales

En el Perú, se ha realizado una estimación de la potencia eólica total y aprovechable. Una estimación de estas características puede realizarse de diversas maneras y parte de unos supuestos que pretenden ser razonables pero que no dejan de ser arbitrarios y que se exponen a continuación.

Para filtrar las zonas con potencial eólico suficiente, se ha calculado la producción en cada uno de los puntos simulados sobre tierra con un generador estándar de 2MW de potencia nominal y 80 m de diámetro de rotor a la altura de 80 m sobre el suelo.

El factor de capacidad mide el rendimiento real de un aerogenerador en un emplazamiento, dado que éstos sólo generan su potencia nominal en un rango determinado de vientos; fuera de ese rango su potencia es inferior. Si se expresa matemáticamente, el factor de capacidad no es más que el cociente entre la potencia media producida y la potencia nominal del aerogenerador.

Para el cálculo de la potencia total se ha considerado dentro del territorio disponible a cada uno de los nodos de simulación, de 1x1 Km, que presentan un factor de capacidad bruto igual o superior al 20%. El valor de corte para realizar la selección depende de la retribución de la energía en la zona en cuestión; en este caso, se ha seleccionado un valor bastante orientativo en países en que el sector está desarrollado.

La última suposición en el cálculo es el número de aerogeneradores de las características supuestas que razonablemente pueden ubicarse en cada recuadro de 1 Km² de la simulación. Según la experiencia del Consultor, un número orientativo serían 3 aerogeneradores. Aunque geoméricamente podrían caber 4, la realidad es que las instalaciones sobre el terreno siempre tienen otros condicionantes, (variaciones del viento dentro del punto de rejilla, dificultades orográficas, etc.), por lo que la cifra real se reduce. Por lo tanto, la hipótesis del Consultor es que en cada nodo se puede instalar 6 MW.

Con estos supuestos, se ha estimado la potencia eólica total del Perú en algo más de 77000 MW, sin considerar zonas off-shore (en el mar), distribuida según lo indicado en la segunda columna (véase la Tabla N°3, en la página "10"), incluida en este numeral. Se considera potencia aprovechable a la fracción de la total que es realizable a nivel teórico por no existir impedimentos prácticos. En un estudio de microescala, del nivel de un parque eólico, el número de estos impedimentos es muy grande y en gran parte debe ser estudiado in situ. Para el presente Atlas Eólico, dado lo extenso del territorio, deben emplearse unos criterios limitados y muy claros para filtrar las ubicaciones en las que por razones evidentes no es posible una instalación eólica. Las zonas que se han eliminado del cálculo con este filtrado tenían uno o varios de los siguientes condicionantes:

- Altitud igual o superior a 3000 m.s.n.m.
- Pendiente igual o superior al 20%.

- Pertener a un centro poblado o al casco urbano de una ciudad.
- Estar ubicado en una zona protegida, arqueológica o de amortiguamiento.
- Existir en ese lugar ríos, quebradas o lagos.

Después de aplicar este filtrado, el cálculo ha arrojado un valor del potencial eólico aprovechable del Perú superior a los 22000 MW, distribuido según la tercera columna (véase la Tabla N°2.1).

**TABLA N° 2.1
POTENCIAL EÓLICO DEL PERÚ**

Departamento	Potencia Total (MW)	Potencia Aprovechable (MW)
Amazonas	1380	6
Ancash	8526	138
Apurímac	0	0
Arequipa	1992	1158
Ayacucho	114	0
Cajamarca	18360	3450
Callao	0	0
Cuzco	0	0
Huancavelica	0	0
Huánuco	54	0
Ica	18360	9144
Junín	48	0
La Libertad	4596	282
Lambayeque	2880	564
Lima	1434	156
Loreto	0	0
Madre de Dios	0	0
Moquegua	144	0
Pasco	0	0
Piura	17628	7554
Puno	162	0
San Martín	504	0
Tacna	942	0
Tumbes	0	0
Ucayali	0	0
TOTAL PERÚ	77394	22452

Con respecto a estas cifras, que son considerables, cabe señalar las siguientes observaciones:

- a. Las restricciones reales suelen ser superiores a las que se hacen en un estudio teórico. Por ejemplo, aquí no se han tenido en cuenta ni las vías de comunicación ni la infraestructura eléctrica. El motivo es que tanto una como la otra infraestructura pueden ser construidas con posterioridad. Sin embargo, muchas veces el coste de estas construcciones es tan elevado que no se llevan nunca a cabo. En resumen, la cifra obtenida puede considerarse más un límite superior que un valor realista del potencial realizable.
- b. Una de las principales diferencias de un parque eólico respecto a las centrales tradicionales es que su eficiencia es, comparativamente, muy baja. Los parques eólicos reales tienen normalmente factores de capacidad que oscilan entre el 20 y el 30%. Por lo tanto, en términos de energía disponible para el Sistema, se debe aplicar este mismo factor de corrección a las cifras presentadas.

Las técnicas convencionales en el estudio del potencial eólico requieren mucho tiempo y, a menudo, dependen de la disponibilidad de costosas torres meteorológicas. Asimismo, los modelos convencionales de flujo de viento son poco precisos ante regímenes de viento muy variables, e incluso en zonas de terreno moderadamente complejo, su exactitud decae sustancialmente con la distancia a la torre de observación más cercana.

Las técnicas de modelización mesoescalares y microescalares ofrecen una solución muy efectiva a todos estos problemas. Combinando la

utilización de un sofisticado modelo de simulación atmosférica, capaz de reproducir los patrones de viento a gran escala, con un modelo de viento microescalar que responda a las características del terreno y a la topografía, se pueden llevar a cabo estudios del potencial eólico en regiones extensas con un nivel de detalle imposible de alcanzar en el pasado. Además, hay que remarcar que no son necesarios datos de viento de entrada para conseguir resultados razonablemente precisos, a pesar de que los datos de torres meteorológicas son imprescindibles para confirmar el potencial eólico previsto en un punto concreto. La modelización atmosférica mesoescalar y microescalar reduce notablemente el coste y el tiempo necesarios para identificar y evaluar regiones potencialmente prometedoras para la implantación de proyectos eólicos.

Para la realización del presente estudio, el Consultor ha empleado el sistema de su propiedad WindSurvey, que consiste en una integración de modelos de simulación de la atmósfera, bases de datos, estaciones de trabajo y unidades de disco de gran capacidad. En el corazón de MesoMap está el MASS (Mesoscale Atmospheric Simulation System), un modelo físico-matemático que simula el comportamiento de la atmósfera a mesoescala. El modelo MASS se acopla a un modelo de microescala simplificado de flujo de viento, WindMap, que se utiliza para refinar la resolución espacial del MASS y para incluir los efectos locales de la rugosidad y la orografía. El MASS simula las condiciones atmosféricas

sobre la región de interés para un total de 366 días elegidos de forma aleatoria de un total de 15 años. Una vez finalizadas las simulaciones, los resultados se introducen en WindMap. En esta Consultoría, el MASS se ha ejecutado hasta llegar a una resolución de 5 Km., mientras que WindMap se ha ejecutado con una resolución de 1 Km. [1].

2.1.2 Antecedentes Internacionales

- a. **JIMENEZ MEJIA Miguel, Pronostico de la velocidad del viento utilizando modelos ARIMA y Redes Neuronales Artificiales.** Tesis pos grado. Morelia Michoacán de México. Universidad Michoacana de San Nicolás de Hidalgo. 2011.

En México cuentan con un gran potencial eólico, solo en el Istmo de Tehuantepec se estima que se podrían instalar más de 6 000MW eólicos. Instituciones gubernamentales como la Comisión Federal de Electricidad y el Instituto de Investigaciones Eléctricas, estudian estos sitios con gran potencial energético y desarrollan proyectos para la explotación de la energía eólica. Para poder hacer una planeación de la explotación de la energía eólica de manera eficiente fue necesario implementar técnicas para el análisis de series de tiempo como herramientas para el pronóstico de la velocidad del viento. [2]

Mencionan varias técnicas para hacer pronóstico de series de tiempo, la metodología ARIMA (Autorregresive Integratec Moving Average) y las Redes Neuronales Artificiales (RNA) han tenido grandes

aplicaciones en el pronóstico de series de tiempo. En este trabajo se comparó estas técnicas para el pronóstico de la velocidad del viento a corto plazo. Algunos autores afirman que la metodología Box-Jenkins (ARIMA) supera a las redes neuronales en el pronóstico a corto plazo, en este trabajo se concluyó que las RNA obtienen mejores resultados que otros métodos empleados inclusive supera al método ARIMA.

El objetivo principal de esta tesis es generar un modelo de predicción utilizando técnicas Estadísticas Autoregresivas y de Redes Neuronales Artificiales (ARIMA y RNA) para proporcionar la velocidad del viento a corto plazo en el Istmo de Tehuantepec Oaxaca, que operen a partir de registros históricos de dicha variable.

Siendo sus objetivos particulares, el de contribuir a la integración de una base tecnológica relacionada con el pronóstico de la velocidad del viento para la explotación masiva del mismo, como fuente de energía en México y el de generar un precedente en el modelado de pronóstico en planta eólica, que pueda ser utilizado como referencia en trabajos de modelado en otras regiones del país y otras disciplinas similares.

Llegando a las siguientes conclusiones, queda demostrado que las técnicas ARIMA y RNA trabajan de manera eficiente con series de tiempo en forma de velocidades del viento en el modelado de pronóstico a corto plazo. Los modelos generados han sido evaluados con medidas de error estadísticas para la correcta comparación e las

técnicas ARIMA y RNA. De la misma manera, se puede plantear la aplicación de dichos modelos para series de tiempo de otra naturaleza. La más importante a la que concluye es que, la preparación de los datos previa al análisis con RNA es determinante para la obtención de buenos resultados en el pronóstico de la velocidad del viento a corto plazo.

- b. **FERNANDEZ JIMENEZ Luis Alfredo, Modelos Avanzados para la Predicción a Corto Plazo de la Producción Eléctrica en Parques Eólicos.** Tesis Doctoral. Logroño. Universidad de Rioja. 2007

La predicción de sucesos futuros ha fascinado al hombre desde los tiempos más remotos, según historiadores, el primer hombre que fue capaz de realizar una predicción fundamental sobre una base científica fue Thales de Mileto, fue capaz de predecir un eclipse total del sol. El desarrollo de modelos de predicción ha sido, un campo científico relativamente activo y en el que se han sucedido continuos avances. La búsqueda de la relación existente entre el valor actual de una variable y los valores pasados de la misma, o de otras variables que puedan influir sobre el valor de la primera, constituye una de las bases del desarrollo de modelo de predicción. La suposición de que entendiendo las relaciones entre los datos del pasado, podremos extrapolar el comportamiento futuro, es la idea que subyace en todos los modelos de predicción.[3]

En las fuentes de energía renovables han surgido un nuevo problema de predicción en corto plazo en el sector de la energía eléctrica: el de la predicción de la energía eléctrica que estas instalaciones van a generar.

El autor se propuso como **objetivo** desarrollar un conjunto de **nuevos modelos avanzados para la predicción a corto plazo** de la producción eléctrica en parques eólicos, capaces de lograr los mejores valores de predicción posible con los datos disponibles.

Los originales modelos de predicción de la generación eléctrica que se han creado han sido desarrollados específicamente para contemplar la mayor o menor cantidad de datos utilizables para dicha predicción.

Los **objetivos específicos**:

- 1) Creación de nuevos modelos de predicción, a corto plazo, de la producción eléctrica en parques eólicos.
- 2) Desarrollo de un sistema de predicción con horizontes de predicción de hasta **72 horas**, para los casos en que se disponga la información relativamente más amplia con suficientes valores históricos de producción eléctrica y predicciones meteorológicas adecuadas.
- 3) Análisis de bondad de las predicciones de producción eléctrica logradas los nuevos modelos creados.

- 4) Desarrollo de sistemas de predicción de la producción eléctrica más económicos.
- 5) Aplicación y desarrollo de herramientas para el diseño de los nuevos modelos de predicción de la producción eléctrica.

También dice que para el sistema de predicción más complejo, a desarrollar para el mayor horizonte de predicción (hasta 72 horas), se hará una evaluación aproximada del valor económico de sus predicciones, en términos del beneficio que se obtendría como consecuencia de utilizar las mismas para realizar ofertas de venta de la energía eléctrica producida en el contexto español.

Como **conclusión** finaliza que, para el caso planteado con mayor información disponible (serie temporal con potencia eléctrica media entregada en la subestación y predicción completa de un modelo atmosférico), **se ha creado un original sistema global de predicción, compuesto por diferentes modelos de predicción, que cubre un horizonte de predicción de hasta 72 horas.** También se ha logrado determinar el modelo idóneo (de entre los que integran el sistema global) que conviene utilizar para cada horizonte de predicción y, además, en función del instante en el que se realiza la predicción. Por último, se ha ilustrado la determinación del valor económico de la predicción de la potencia eléctrica, mostrando el incremento del beneficio económico obtenido por el gestor del parque

eólico, en la venta de la energía eléctrica producida en dicho parque eólico, bajo opciones de venta posible en España.

- c. PONCELA BLANCO Marta, **Nuevos modelos de predicción eólica basados en series temporales**. Tesis Doctoral. Valladolid. Universidad de Valladolid. 2012

El trabajo que desarrolló, se centra en la mejora de los modelos estadísticos presentes en herramientas de predicción eólica actualmente existentes. Para ello, partieron de un caso particular como es Sipreolico. Las principales aportaciones de este trabajo son, en primer lugar, partiendo de los modelos basados en los filtros de Kalman hasta ahora desarrollados, la estimación de las matrices del sistema por máxima verosimilitud, las características del parque eólico a la predicción y se adapta la herramienta de predicción a cada parque.[4]

En segundo lugar, estudió la aplicación de técnicas multivariantes que permitan captar, a través de las correlaciones entre parámetros, relaciones no recogidas explícitamente en los modelos.

Como el problema objeto del estudio es no lineal y variante con el tiempo, es usual en predicción eólica recurrir a un conjunto de modelos en paralelo, ya que cada uno de ellos funcionará adecuadamente en determinados instantes de tiempo, dado que cada modelo capta mejor unas dinámicas que otras, por ello se plantea, de manera novedosa, distintas alternativas multivariantes de reducción

de datos como estrategia de combinación de predicción para la obtención de la predicción final

El objetivo de esta tesis fue, encontrar una familia de modelos que permita que en torno a ella se pueda desarrollar una metodología automática de modelado de series horarias de generación en parques eólicos. Se pretende conseguir modelos de alta prestaciones en la predicción a corto plazo (1 a 2 días). Todo ello dentro de un marco general de desarrollo y análisis.

Los objetivos particulares fueron, explorar las principales características particulares que presentan las series objeto de estudio, analizar la bondad de los resultados de la aplicación a estas series de la metodología y técnica actualmente existente, estudiar nuevos planteamientos y enfoques de modelado buscando el desarrollo de una metodología completa aplicada a herramientas que operan en tiempo real y con intervención mínima de parte del operario.

Dentro de esta metodología general, profundizó en el estudio de aquellas características particulares de especial impacto. Se tendrán que diagnosticar aquellos aspectos principales a analizar de forma diferenciada, es decir, el tratamiento dentro del modelo de lo siguiente:

- Efecto de la predicción meteorológica

- La combinación de predicciones
- El ensayo de nuevas familias de modelos multivariantes

Procediendo al desarrollo y programación de todo el marco anterior de análisis, con la mirada puesta en la mayor automatización posible de todo el proceso, de forma que permita de forma simple un tratamiento masivo de series.

Mostrar la bondad de los resultados en un parque eólico como el de Sotavento situado en terreno complejo y con la dificultad añadida de estar formado por aerogeneradores de distinta tecnología.

Como **conclusión** finaliza que, a la propuesta de técnicas para mejorar los modelos paramétricos que están actualmente operativos en herramientas de predicción eólica y de una manera novedosa en el área de predicción eólica, las técnicas multivariantes de reducción de datos para la realización de la combinación de las predicciones disponibles, con el objetivo de calcular una única predicción final. La ventaja de estas técnicas de combinación frente a las técnicas clásicas, es que estas últimas calculan los pesos o coeficientes de ponderación en función del error de predicción.

En las técnicas que propuso, la ponderación de cada predictor no se obtiene del error cometido hasta ese momento, sino en la cantidad de información, respecto al resto de predictores, que puede aportar. Las técnicas analizadas han sido componentes principales y mínimos cuadrados parciales, demostrándose que los resultados obtenidos con

ambas técnicas son similares. El método en dos etapas propuesto permite, para cada instante t , obtener una predicción mejor que los predictores de partida, demostrándose la utilidad de la metodología de combinación planteada en dicho capítulo y la ventaja de disponer de varios modelos de predicción operando en paralelo en problemas donde la naturaleza no lineal del mismo hace que los modelos sean aproximaciones más o menos válidas para cada rango de datos y valores de las entradas al sistema.

- d. PALOMARES LOSADA Ana María, **Caracterización del régimen de viento y desarrollo de un modelo de predicción eólica a escala local en el estrecho de Gibraltar**. Tesis Doctoral. Madrid. Universidad Complutense de Madrid. 2002.

Se realizó una caracterización del régimen de vientos en el Estrecho de Gibraltar, utilizando series anuales de observaciones en dos estaciones meteorológicas próximas. Asimismo, se elaboró un modelo de ajuste estadístico entre datos de salida de un modelo global de predicción meteorológica y las observaciones locales, que permite mejorar significativamente las previsiones de intensidad de viento en dicha región, donde se ubican diversos parques eólicos con un potencial conjunto de generación de energía eléctrica que supera actualmente los 130 MW.[5]

Existen dos objetivos fundamentales, que engloban otros objetivos secundarios:

- 1) Caracterización del régimen de vientos en el Estrecho de Gibraltar
 - a) Clasificación de las situaciones meteorológicas a escala sinóptica en superficie, que afectan al Estrecho y determinación de su influencia sobre el viento resultante.
 - b) Análisis estadístico pormenorizado de las variables que definen el viento.
 - c) Determinación de la relación existente entre las características del viento medido en los tres niveles de altura (10, 20 y 30 m).
 - d) Determinación de las curvas de duración de velocidad y ajuste a las curvas teóricas de Weibull
 - e) Estudio detallado de la curva de evolución diaria, y sus posibles causas.
 - f) Comprobación de que el régimen general del Estrecho domina sobre el régimen local y se hace extensible a toda la zona de su influencia.
- 2) Elaboración de un modelo de predicción del potencial eólico a escala local en el Estrecho.
 - a) Valoración de las posibles herramientas de predicción aplicables en la zona y elección de la más adecuada.
 - b) Desarrollo del modelo y validación del mismo.

Como conclusión finaliza que, es posible la aplicación de un modelo estadístico para la predicción operativa a escala local del módulo de

viento en Tarifa a 24, 36 y 48 horas, a partir de las predicciones que proporciona el modelo global del CEPMPM, con unos resultados bastante aceptables. No obstante, en la predicción de las calmas y algunos casos de ponientes de Verano, que es posible identificar a priori, la calidad de la predicción no es satisfactoria. El modelo estadístico podría ser optimizado utilizando valores medios en lugar de medidas instantáneas, lo cual constituye una de las tareas más inmediatas a realizar en el futuro.

- e. GARCIA LOBO Miguel, Métodos de predicción de la generación agregada de energía eólica. Tesis Doctoral. Leganés. Universidad Carlos III de Madrid. 2010

En esta tesis se propuso un método de predicción de la generación agregada de energía eólica en una región, mediante la búsqueda de similitudes entre el vector de viento previsto en algunas coordenadas de la región y otros vectores históricos de predicción de viento. El modelo propuesto se basa en modelos locales de suavizado de Media Ponderada y de Regresión Local Ponderada. Estos modelos ajustan de forma no paramétrica un modelo de predicción para cada punto de entrada al problema, que en nuestro caso será un vector de viento previsto.[6]

Para hacer el suavizado de los datos, se proponen y comparan distintos procedimientos para tener en cuenta la medida de distancias

entre vectores, el método de selección de datos cercanos, y las funciones de ponderación por distancia y antigüedad de los datos seleccionados.

Mediante un procedimiento de estimación adaptativa de parámetros, se conseguirá que el modelo de predicción se adecue a la evolución temporal de la compleja relación no lineal existente entre el viento previsto en la región y la potencia eólica total generada.

Las predicciones de potencia obtenidas con los modelos de Media Ponderada y Regresión Local Ponderada, se combinan para aprovechar las ventajas que ofrecen ambas aproximaciones en función de la complejidad del problema de predicción en cada momento.

El modelo propuesto se validará comparando sus predicciones con la agregación de predicciones que se obtienen con Sipreólico, una herramienta que calcula predicciones para cada uno de los parques en la región y después las suma.

Por último, para dar mayor valor a las predicciones agregadas, se propone un modelo de predicción probabilista condicionada a los vectores de viento previsto mediante estimación de densidades con Kernel, para el cual se utilizan las medidas de distancia y funciones de ponderación propuestas para el modelo de predicción de potencia

El objetivo fundamental que persiguió esta tesis consiste en desarrollar un modelo de predicción a corto plazo de la generación agregada de energía eólica en una región con varios parques eólicos instalados.

Dada la habitual carencia en la disponibilidad de datos específicos, como las características individuales de cada parque y las telemidas de potencia o viento real registrado en cada parque eólico, se desea un sistema de predicción que necesite la menor cantidad de datos posible. En este sentido, el modelo de predicción propuesto necesitará únicamente un conjunto histórico de predicciones de velocidad de viento en algunas coordenadas distribuidas por la región, y las correspondientes medidas de potencia eólica total registrada. Con esta información, el modelo podrá proporcionar predicciones de potencia eólica total en la región cada vez que reciba un nuevo vector de predicciones de viento en las coordenadas.

El algoritmo de predicción se basará en dos métodos no paramétricos de modelado local por suavizado, la Media Ponderada y la Regresión Local Ponderada, basándose se hará el cálculo en medidas de distancia entre vectores de predicción de viento. Para llevar a cabo el suavizado de los datos, se propondrán y evaluarán varias formas de medir la distancia entre vectores de viento, diferentes procedimientos de selección de datos cercanos, y varias funciones de ponderación en

base a la distancia y a la antigüedad de los datos. Uno de los trabajos principales de la tesis consistirá precisamente en evaluar todas estas alternativas, y seleccionar aquellas que se adapten mejor al problema de predicción regional eólica.

Con el objetivo de que el algoritmo de predicción pueda ser implementado en un entorno real de funcionamiento sin supervisión, se propondrá un método de estimación adaptativa de los parámetros del modelo de predicción de forma automática. Frente a una estimación global de parámetros con un periodo de entrenamiento, esta estimación adaptativa también puede mejorar las predicciones, gracias a la adaptación del modelo a la evolución de las condiciones climáticas a lo largo del año, o a los cambios en las características de los parques instalados en la región.

Debido a la complejidad del problema de predicción, y a las ventajas e inconvenientes que proporcionan las distintas configuraciones alternativas de modelado local, se establecerá un proceso de combinación en dos etapas de las predicciones obtenidas con los distintos modelos alternativos. Este proceso de combinación también podrá ser utilizado en el caso de disponer de varias fuentes de predicción de viento como entradas al modelo.

Como conclusión finaliza que, Para aprovechar las ventajas que presentan tanto los modelos de Media Ponderada como de Regresión Local Ponderada, en esta tesis se propone combinar las predicciones

obtenidas con ambos modelos, que al ser de distinto orden se adaptarán mejor o peor al problema de predicción dependiendo de la complejidad local del mismo en cada instante.

Mediante un procedimiento de combinación en dos etapas que se puede encontrar en la literatura, se ha comprobado que se obtienen mejores resultados cuando se combinan las predicciones obtenidas con ambas aproximaciones. La idoneidad de combinar predicciones también se ha demostrado utilizando como datos de entrada al modelo las predicciones de viento de dos agencias meteorológicas distintas.

Como evaluación final, se han comparado las predicciones obtenidas con el modelo de predicción propuesto, las calculadas por la herramienta de predicción Sipreólico, y las proporcionadas por otra agencia de predicción de potencia agregada. Se ha comprobado que los errores del modelo agregado propuesto son menores que los de la otra agencia de predicción agregada, y entre un 12 y un 16 % inferiores a los obtenidos con Sipreólico. Como ventaja añadida, el tiempo de cálculo del modelo de predicción de potencia agregada resulta ser de apenas 30 segundos cuando calcula predicciones para horizontes de 1 a 48 horas, unas 40 veces inferior al empleado por la herramienta Sipreólico. Para dar mayor utilidad a las predicciones de potencia es necesario que vayan acompañadas de una medida de incertidumbre. A tal fin, se ha propuesto un modelo de predicción

probabilista basado en técnicas no paramétricas de estimación de densidades con funciones Kernel. Se propone una predicción probabilista condicionada al vector de viento previsto en la región, que utiliza las medidas de distancia y las funciones de ponderación utilizadas en el modelo de predicción de potencia. El modelo proporciona la función de distribución del error de predicción, a partir de la cual se puede obtener cualquier percentil o intervalo de predicción deseado.

Se ha comprobado finalmente que la obtención de la función de distribución condicionada a los vectores de viento es un método muy atractivo para el cálculo de predicciones probabilistas, pues proporcionando la misma fiabilidad y agudeza, se consigue aumentar la resolución de los intervalos de predicción obtenidos respecto de otros modelos probabilistas no condicionados o condicionados al nivel de potencia prevista.

2.2 Bases teóricas

2.2.1 Parque eólico

a. Parques eólicos en el mundo

Los tres grandes parques eólicos del mundo se encuentran ubicado el primero en la Republica China y los dos siguientes en los estados unidos, el Gansu Wind Farm con 5.160 MW , ubicada en China, se encuentra en construcción [7].

1) **Gansu Wind Farm**

También llamado **Jiuquan Base Wind Power**, con 5.160 MW construidos, es un grupo de grandes parques eólicos en construcción en el oeste de la provincia de Gansu en China. El Proyecto de Granja Gansu viento, se encuentra en las zonas desérticas cerca de la ciudad de Jiuquan, en dos localidades del Condado Guazhou y también cerca de la ciudad de Yumen, en la provincia noroeste de Gansu, que tiene una abundancia de recursos eólicos.

El proyecto es uno de los seis nacionales de energía eólica megaproyectos aprobados por el gobierno chino. Se espera que crezca a 20.000 megavatios en 2020, a un costo estimado de 120 millones de yuanes chinos (\$ 17,5 mil millones). El proyecto está siendo construido por más de 20 desarrolladores en dos localidades en el condado de Guazhou y también cerca de la ciudad de yumen.

2) **Centro de Energía Eólica Alta:**

El **Centro de Energía Eólica Alta (AWEC, Alta Wind Energy Centre)** situado en Tehachapi, (Condado de Kern), en California, Estados Unidos, es actualmente **el mayor parque eólico en operación del mundo, con una capacidad operativa de 1.020 MW**. El parque eólico terrestre es operado por los ingenieros de Terra-Gen Power, quienes se encuentran inmersos actualmente en una nueva ampliación para incrementar la capacidad del parque eólico a **1.550 MW**.

Las primeras cinco unidades de AWEC fueron terminadas en 2011, instalándose dos unidades adicionales al año siguiente. La primera unidad estaba formada por 100 turbinas GE 1.5-MW SLE, mientras que las otras seis unidades operativas fueron instaladas con turbinas Vestas V90-3.0MW. A partir de 2013 se iniciaron las fases para implementar otras cuatro unidades más a AWEC, siendo la octava y novena unidad integradas por aerogeneradores de Vestas, mientras que las dos últimas unidades serán instaladas con turbinas GE 1.7-MW y GE 2.85-MW de General Electric. Cuando se combinen, las 11 unidades del parque eólico estarán formadas por 586 turbinas en total.

3) Parque Eólico Shepherds Flat:

El **Parque Eólico Shepherds Flat** situado cerca de Arlington, al este de Oregón, en Estados Unidos, es el segundo parque eólico más grande del mundo con una capacidad instalada de **845 MW**. Desarrollado por los ingenieros de Caithness Energy, las instalaciones cubren más de 77 km² entre los condados de Gilliam y Morrow. El proyecto, desarrollado por los ingenieros de Caithness Energy en un área de más de 77 km² entre los condados de Gilliam y Morrow, comenzó a construirse en 2009 con un coste estimado en 2 mil millones de dólares (1,4 mil millones de euros), recibiendo una garantía de préstamo de 1,3 mil millones de dólares del Departamento de Energía de EE.UU. en octubre de 2010, lo que supuso la mayor

financiación jamás llevada a cabo en el mundo para la construcción de un parque eólico.

El parque eólico se encuentra en funcionamiento desde septiembre de 2012, el cual lo integran 338 turbinas GE2.5XL, cada una con una capacidad nominal de 2,5 MW cuya energía producida es suministrada a la Southern California Edison para su distribución. En términos generales, la energía renovable generada por el parque eólico es suficiente como para satisfacer las necesidades eléctricas de más de 235.000 hogares.

b. Parques eólicos en el Perú

En el Perú, existen tres parques eólicos, en proyecto y uno en operación que es la Central Eólica de Marcona

1) Central Eólica de Marcona

El primer parque eólico del Perú, integrada al SEIN, es la Central Eólica de Marcona, su operación lo realiza en coordinación con el COES¹, tiene una capacidad de 32MW, se encuentra ubicado en el departamento de Ica, Provincia de Marcona, Distrito de Marcona a una altitud de 200 msnm, energía anual ofertada 148 378MWh; como información relevante de la central parque eólico de Marcona es la siguiente: [8].

- El proyecto contempla la instalación de 11 Aerogeneradores (8 x 3,15 MW + 3 x 2,3 MW).

¹Comité de Operación Económica del Sistema Interconectado Nacional

- Los 11 Aerogeneradores tendrán una tensión de generación de 0,69 kV. Se agruparán en 3 bloques, interconectados entre sí. Luego, cada uno de estos bloques se conectará mediante 3 circuitos a la S.E. de la Central Eólica en el nivel de 20 kV. En esta S.E. se instalará un transformador 20/220 kV, el cual se conectará a la S.E. Marcona existente.
- La potencia generada se inyectará al SEIN mediante una L.T. de 31 km en 220 kV que conectará la S.E. Central Parque Eólico Marcona con la S.E. Marcona existente.
- El Estudio de Impacto Ambiental fue aprobado en junio 2011. El 13.08.2012, mediante R.S. N° 097-2012-EM, se otorgó la Concesión Definitiva de Generación. El 31.10.2012, mediante R.S. N° 104-2012-EM, se otorgó la Concesión Definitiva de Transmisión.

Sus parámetros meteorológicos para la central eólica de Marcona es la siguiente: [9].

En la tabla (véase la tabla N°2.2, en la página "33"), se muestran los valores mínimos, máximos y promedios de los parámetros meteorológicos registrados durante los días de monitoreo. La estación meteorológica portátil fue instalada en el punto céntrico del parque eólico, en el distrito de San Juan de Marcona.

TABLA N°2.2
RESUMEN DE PARÁMETROS METEOROLÓGICOS

Puntos de registro	Temperatura (°C)			Humedad Relativa (%)			Velocidad del viento (m/s)			Dirección predominante del viento
	Min	Prom	Máx	Min	Prom	Máx	Min	Prom	Máx	
24/06/10	11,6	15,9	18,5	60,9	70,9	78,4	3,7	4,7	5,7	S, SSW, SW
25/06/10	13,9	15,8	18,3	59,5	70,0	79,7	3,6	4,9	6,4	S
26/06/10	14,6	16,5	19,1	59,6	70,3	81,4	4,1	5,4	6,4	SSW
27/06/10	14,9	16,1	18,4	65,8	76,7	82,9	5,2	6,2	7,1	SSW

Elaboración: Walsh Perú, 2010

La temperatura ambiental o del aire es un parámetro que varía con la hora del día y con la estación del año. Durante el periodo de muestreo el comportamiento de la temperatura presentó un valor mínimo de 11.6 °C y un valor máximo de 19.1 °C. En general los valores registrados son típicos de los meses de invierno. Los más altos valores ocurrieron entre las 12:00 y 14:00 horas.

La humedad relativa disminuyó conforme avanzaban las horas de sol y por el aumento de la temperatura en el aire. Luego se incrementó lentamente hasta alcanzar sus valores máximos, hasta 82.9%, en horas de la madrugada. Este fenómeno se explica porque durante el día la temperatura aumenta, provocando que el vapor de agua y la capacidad de retención de humedad del aire disminuyan.

El viento presentó mayor porcentaje de ocurrencia para velocidades entre 3.6 y 5.7 m/s. No se registraron periodos de calma (velocidades menores a 0.6 m/s, según la escala de Beaufort). Asimismo, predominaron vientos provenientes del Sur-suroeste (SSW).

2) Central eólica Cupisnique

La Central Eólica de Cupisnique, en proyecto, tiene una capacidad de 80MW, se encuentra ubicado en el departamento de La Libertad, Provincia de Pacasmayo, Distrito de Cupisnique a una altitud de 20 msnm, tiene un factor de planta de 43%; como información relevante de la central parque eólico de Cupisnique [10] es la siguiente:

- La central se encuentra ubicada en la costa peruana, entre los puertos de Malabrigo y Puémape, en las pampas de Cupisnique.
- La central contempla la instalación de 45 Aerogeneradores de 1,8 MW c/u. (3 en Reserva) dispuestos estratégicamente en 6 Celdas de Transformación de 30 kV.
- Cada aerogenerador está constituido por una turbina y un generador eléctrico situados en lo alto de una torre de acero, cimentada sobre una zapata de concreto armado. Además, en el interior de la torre se incluye un centro de transformación (de baja a media tensión).
- La marca y modelo de los aerogeneradores instalados es VESTAS, modelo V100 de 1,8 MW, el cual posee un rotor a barlovento equipado con tres Palas a 120° entre ellas. Dentro de cada aerogenerador, se encuentran los equipos de control y conexión en baja tensión, el transformador de baja tensión (generación) a media tensión, y las celdas de media tensión.

- La potencia individual de cada aerogenerador es de 1,8 MW. Las 45 máquinas que componen el parque se disponen en seis circuitos independientes, agrupados estos últimos en dos barras colectoras de media tensión independientes, cada una de las cuales corresponde a un transformador de evacuación.
- El EIA se aprobó mediante R.D. N° 008-2011-MEM/AE del 07.01.11.
- La nueva fecha de operación de la central, según el MINEM, es el 20.02.2014. La concesionaria ha solicitado una ampliación de plazo de 192 días adicionales al 20.02.2014 estableciendo el 31.08.2014 como fecha POC debido a la demora en la fabricación de equipos y materiales retenidos por Montealto y a la suspensión de actividades en la SET Cupisnique.

3) Central eólica Talara

La Central Eólica de Talara, por inaugurarse, tiene una capacidad de 30MW, se encuentra ubicado en el departamento de Piura, Provincia de Pariñas, Distrito de Talara, a una altitud de 11 msnm, su energía anual ofertada es 119 673 MWh; como información relevante de la central parque eólico de Talara [11] es la siguiente:

- Contempla la instalación de 17 aerogeneradores, modelo V100 de 1,8 MW c/u; formado por 3 palas de 49 metros de longitud c/u y un ángulo de 120° entre ellas.

- Cada aerogenerador está constituido por una turbina y un generador eléctrico situados en lo alto de una torre de acero, cimentada sobre una zapata de concreto armado. Además, en el interior de la torre se incluye un centro de transformación.
- Cada aerogenerador V100 de 1,8 MW está provisto de un transformador trifásico tipo seco, con refrigeración forzada por aire, ubicado en la parte trasera de la góndola. Sus características principales serán: potencia primario 2 100 kVA, potencia secundario (690 V) 1 900 kVA, potencia secundario (480 V) 200 kVA, frecuencia 60 Hz, tensión primaria 30 kV y tensión secundaria 690/480 V.
- La torre metálica que soporta al aerogenerador es de 80 m.
- Cada aerogenerador ocupa un área de 80 m², distanciados a 1,5 veces en diámetro que forman las palas (150 m) y 3 veces del diámetro en paralelo (300 m), distancia que permite evitar el efecto Estela (Cola de viento turbulento).
- La energía generada se inyectará al SEIN a través de la L.T. 220kV que parte de la S.E. Central Eólica Talara (30/220 kV) y se interconectará a la S.E. Pariñas. El proyecto de interconexión contempla que se instale un sistema de simple barra en la subestación de la central y una celda en la S.E. Pariñas.
- El EIA se aprobó mediante R.D. N° 016-2011-MEM/AE del 19.01.11; asimismo, mediante R.S. N° 033-2011-EM del 26.04.11

se otorgó la Concesión Definitiva de Generación y con R.S. N° 045-2011-EM del 01.06.11 se aprobó la Concesión Definitiva de la L.T.

- Establecido el 31.07.2014 como fecha POC debido a la demora en la fabricación de equipos y materiales retenidos por Montealto,
- Con carta COES/D/DP-344-2014 del 17.03.2014 se aprobó el estudio de operatividad de la C.E. Talara-Etapa 1.

4) Central eólica Tres hermanas

La Central Eólica de Tres hermanas, en proyecto, tiene una capacidad de 90MW, se encuentra ubicado en el departamento de Ica, Provincia de Nazca, Distrito de Marcona a una altitud de 496 msnm, con energía anual de 415 760 MWh; como información relevante de la central parque eólico de Tres hermanas [12] es la siguiente:

- La central consta de un conjunto de 33 aerogeneradores dispuestos en alineaciones distribuidas perpendicularmente a los vientos dominantes en la zona.
- Los circuitos eléctricos de media tensión del parque eólico se proyectan en 20 kV y conectan transformadores de cada turbina con la S.E. de la central.
- El punto determinado para la conexión de la Central Eólica Tres Hermanas es la barra en 220 kV de la S.E. Marcona de REP en el SEIN.

- El proyecto contempla implementar una bahía de salida en el lado de 220 kV de la S.E. C.E. Parque Eólico Marcona para recepcionar la conexión proveniente de la C.E. Tres Hermanas.
- Con R.D. N° 251-2013-MEM/AAE del 28.08.2013 se aprobó el Estudio de Impacto Ambiental.
- la POC del 31.12.2014 al 31.12.2015. Esto se debió a demoras en la aprobación del EIA. El inicio de la obra está previsto para julio del 2014.

c. Componentes de un parque eólico

El parque eólico consta de varios aerogeneradores que integrados inyectan energía y potencia al Sistema Eléctrico. Así mismo estos grupos de aerogeneradores van acompañadas de otros componentes como los siguientes:

- 1) Generadores eólicos, de 2.0 MW de potencia cada uno (30 unidades), haciendo una potencia instalada de 60MW.
 - Rotor.
 - Caja de engranajes.
 - Generadores eléctricos.
 - Sistemas de regulación de potencia y de velocidad.
 - Sistemas de orientación.
 - Sistemas de conexión a red.
 - Sistemas de seguridad.
 - Controladores electrónicos locales.

- Elementos de acoplamiento mecánico.
 - Chasis principal.
 - Torres.
- 2) Transformador de potencia $220 \pm 10 \times 1\% / 34,5 \text{KV}$
 - 3) Transformadores de $34,5 \pm 10 \times 1\% / 0,69 \text{KV}$
 - 4) Líneas de transmisión
 - 5) Equipos de protección (Relés, contactores, interruptores, etc)
 - 6) Centro de mando, supervisión y control.
 - 7) Equipos de control y supervisión (SCADA)

La energía o potencia aprovechable de un aerogenerador, se puede evaluar mediante la densidad, es decir, la potencia del viento por unidad de área de impacto, en las aspas de un aerogenerador.

En consecuencia se dice que la velocidad del viento, influyen directamente en la producción de energía de un parque eólico; la temperatura, presión atmosférica y la humedad tienen relación directa con el viento, el mismo que forma parte del presente trabajo.

d. El Viento

1) Naturaleza del viento

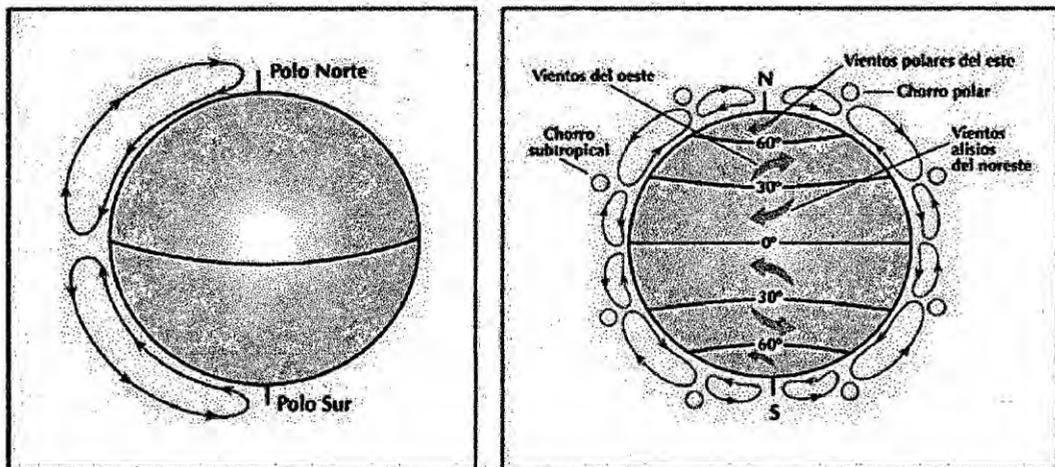
Todas las fuentes de energía renovables (excepto la maremotriz y la geotérmica), incluyendo la energía de los combustibles fósiles, provienen, en último término, del sol. La Tierra recibe $1,74 \times 10^{14} \text{kW}$ de potencia del sol.

Alrededor de un 1 a un 2% de la energía proveniente del sol es convertida en energía eólica.

Esto supone una energía alrededor de 50 a 100 veces superior a la convertida en biomasa por todas las plantas de la tierra.

El fenómeno conocido como viento está constituido por las corrientes de aire generadas a consecuencia del desigual calentamiento de la superficie de la tierra. La no uniformidad del flujo de radiación solar incidente hace que unas zonas se calienten más que otras, provocando movimientos convectivos de la masa atmosférica. El aire caliente asciende, arrastrando aire más frío proveniente de una región vecina. Al subir se enfría, por lo que aumenta su densidad, descendiendo para volver a repetir el ciclo. [13].

FIGURA N° 2.1
VARIACIÓN DEL VIENTO

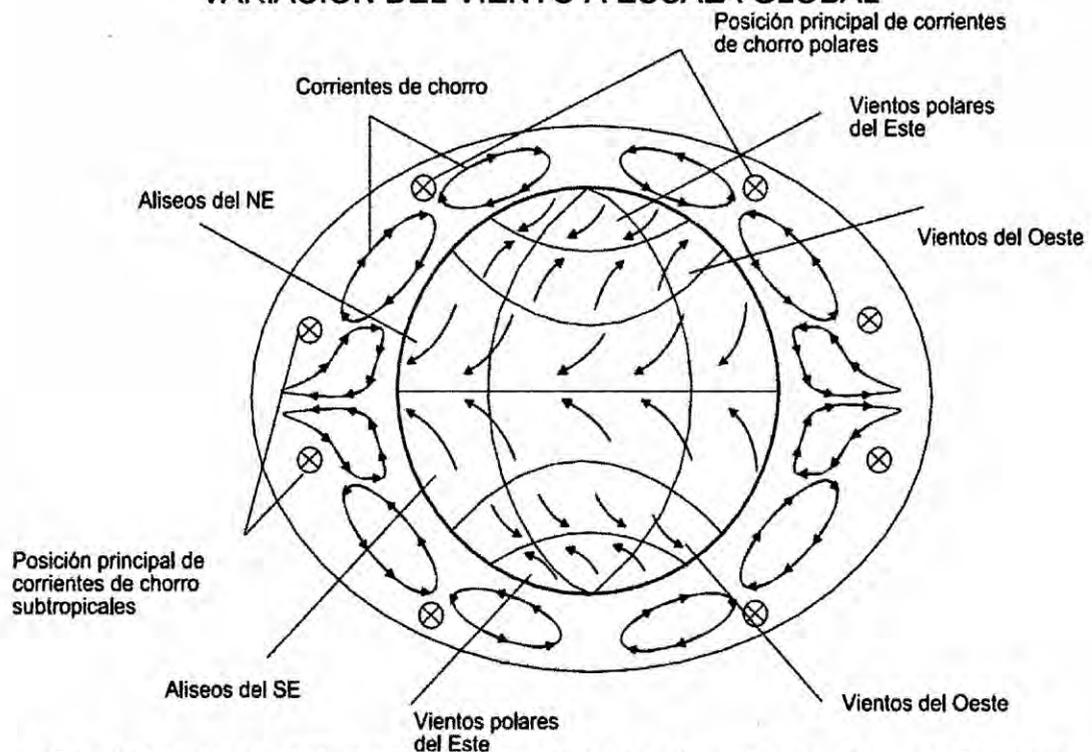


A altitudes de hasta 100 metros sobre la superficie terrestre, los vientos están muy influenciados por las características de dicha superficie. El viento es frenado por la rugosidad de la superficie de la tierra y por los

obstáculos. Tratándose de energía eólica interesará conocer estos vientos de superficie y cómo calcular la energía aprovechable del viento.

Aunque los vientos globales son importantes en la determinación de los vientos dominantes de un área determinada, las condiciones climáticas locales pueden influir en las direcciones de viento más comunes. Los vientos locales siempre se superponen en los sistemas eólicos a gran escala, esto es, la dirección del viento es influenciada por la suma de los efectos global y local. Cuando los vientos a gran escala son suaves, los vientos locales pueden dominar los regímenes de viento. [14]

**FIGURA N° 2.2
VARIACIÓN DEL VIENTO A ESCALA GLOBAL**



En las cinco estaciones, se realizó el registro de parámetros meteorológicos como velocidad y dirección del viento, temperatura; con

respecto a la humedad relativa y presión atmosférica, los datos son del SENAMHI¹. Estos datos caracterizan las condiciones atmosféricas durante el muestreo de calidad del aire.

La instrumentación es instalada en la torre del comunicaciones en el Cerro Chocan – Piura - Perú, se colocaron dos instrumentos de la velocidad del viento en el nivel de 58m, dos instrumentos de la velocidad del viento en el nivel de 49m y un instrumento de velocidad del viento en el nivel de 24m. Las velocidades promedio del viento cada hora por el periodo entero de registro (Noviembre 2007 – Febrero 2009).

Véase la tabla N°2.3, muestra los parámetros de los sensores meteorológicos que conformaron la estación portátil.

**TABLA N° 2.3
CARACTERÍSTICAS DE LOS SENSORES
METEOROLÓGICOS**

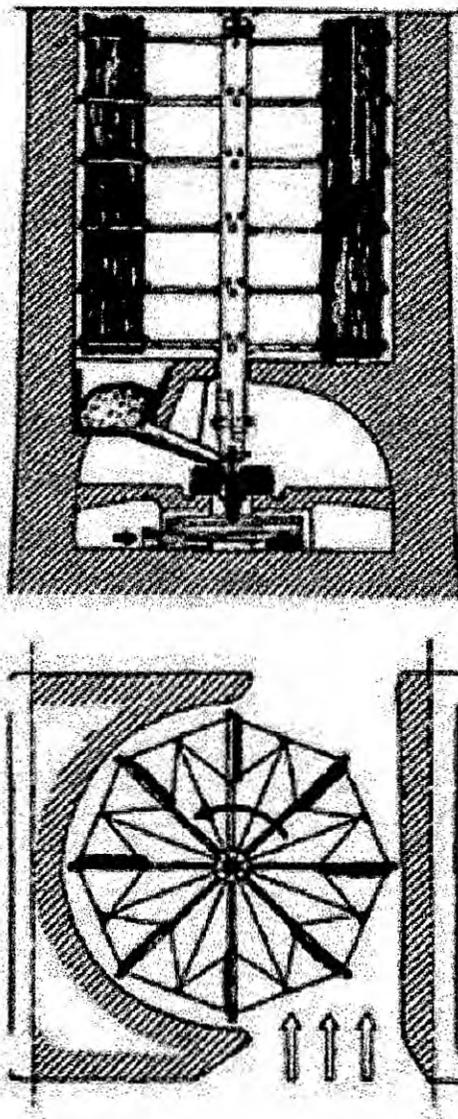
Variable	Unidades	Estaciones
Velocidad del viento	m/s	Particular
Dirección del viento	-	Particular
Temperatura	°C	Particular
Presión Atmosférica	Mbar	SENAMHI
Humedad relativa	%	SENAMHI

2) Historia de la utilización del Viento

La primera y más sencilla aplicación de la energía del viento corresponde al uso de las velas en la navegación. La referencia más antigua se encuentra en un grabado egipcio que data del III milenio a.C. Las peculiares características de los vientos, su comportamiento

irregular tanto en intensidad como en dirección, exigen para su aprovechamiento una tecnología capaz de desarrollar los mecanismos de regulación y orientación, más complejos que los de los sistemas hidráulicos. [14]

FIGURA 2.3
MOLINO PERSA



Los primeros ingenios eólicos debieron desarrollarse en la antigua Persia, cuyos territorios eran muy ventosos y donde se reunían los

más avanzados conocimientos técnicos del momento, tanto del extremo oriente como de la cultura helénica. Primeramente aparecieron los molinos de eje vertical: Un número determinado de velas montadas verticalmente unidas a un eje y empujadas por el aire reemplazaron el accionamiento animal para proporcionar un movimiento giratorio. Una ilustración de este antiguo molino se muestra (véase la figura 2.3, en la página "43"). Desgraciadamente no se conservan detalles de molinos persas, pero sí es bien conocido el hecho de que a mediados del siglo VII a.C. los molinos de viento eran máquinas bien conocidas en esa parte del mundo, aunque se tratara de diseños bastos y mecánicamente ineficientes.

Los chinos utilizaban desde tiempos inmemoriales los molinos de viento llamados panémonas, que se usaban para bombear agua en las salinas, e incluso algunos historiadores apuntan hacia la posibilidad de que pudieron ser los precursores de los molinos persas. Las panémonas eran también de eje vertical y sus palas estaban construidas a base de telas sujetas a largueros de madera. La posición de las palas podía variarse para regular la acción del viento sobre el molino.

Hay una gran distancia entre los escasos datos acerca de los molinos persas y aquellos llamados molinos europeos con velas montadas en un eje horizontal, alguno de los cuales se utiliza hasta el día de hoy. Los primeros molinos de eje horizontal tenían una serie de lonas

dispuestas a lo largo de una estructura de madera que se debía de orientar hacia el viento incidente. Esta orientación se conseguía haciendo girar el rotor sobre el poste donde se suspendía, guiado por una veleta.

Las primeras referencias de molinos de viento en España datan de la Edad Media. En un principio aparecieron en zonas cristianas y posteriormente incrementaron su presencia. Poco después fueron realizadas variaciones en los diseños dependiendo de la zona y del uso al que se destinaban. Del período comprendido entre el siglo XVI y el siglo XIX quedan bastantes restos de sistemas eólicos en España: molinos manchegos y andaluces empleados para moler cereales, molinos cartageneros empleados en irrigación y molinos baleares empleados tanto para molienda como para irrigación.

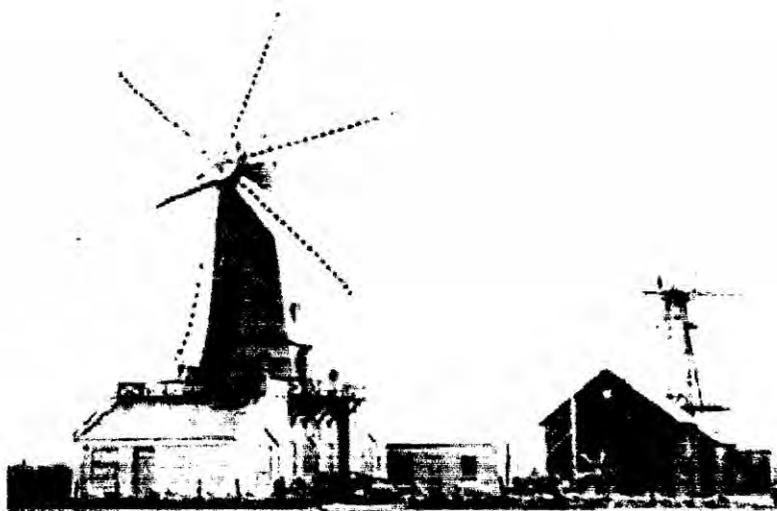
El siglo XVII es un siglo de grandes avances científicos y tecnológicos. Sin embargo, a principios del mismo, los principios teóricos de los molinos son todavía apenas conocidos. Las innovaciones no alteraron el formato exterior de los molinos, que se mantuvo sin demasiadas modificaciones, pero en cambio mejoraron los detalles de diseño y construcción apareciendo los sistemas mecánicos de orientación y regulación.

Las palas de los molinos anteriores al siglo XVII se construían con un entramado de varillas a ambos lados de un mástil principal, cubriéndose posteriormente con una tela. Más tarde el mástil se

colocó en el borde de ataque de la pala, de forma que soportara mejor la entrada de aire. Este sistema era también el más adecuado para dotar de cierta torsión a la pala a lo largo de la envergadura, con el fin de mejorar su rendimiento aerodinámico. Las palas con torsión se desarrollaron en el siglo XVII y la incorporación de los sistemas de regulación se llevó a cabo en el siglo siguiente.

Los molinos de viento evolucionaron en su desarrollo hasta mediados del siglo XIX, introduciéndose continuas mejoras tecnológicas a partir de elementos mecánicos. Las primeras bombas eólicas aparecen hacia 1854, desarrolladas por Daniel Halladay. Son rotores multipalas acoplados mediante un sistema biela-manivela a una bomba de pistón. Hacia 1890 se empieza a fabricar, con álabes metálicos, el conocido molino de bombeo americano, llegando a convertirse en el molino de viento más extendido de cuantos hayan existido.

FIGURA 2.4
MOLINO PARA BOMBEO



En España es interesante destacar la masiva presencia de estas máquinas multipala en la isla de Fuerteventura, con importantes modificaciones respecto de las utilizadas anteriormente importadas de los Estados Unidos.

Fue en 1892, cuando el profesor La Cour diseñó el primer prototipo de aerogenerador eléctrico. Los trabajos de La Cour constituyeron los primeros pasos en el campo de los aerogeneradores modernos.

Poul La Cour (1846-1908), inicialmente meteorólogo, fue el pionero de las modernas turbinas eólicas generadoras de electricidad. Construyó la primera turbina eólica generadora de electricidad del mundo en 1891. Poul La Cour realizó sus experimentos de la aerodinámica de las palas en un túnel de viento construido por él mismo. Uno de los primeros lugares donde utilizó la producción de electricidad de sus máquinas eólicas fue en su propia escuela. Utilizaba dicha electricidad para producir electrólisis y obtener así hidrógeno para las lámparas de gas de su escuela. El principal inconveniente que esto suponía es que tuvo que sustituir las ventanas de varios de los edificios de la escuela en diversas ocasiones, pues el hidrógeno explosionaba debido a las pequeñas cantidades de oxígeno que había en él.

Los primeros aerogeneradores de corriente alterna surgieron en los años 50 de la mano del ingeniero Johannes Juul. Concretamente en 1956 se desarrolló el aerogenerador de Gedser (Dinamarca) de 200 kW que representa la antesala de los actuales aerogeneradores.

Otra máquina eólica de trascendencia fue la construida por el profesor Utrich Hutter en 1960 con una potencia de 100 kW y un diámetro de 34 m. Estas máquinas representan los comienzos de diferentes facetas en el desarrollo de la energía eólica: máquinas con potencias del orden de los megavatios las cuales están empezando a aparecer en la escena comercial y los diseños caracterizados por estructuras más livianas que todavía representan el futuro de las máquinas eólicas.

Después de la primera crisis del petróleo de 1973, muchos países despertaron su interés en la energía eólica. En un principio las compañías de energía dirigieron inmediatamente su atención a la construcción de grandes aerogeneradores, tomando como punto de partida el aerogenerador de Gedser. En 1979 construyeron dos aerogeneradores de 630 kW. Estos diseños resultaron extremadamente caros y, en consecuencia, el alto precio de la energía devino un argumento clave en contra de la energía eólica.

En España, el desarrollo actual de la energía eólica comienza en los años 1978-1979, partiendo la iniciativa del Ministerio de Industria, Comercio y Turismo, dando lugar a la instalación de un prototipo de 100 kW de potencia, situado en Tarifa (Cádiz). Siguiendo la tendencia general de la energía eólica, en el periodo comprendido entre 1981 y 1986 se desarrollaron fundamentalmente máquinas de pequeña y mediana potencia que permitían dar un contenido tecnológico básico

a la naciente industria del sector. Es en 1981 cuando surgen las primeras máquinas de pequeña potencia.

Durante este periodo la mayor implantación de sistemas eólicos se produjo en EEUU. Miles de máquinas fueron instaladas en el programa eólico de California a principios de los 80. La Micon de 55 kW (máquina eólica de origen danés) es un ejemplo de tales máquinas, instalada en un enorme parque eólico de más de 1.000 máquinas en Palm Springs (California).

En España, el Plan de Energías Renovables de 1986 puso en práctica una serie de medidas que facilitaron la introducción de energías a gran escala como la eólica. De esta manera, la iniciativa pública hizo posible la instalación de los primeros parques eólicos de pequeña dimensión (oscilando entre 300 y 600 kW de potencia instalada a partir de máquinas de 30 kW), donde el Ministerio de Industria Comercio y Turismo, a través del IDAE¹ y convenios con comunidades autónomas, compañías eléctricas y empresas privadas fueron los promotores y ejecutores de los mismos. De esta forma se pusieron en marcha los parques eólicos de Granadilla (Tenerife), La Muela (Zaragoza), Estaca de Bares (La Coruña), Ontalafia (Albacete) y Tarifa (Cádiz). A pesar de estos proyectos, a finales de 1988 la participación de la energía eólica en el contexto de las energías renovables era todavía muy baja.

¹Instituto para la Diversificación y Ahorro de la Energía

Fue en 1989, con el segundo Plan de Energías Renovables, cuando la energía eólica adquirió una mayor presencia. En esta etapa se realizó la instalación en Cabo Villano (La Coruña) y en el parque eólico de Monteahumada, en Tarifa (Cádiz), de un aerogenerador de 1200 kW, a finales de 1990, en el cuarto lugar de Europa en cuanto a potencia instalada. Es en esta fecha cuando se apuesta decididamente por la implantación en España de parques eólicos, conectados a la red de distribución eléctrica en aquellas zonas geográficas nacionales con alto potencial (Canarias, Andalucía, Galicia), utilizando máquinas de potencias superiores a 100 kW.

El desarrollo creciente de las energías renovables y en particular de la energía eólica observado a lo largo de las últimas décadas y previsto para los próximos años para España en general, se basa en un continuo esfuerzo de investigación e innovación. En este marco, la administración juega un papel importante con la creación de infraestructuras de investigación y desarrollo (I+DT). Esta I+DT resulta imprescindible para dotar al tejido empresarial de recursos y capacidades necesarias para afrontar con éxito la adaptación a una dinámica de crecimiento que demanda altos niveles de actividad innovadora, en un mercado cada día más abierto y competitivo.

En un sentido amplio, el lanzamiento de nuevos productos o de mayor fiabilidad y eficiencia se inicia con un proceso de innovación tecnológico, seguido por una fase de demostración y difusión,

alcanzando la madurez comercial al lograr un volumen suficiente de producción. Las energías renovables, y en particular la energía eólica, se encuentran en diferentes estados de esta cadena, que por el bien común deben consolidarse hasta llegar a constituir recursos plenamente desarrollados. Gracias a este apoyo, en el año 2002 España ocupaba ya el segundo puesto mundial en producción de energía eólica (detrás de Alemania).

Actualmente en España, las líneas prioritarias de actuación en este campo están orientadas fundamentalmente, por un lado a disminuir los costes del kW instalado y los asociados al mantenimiento de las instalaciones, y por otro a impulsar el desarrollo de máquinas en la gama del megavatio.[14]

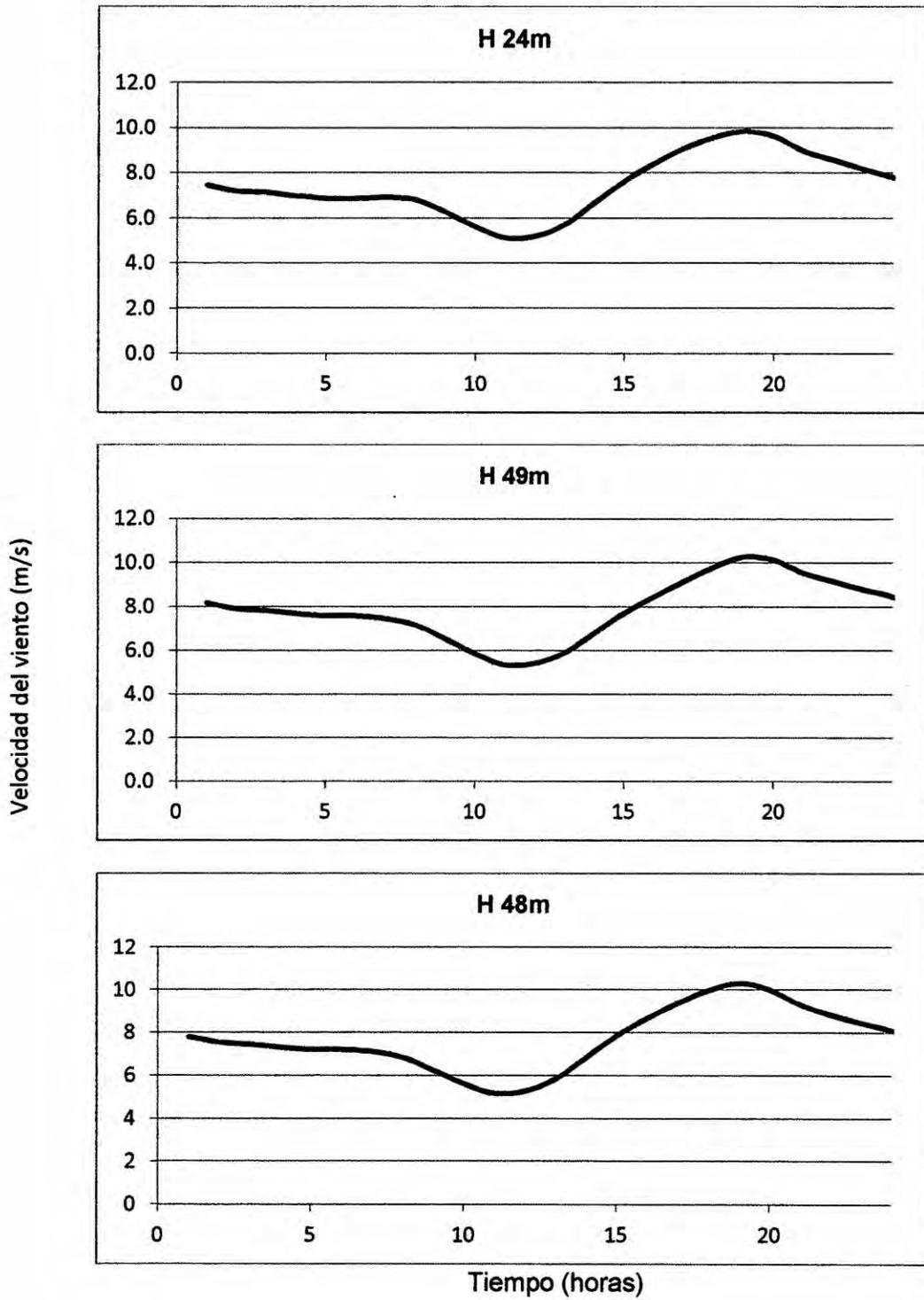
3) Característica del viento

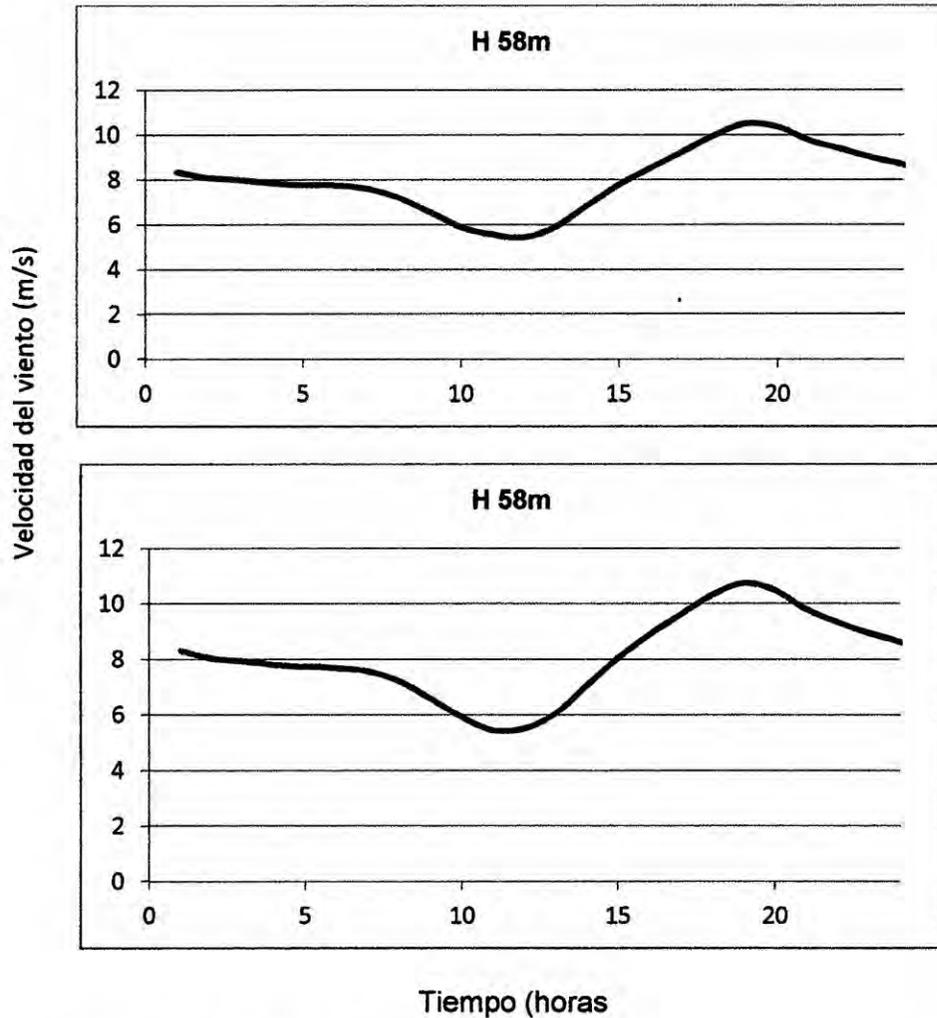
La velocidad del viento varía en el tiempo y el espacio. Las variaciones espaciales son generalmente depende de la altura sobre el suelo y las condiciones geográficas locales y globales. Después de que los datos de recurso eólico, como la velocidad del viento, dirección y temperatura son conocidos en un sitio en particular, entonces los siguientes parámetros se determinan utilizando la información anterior.

a) Velocidad del viento media anual

Esta es la media de la velocidad del viento registrado durante Noviembre 2007 – Febrero 2009, a 24m, 48 m, 49m y 58m de altura en el cerro Chocan-Piura-Perú.

GRAFICO N° 2.1
VARIACIÓN EN EL DÍA DE LA VELOCIDAD ANUAL DEL VIENTO PARA
DIFERENTES ALTURAS DE LAS ESTACIONES





b) Velocidad del viento diurno anual

Esto muestra cómo la velocidad del viento varía durante un día, considerando los datos de velocidad del viento de todo el año. Los diferentes datos de la velocidad del viento por hora es un promedio para el año y se dibuja (véase el gráfico N°2.1, en la página "52"). Variación diaria en la velocidad del viento es debido al calentamiento diferencial de la superficie de la tierra durante el ciclo de radiación diaria.

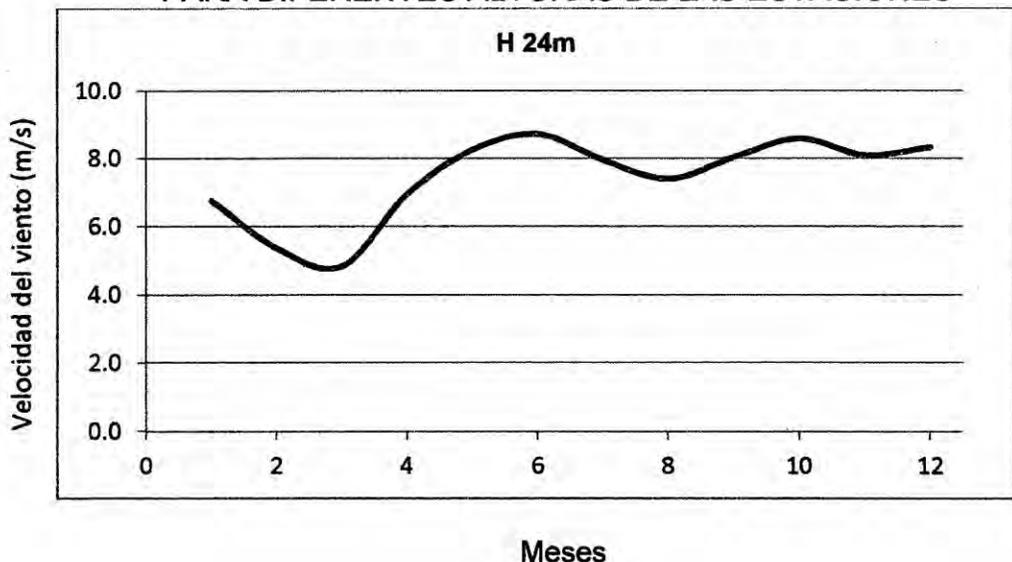
c) Velocidad del viento diurno mensual

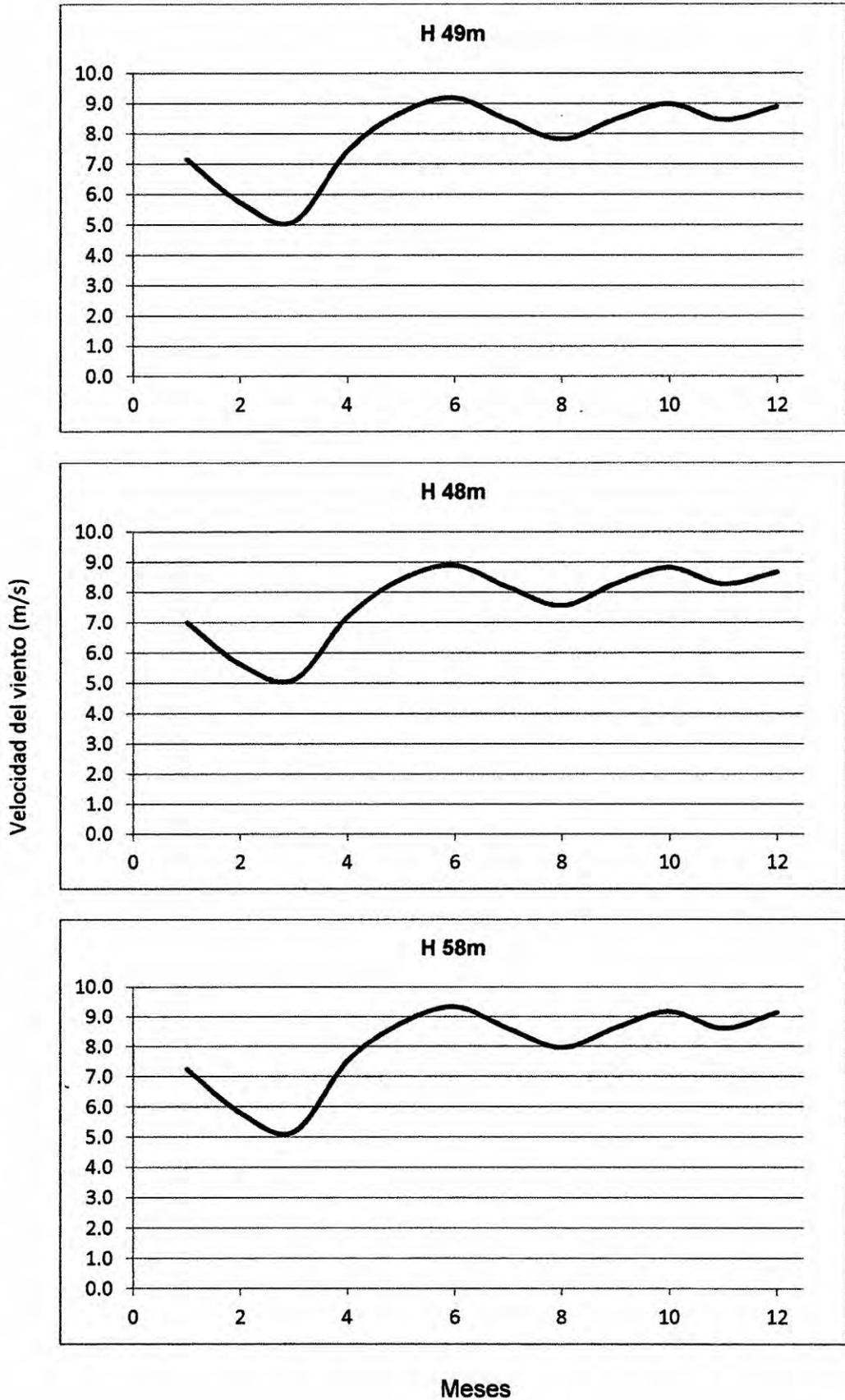
Esto muestra la variación diaria de la velocidad del viento en diferentes meses a lo largo de un año. Esto será similar a la grafico (véase el grafico N°2.1, en la página"52"), pero en lugar de la gráfica durante todo un año, sería la variación diurna de la velocidad del viento para diferentes meses. Esta información ayudará al planificador de viento para que coincida con la carga de todos los días en diferentes meses y el recurso eólico de cada día.

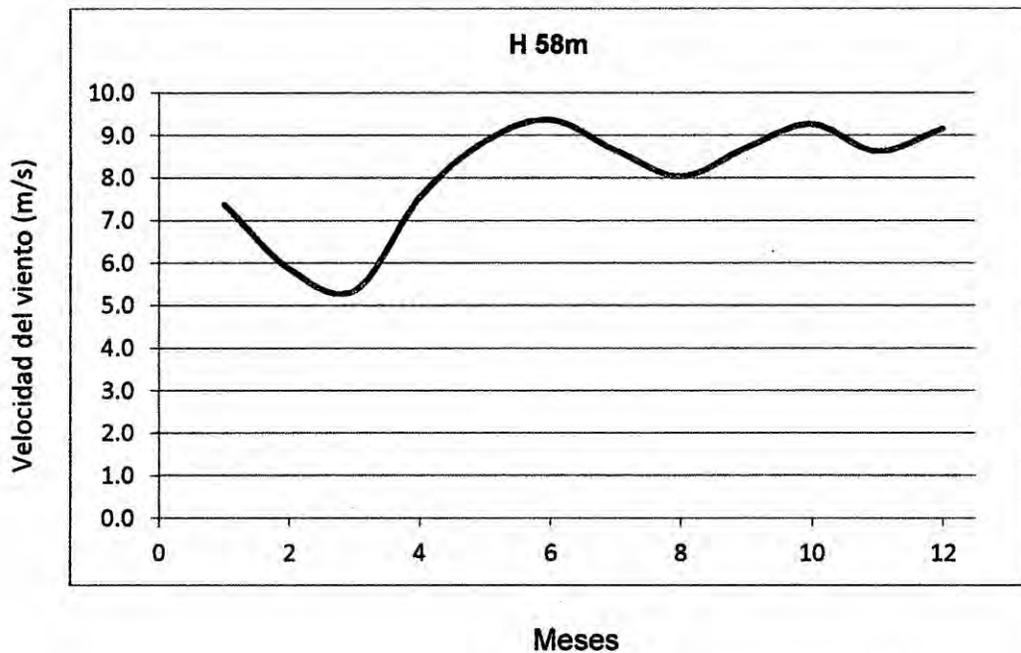
d) Variación mensual de la velocidad del viento

Esto muestra cómo las velocidades del viento están cambiando en diferentes meses en un año, (véase el grafico N°2.2). Nos ayuda a saber qué meses tienen altas velocidades del viento y que tiene bajo, por consiguiente, ayudar a decidir el tipo de sistema de almacenamiento que sería necesaria.

**GRAFICO N°2.2
VARIACIÓN MENSUAL DE LA VELOCIDAD DEL VIENTO
PARA DIFERENTES ALTURAS DE LAS ESTACIONES**







e) Distribución de frecuencia de la velocidad del viento

Esta es la más importante de todas las características del viento. Con esta información la energía eólica anual de un sitio es estimada. Para obtener este gráfico, se agrupa los datos de la velocidad del viento.

Es decir, contenedores de 0-1 m / s, 1-2 m / s, 2-3 m / s, etc, se grafica la velocidad del viento en el año, en este caso se tiene 24 datos al día por 365 días al año haciendo un total de 8760 datos y se han agrupado en los respectivos contenedores de cada rango. A continuación se determina la frecuencia. Esto se puede hacer en Microsoft Excel, donde esto se hace de forma automática y se obtiene la curva de frecuencia de la velocidad del viento.

Se muestra la velocidad del viento y la frecuencia (véase la tabla 2.4), el grafico N°2.3 y el grafico 2.4 se obtiene utilizando la tabla (véase la tabla N°2.4)

GRAFICO N°2.3
FRECUENCIA DEL VIENTO ANUAL PARA UNA ALTURA DE 24m

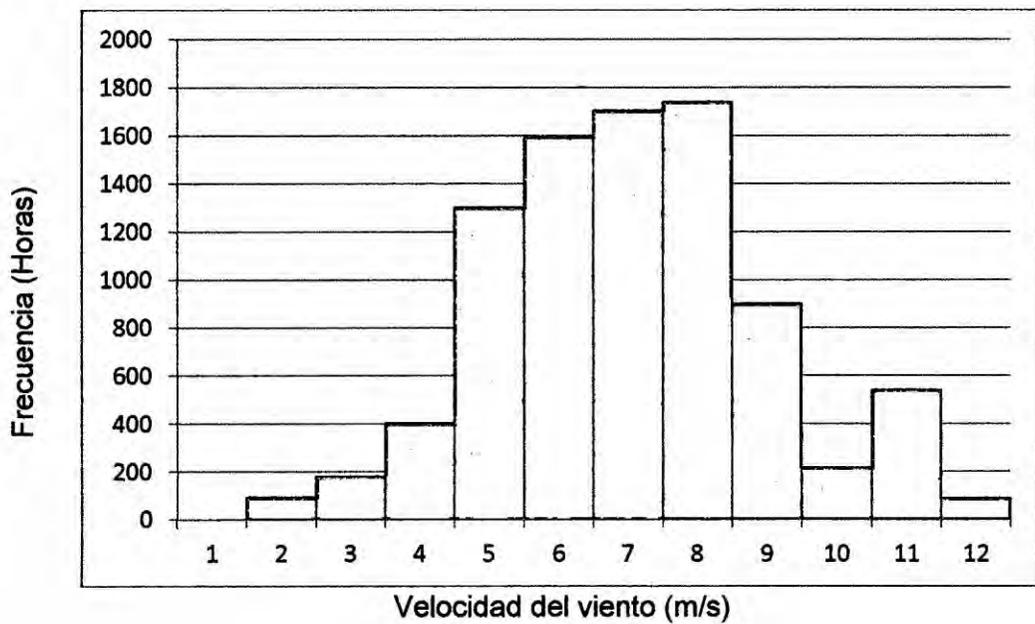


TABLA N°2.4
VELOCIDAD DEL VIENTO Y LA FRECUENCIA

Velocidad del Viento(m/s)		Frecuencia (horas)
0.0	1.0	0
1.0	2.0	93
2.0	3.0	180
3.0	4.0	403
4.0	5.0	1302
5.0	6.0	1590
6.0	7.0	1705
7.0	8.0	1740
8.0	9.0	900
9.0	10.0	217
10.0	11.0	540
11.0	12.0	90

GRAFICO N° 2.4
PROBABILIDAD DEL VIENTO ANUAL PARA UNA ALTURA DE 24m

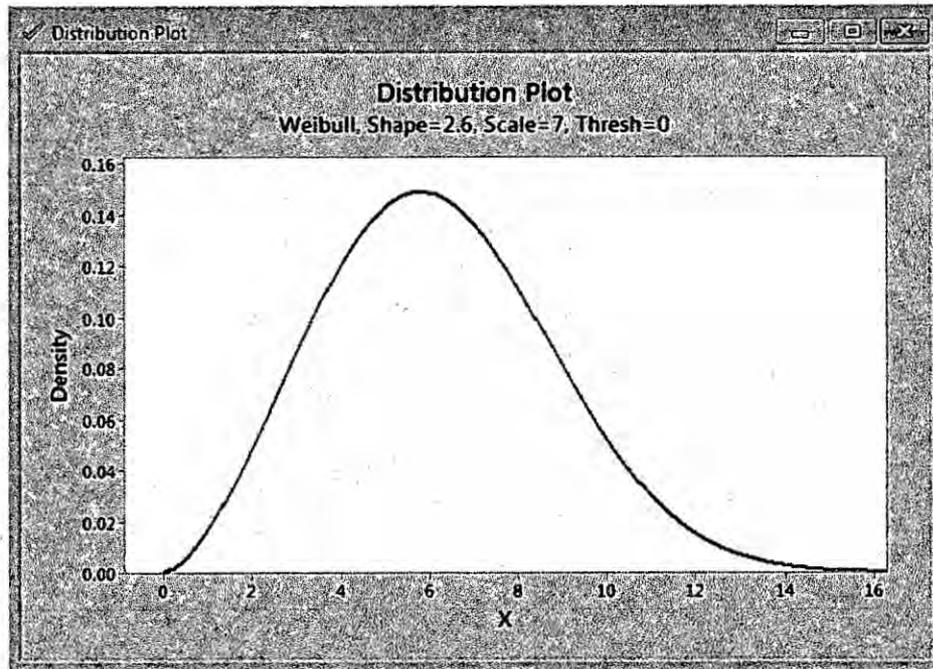
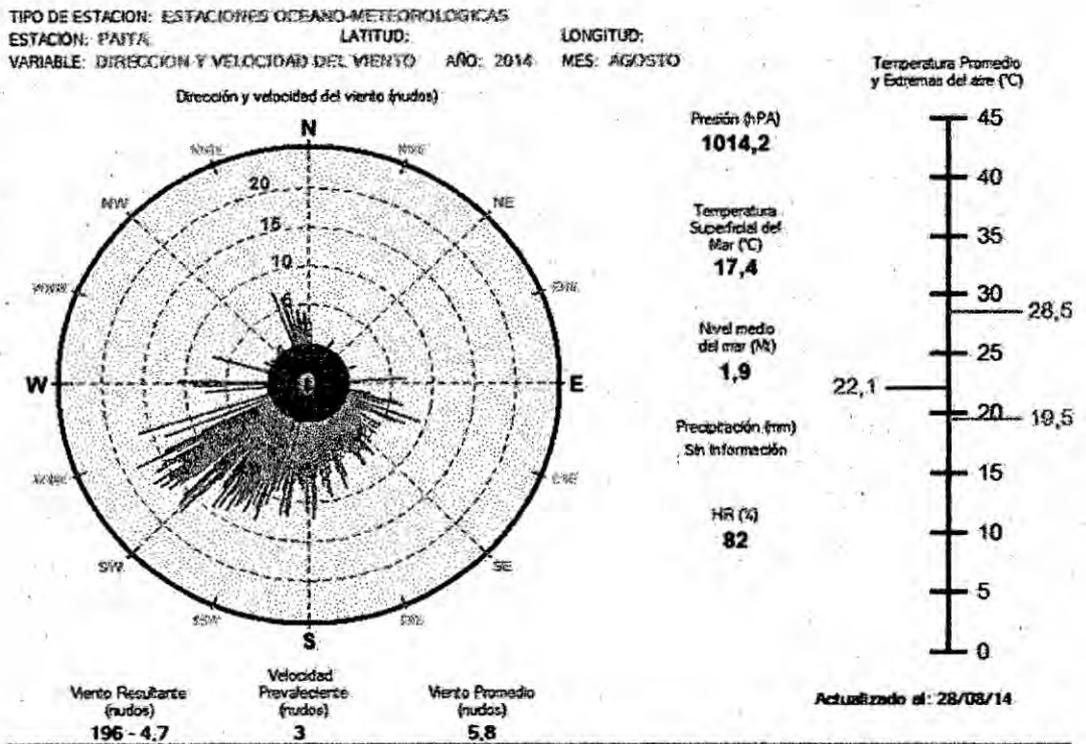


FIGURA N°2.5
ROSA DE VIENTOS



f) Dirección del viento

La Figura (véase la figura N°2.5, en la página "58") muestra la rosa de vientos, registrada durante el mes de agosto en la Estación Paita, los datos de esta estación son procesadas por el proyecto NAYLAMP, de la Dirección de Hidrografía y Navegación de la Marina de Guerra del Perú. Sus datos son medidos a nivel del mar. [15]

4) Tipos de viento

a) Brisas marinas

Durante el día la tierra se calienta más rápidamente que el mar por efecto del sol. El aire sube, circula hacia el mar, y crea una depresión a nivel del suelo que atrae el aire frío del mar. Esto es lo que se llama brisa marina. A menudo hay un periodo de calma al anochecer, cuando las temperaturas del suelo y del mar se igualan. Durante la noche los vientos soplan en sentido contrario. Normalmente durante la noche la brisa terrestre tiene velocidades inferiores, debido a que la diferencia de temperaturas entre la tierra y el mar es más pequeña. [13]

b) Vientos de montaña

Las regiones montañosas muestran modelos de clima muy interesantes. Un ejemplo es el viento del valle que se origina en las laderas que dan al sur (o en las que dan al norte en el hemisferio sur). Cuando las laderas y el aire próximo a ellas están calientes la

densidad del aire disminuye, y el aire asciende hasta la cima siguiendo la superficie de la ladera. Durante la noche la dirección del viento se invierte, convirtiéndose en un viento que fluye ladera abajo. Si el fondo del valle está inclinado, el aire puede ascender y descender por el valle. Este efecto es conocido como viento de cañón. Los vientos que soplan en las laderas a sotavento pueden ser bastante potentes.

c) Variaciones diurnas (noche y día) del viento

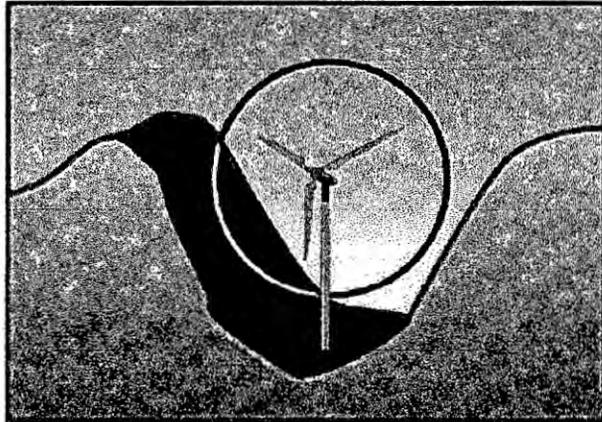
En la mayoría de las localizaciones del planeta el viento sopla más fuerte durante el día que durante la noche. El viento presenta también más turbulencias y tiende a cambiar de dirección más rápidamente durante el día que durante la noche. Desde el punto de vista de los propietarios de aerogeneradores, el hecho de que la mayor parte de la energía eólica se produzca durante el día es una ventaja, ya que el consumo de energía entonces es mayor que durante la noche. Muchas compañías eléctricas pagan más por la electricidad producida durante las horas en las que hay picos de carga (cuando hay una falta de capacidad generadora barata).

d) Efecto túnel

Si tomamos un camino entre dos edificios altos o en un paso estrecho entre montañas observaremos que el aire al pasar a su través se comprime en la parte de los edificios o de la montaña que está expuesta al viento, y su velocidad crece considerablemente

entre los obstáculos del viento. Esto es lo que se conoce como efecto túnel.

FIGURA N°2.6.
EFECTO TÚNEL



Así pues, incluso si la velocidad normal del viento en un terreno abierto es de 6 metros por segundo, en un "túnel" natural, ésta puede fácilmente alcanzar los 9 metros por segundo. Situar un aerogenerador en un túnel de este tipo es una forma inteligente de obtener velocidades del viento superiores a las de las áreas colindantes. Para obtener un buen efecto túnel, debe estar suavemente enclavado en el paisaje. En el caso de que las colinas sean muy accidentadas, puede haber muchas turbulencias en esa área, que pueden causar roturas y desgastes innecesarios en el aerogenerador.

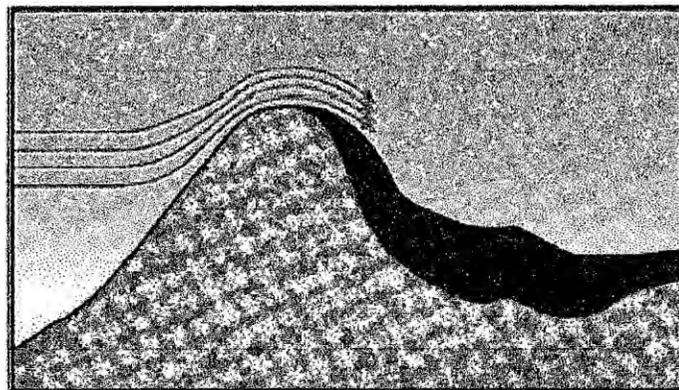
e) Efecto de la colina

Una forma corriente de emplazar aerogeneradores es situándolos en colinas o estribaciones dominando el paisaje circundante. En particular, siempre supone una ventaja tener una vista lo más

amplia posible en la dirección del viento dominante en el área. En las colinas, siempre se aprecian velocidades de viento superiores a las de las áreas circundantes.

Tal y como se puede observar (véase la figura N°2.7), el viento empieza a inclinarse algún tiempo antes de alcanzar la colina. También se aprecia que el viento se hace muy irregular una vez pasa a través del rotor del aerogenerador. Al igual que ocurría anteriormente, si la colina es escarpada o tiene una superficie accidentada, puede haber una cantidad de turbulencias significativa, que puede anular la ventaja que supone tener unas velocidades de viento mayores.

FIGURA N°2.7.
EFECTO COLINA



De este modo, el viento, al considerarlo como recurso energético y desde el punto de vista de su disponibilidad como suministro, tiene sus características específicas: es una fuente con sustanciales

variaciones temporales, a pequeña y gran escala de tiempo, y espaciales, tanto en superficie como en altura, sin olvidar una componente aleatoria que afecta en gran parte a su variación total. Al mismo tiempo hay que considerar que la energía disponible a partir del viento depende de la velocidad del mismo al cubo, por lo que pequeñas variaciones en este parámetro afectarán en gran medida al resultado final de energía obtenida.

La energía cinética de un cuerpo en movimiento es proporcional a su masa (o peso). Así, la energía cinética del viento depende de la densidad del aire, es decir, de su masa por unidad de volumen.

En otras palabras, cuanto "más pesado" sea el aire más energía recibirá la turbina.

A presión atmosférica normal y a 15° C el aire pesa unos 1,225 kilogramos por metro cúbico, aunque la densidad disminuye ligeramente con el aumento de la humedad.

Además, el aire es más denso cuando hace frío que cuando hace calor. A grandes altitudes (en las montañas) la presión del aire es más baja y el aire es menos denso.

f) Variación vertical del viento.

Las estaciones anemométricas en el Cerro Chocan-Piura-Perú donde se efectuó las medidas a 24, 49 y 58 metros de altura, mientras que el eje de los aerogeneradores normalmente se

encuentra a una altura superior. Dado que el viento aumenta con la altura se realizó la corrección a la altura requerida.

Uno de los modelos más empleados para realizar la corrección del viento en altura es el potencial modificado propuesto por Mikhail y Justus [16]. Viene dado por:

$$V_h = V_a (H_h / H_a)^\alpha \quad (2.1)$$

Siendo:

V_h = velocidad del viento a calcular (m/s).

V_a = velocidad conocida del viento (m/s).

H_h = altura a la que se quieren calcular los datos de viento (m).

H_a = altura a la que se han obtenido los datos de viento (m).

El exponente α se calcula de la siguiente manera:

$$\alpha = \frac{1}{\ln\left(\frac{H_g}{H_o}\right)} - \frac{0.0881}{1 - 0.0881 \cdot \ln\frac{H_a}{10}} \cdot \ln\left(\frac{\bar{V}_a}{6}\right) \quad (2.2)$$

Siendo:

H_g = altura media geométrica entre las distintas alturas.

$$H_g = \sqrt{H_a \cdot H_h} \quad (2.3)$$

H_o = longitud de la rugosidad estimada del terreno (m).

Se ha adoptado una rugosidad de 0,01m, que corresponde a un terreno llano de los posibles terrenos circundantes al emplazamiento de los aerogeneradores del parque.

El método seguido tradicionalmente es suponer atmósfera neutra y hacer una estimación de la longitud de rugosidad, y a partir de las medidas de viento a un nivel z , estimar los resultados a otra altura z' , con la expresión:

$$V(z)/V(z') = (Lnz' - Lnz_0) / (Lnz - Lnz_0) \quad (2.4)$$

Las estaciones del SENAMHI cercanas a la zona de estudio son las siguientes:

- Estación Paita-000204, se encuentra ubicado en el departamento de Piura, provincia de Paita, distrito de Paita, en la latitud $5^{\circ} 7' 1''$ y en la longitud $81^{\circ} 8' 1''$.
- Estación Montegrando - 000249, se encuentra ubicado en el departamento de Piura, provincia de Piura, distrito La arena, en la latitud $5^{\circ} 21' 11''$ y en la longitud $80^{\circ} 42' 1''$.
- Estación La Esperanza-000230, se encuentra ubicado en el departamento de Piura, provincia de Paita, distrito de Colan, en la latitud $4^{\circ} 55' 55''$ y en la longitud $81^{\circ} 4' 4''$.

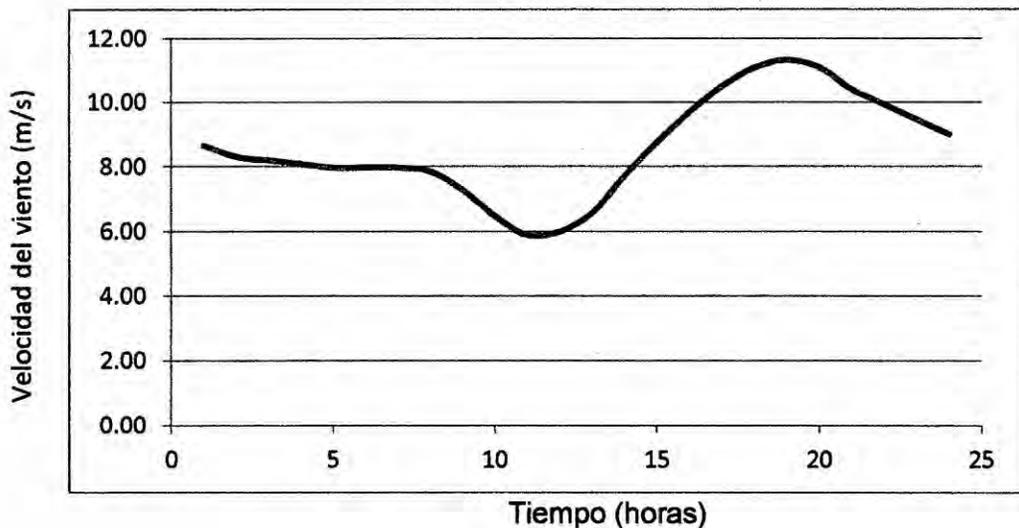
- Estación Bernal-152100, se encuentra ubicado en el departamento de Piura, provincia de Sechura, distrito de Bernal, en la latitud 5 29' 29'' y en la longitud 80 45' 45''.

La zona de estudio se encuentra ubicado en el cerro Chocan, aproximadamente tiene la siguiente coordenada, latitud 5 11' 52'' y en la longitud 80 54' 36''.

La estación Paita y la estación Montegrande se encuentran distantes y al trazar una línea imaginaria entre ambas estaciones pasa por la zona de estudio; lo mismo con la estación La Esperanza y la estación Bernal la línea imaginaria pasa por la zona de estudio. Según anexo N°3.

De la ecuación (1), se halló la velocidad del viento para una altura de 80m que corresponde a la altura del aerogenerador VESTA V90 de 2MW, (véase la gráfica N° 2.5).

GRAFICO N° 2.5
VARIACIÓN EN EL DÍA DE LA VELOCIDAD MEDIA ANUAL
DEL VIENTO A 80m



5) Clasificación del viento

El viento se clasifica de acuerdo a su velocidad desde menos de 1 Km/H hasta más de 117Km/H, que corresponden a viento calmo y huracán respectivamente. [17]

En el caso de la zona de estudio corresponde al promedio de brisa moderada a brisa fresca. (Véase la tabla N°2.5).

**TABLA N° 2.5
CLASIFICACION DEL VIENTO**

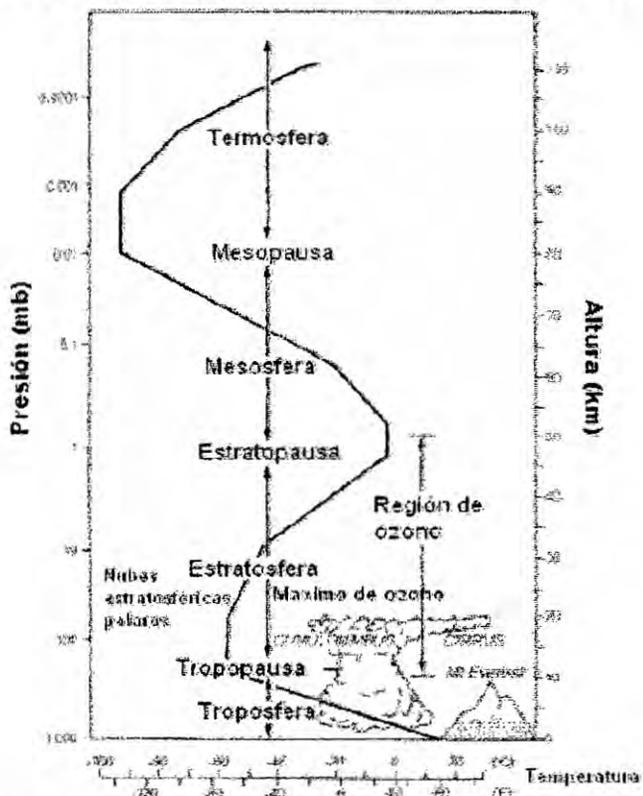
Velocidad del Viento(Km/Hr)	Descripción Internacional
<1	Calmo
1-5	Brisa suave
6-11	Brisa suave
12-19	Brisa moderada
20-28	Brisa moderada
29-38	Brisa fresca
39-49	Ventarrones
50-61	Brisca
62-74	Ventarrón
75-88	Ventarrón
89-102	Ventarrón constante
103-117	Tormenta
>117	Huracán

e, Temperatura del aire, estructura térmica de la atmosfera terrestre.

En términos generales la atmosfera terrestre se divide en la baja atmosfera y la alta atmosfera. Se considera que la baja atmosfera se extiende hasta la cima de la estratosfera, aproximadamente hasta unos 50 Km. Desde el punto de vista térmico, la atmosfera se divide en diferentes capas, dentro de las cuales la temperatura, la presión y la densidad varían con la altura. Estas variaciones se puede observar con detalle en (véase la figura N°2.8, en la página"69"), donde se muestra un perfil típico para latitudes medias. Según la nomenclatura estándar definida por la Unión Internacional de Geodesia y Geofísica (IUGG) en 1960, el perfil vertical se divide en cuatro capas distintas. Estas capas son la troposfera, la atmosfera, la mesosfera y la termosfera. Las cimas de estas capas se denominan tropopausa, estratopausa y termopausa, respectivamente.

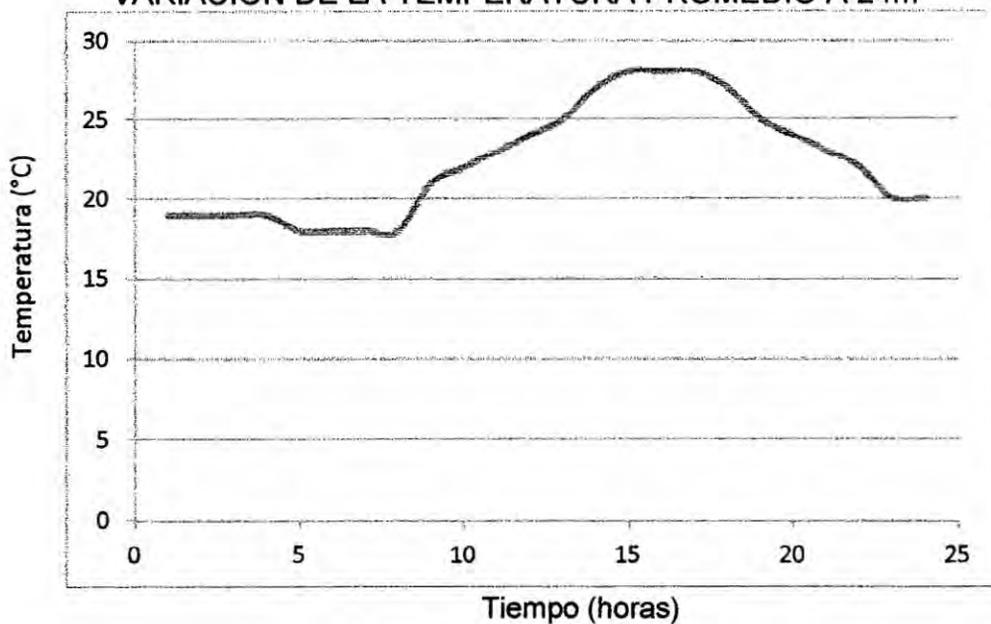
La troposfera se caracteriza por un decaimiento de la temperatura con la altura, de aproximadamente $6.5\text{ }^{\circ}\text{C}/\text{km}$. La temperatura en esta capa es consecuencia directa del balance radiativo y de los procesos convectivos de transporte de energía desde la superficie a la atmosfera. Prácticamente todo el vapor de agua, nubes y precipitación están confinados dentro de esta capa. [18]

FIGURA N° 2.8
PERFIL VERTICAL DE TEMPERATURA PARA LA ATMOSFERA



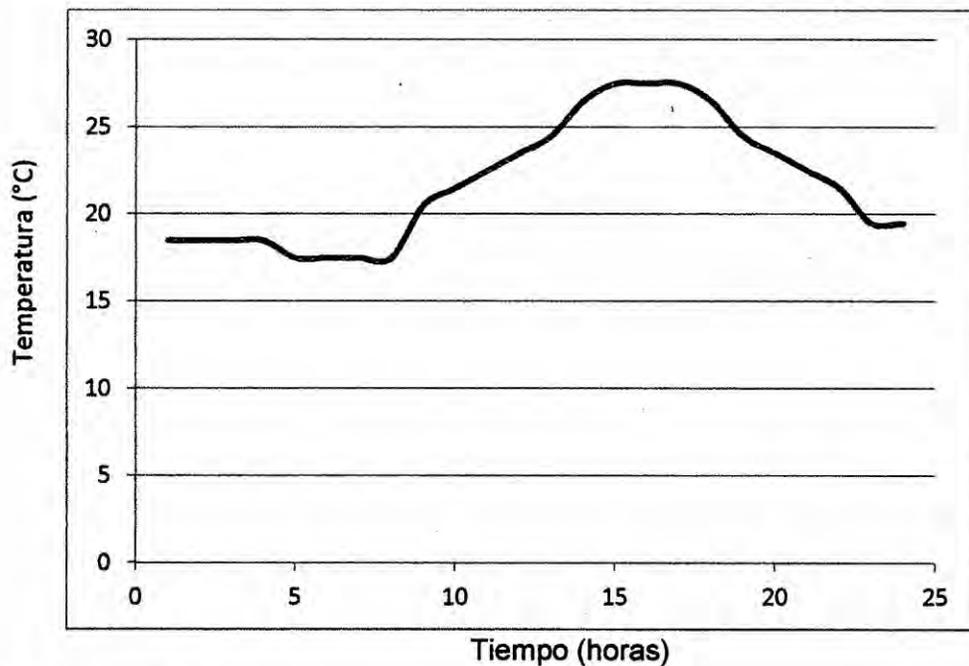
La temperatura promedio en el cerro Chocan (véase la gráfica N°2.6).

GRAFICO N° 2.6
VARIACIÓN DE LA TEMPERATURA PROMEDIO A 24m



A 80m de altura la temperatura descenderá a razón de $6.5\text{ }^{\circ}\text{C}/\text{km}$ por lo tanto el nuevo promedio anual de temperatura será (véase el grafico N°2.7).

GRAFICO N° 2.7
VARIACIÓN DE LA TEMPERATURA PROMEDIO A 80m



f. Rugosidad

Numerosos científicos dedicados al estudio del recurso eólico, enfatizan que la influencia de las características del terreno en la producción de energía de una turbina de viento, puede ser tan grande que la economía del proyecto completo dependerá de la propia elección del sitio [19]

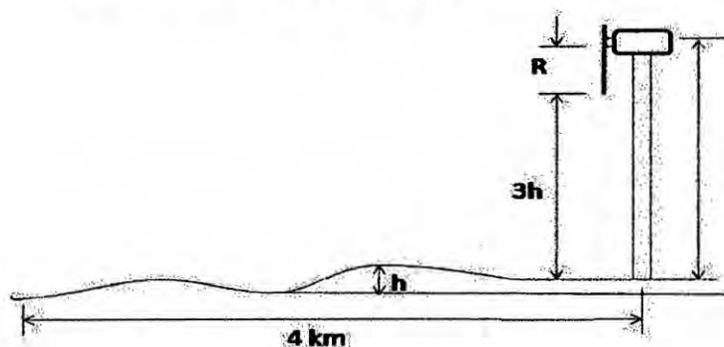
1) Clasificación del terreno.

La definición básica del terreno lo divide en llano y no llano. Algunos autores definen el terreno no llano como un terreno complejo (este

es definido como una área en donde los efectos del terreno son evidentes en el flujo que se presentan sobre la tierra considerada). Un terreno llano es aquel que tiene pequeñas irregularidades, como un bosque o un pastizal. Un terreno no llano tiene elevaciones o depresiones de gran escala, como montañas, cerros, valles y cañones. Para cuantificar que tan llano es el terreno, se debe considerar la siguiente: [2]

- Las diferencias de elevación entre el lugar en donde se pretende instalar la turbina y el terreno circundante no debe ser más de 60 metros en cualquier lugar dentro de un diámetro de 11.5Km, alrededor de esta.
- Ninguna elevación debe tener una proporción de aspecto (alto-ancho) tan grande como 1/50 dentro de 4 Km flujo arriba y flujo abajo del sitio.
- Las diferencias de elevación entre la parte más baja del rotor y la más baja elevación del terreno, debe ser tres veces la máxima elevación (h) dentro de 4Km flujo arriba, (véase la figura N°2.9).

FIGURA N°2.9
DETERMINACION DE UN TERRENO LLANO



Un terreno no llano o complejo de acuerdo con Hyster y Pennel, depende de las elevaciones aisladas o depresiones y de los terrenos montañosos. Las condiciones de flujo en terrenos montañosos son complejas, ya que las elevaciones y depresiones ocurren en forma aleatoria, por consiguiente el flujo de dichos terrenos es dividido también en pequeña y gran escala. La diferencia entre las dos, está hecha en comparación con la capa límite planetaria, la cual se asume es de cerca de 1 Km.

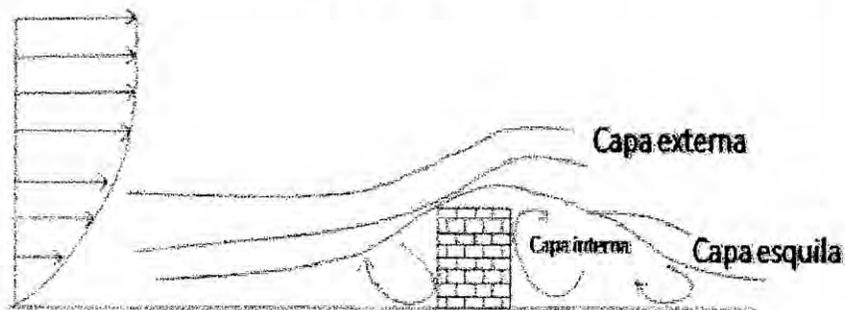
Para realizar la clasificación anterior, es importante conocer la dirección del viento, si una montaña aislada (ej. 200m de altura y 1000m de ancho), estuviera situada a 1 Km al sur del sitio propuesto, el sitio podría ser clasificado como no llano. Sin embargo, si el viento sopla solamente el 5% en esa dirección, con una velocidad media baja de 2 m/s aproximadamente, el terreno puede ser clasificado como llano.

a) Flujo sobre terreno con obstáculos.

Para el estudio de los obstáculos creados por el hombre es común proponer un bloque rectangular (véase la figura N°2.10, en la página "73") y considerar el flujo en dos dimensiones, simulando el problema en un programa de computo. Este tipo de obstáculo produce una perturbación en las líneas de flujo, formando un remolino antes y después del paso del mismo.

Dicha perturbación debe ser considerada en el momento de instalar una turbina de viento, conociendo el punto donde las líneas e flujo se estabilizan nuevamente, esto ocurre generalmente a una distancia de 15 veces la altura del edificio.

FIGURA N°2.10
ESTELA DE LA VELOCIDAD DEL VIENTO CAUSADO POR LOS OBSTACULOS



Los valores de rugosidad asociados a cada una de las superficies (véase el grafico N°2.8).[20]

GRAFICO N°2.8
ESCALA DE RUGOSIDADES

Zo (m)	Superficie
1.00	Ciudad
0.50	Bosque
0.30	Suburbios
0.20	Cinturones verdes
0.10	Arbolado abundante
0.05	Campo
0.03	Campo abierto
0.01	campo sin construcciones sin arbolado
5.10 ⁻³	Pistas de aeropuertos
10 ⁻³	hierba cortada
3.10 ⁻⁴	Terreno descubierto
10 ⁻⁴	Superficies nevadas
3.10 ⁻⁴	Superficies de arena
10 ⁻⁴	Agua (lagos, mar)

g. La humedad y su influencia con la temperatura

La humedad relativa en el medio ambiente es una cantidad numérica que indica el cociente entre la humedad absoluta del medio ambiente y la cantidad máxima de vapor de agua que admite el aire por unidad de volumen, esta cantidad se mide en porcentaje de humedad relativa (%) y su valor está normalizado de tal forma que la humedad relativa máxima posible del medio ambiente es del 100%, la determinación del valor de la humedad relativa del medio ambiente está muy ligada a la temperatura del medio ambiente en el momento de la medición, por lo que es común considerar y realizar la medición de ambas variables al mismo tiempo.[21]

Existen diversas formas para medir el contenido de vapor de agua en la atmósfera, La medición más frecuente es la de la humedad relativa, que corresponde a la fracción porcentual entre la presión parcial del vapor de agua y la presión de vapor de agua en el punto de saturación a la temperatura ambiente.

h. Variación de la densidad vertical

Para el cálculo de generación eoloelectrónica, se debe considerar una corrección por la variación de la densidad del aire en función de la altura. Con base en la norma IEC-61400-12 (IEC,1998) de la Comisión Electrotécnica Internacional, para aquellos aerogeneradores con control por desprendimiento de flujo (stall regulation), que son aquellos que mediante las propiedades aerodinámicas de las aspas limitan el torque

producido a altas velocidades, la corrección de la curva de potencia se debe hacer de acuerdo con la ecuación:

$$P_c = P_n(\rho / \rho_{ref}) \quad (2.5)$$

donde P_c [W] es la potencia generada como función de la densidad del aire atmosférico del sitio, P_n [W] es la curva de potencia del aerogenerador, ρ [kg/m³] es la densidad del aire de referencia (comúnmente 1.225 kg/m³) y ρ [kg/m³] es la densidad del aire de la localidad.

Para aquellos aerogeneradores con control de ángulo de ataque (pitch regulation), que son aquellos que reducen el ángulo de incidencia del viento para evitar daños al mismo cuando se presentan grandes velocidades de viento, se debe aplicar la corrección de acuerdo con la expresión:

$$V_c = V_n / (\rho / \rho_{ref})^{1/3} \quad (2.6)$$

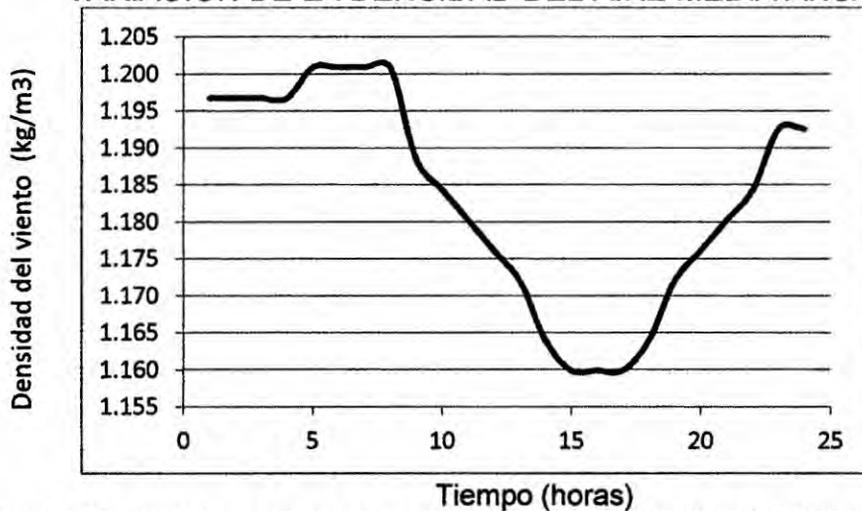
donde V_c [m/s] es la velocidad corregida y V_n [m/s] es la velocidad a la densidad de referencia. Es importante indicar que la corrección que se lleva a cabo mediante la ecuación 2.11, se trata de una corrección aplicada a la velocidad del viento, por lo que la curva potencia se desplaza sobre el eje horizontal.

Para el cálculo de la densidad del aire en función de la altura y la temperatura promedio, ésta se puede calcular mediante (Burton, 2001):

$$\rho = 1.225 \exp\left(\frac{-h}{8435} - \frac{(T-15)}{288}\right) \quad (2.7)$$

donde $h[m]$ es la altura del sitio sobre el nivel del mar (S.N.M.) y $T [^{\circ}C]$ es la temperatura ambiente promedio del sitio.[22]

GRAFICO N° 2.9
VARIACION DE LA DENSIDAD DEL AIRE MEDIA ANUAL



i. Tipos de modelos que se usan para la predicción del tiempo

1) Modelos globales

Como abarcan la totalidad de la atmósfera planetaria, con este tipo de modelos se puede realizar la predicción del tiempo en cualquier lugar del mundo. El inconveniente es que el tamaño de las celdillas en que dividen la atmósfera debe ser del orden de 40-60 km con la potencia de cálculo que proporcionan los superordenadores actuales. Esto impide realizar la predicción con el detalle necesario en regiones delimitadas del planeta.

2) Modelos regionales

Consideran solo el área del planeta donde se quiere realizar la predicción. Como el tamaño de celdillas de estos modelos es de 5 a 20 km, la predicción meteorológica resulta bastante más detallada. Para tener en

cuenta como evoluciona la atmósfera fuera del dominio geográfico donde se aplican, los modelos regionales hacen uso de los resultados de la predicción realizada con los modelos globales en las zonas que bordean el dominio considerado.

2.2.2 Modelamiento de la potencia de inyección

El viento es una masa de aire en movimiento, y como tal, contiene cierta cantidad de energía cinética, que depende de la velocidad de la masa de aire. Esta última, vendrá determinada por la temperatura y presión atmosféricas o, en definitiva, por la densidad del aire:

$$E = (1/2)m_{air}v^2 \quad (2.8)$$

donde m_{air} es la masa del aire y v es su velocidad.

Un aerogenerador convierte en energía eléctrica la energía cinética contenida en el viento. Teniendo en cuenta la energía cinética contenida en una masa de aire en movimiento que atraviesa un área determinada por el diámetro del rotor de un aerogenerador, se puede llegar a la relación entre la velocidad del viento y la potencia eléctrica entregada por el aerogenerador como la energía que lo atraviesa por unidad de tiempo:

$$P_e = (1/2)C_p \rho_{air}A_r v^3 \quad (2.9)$$

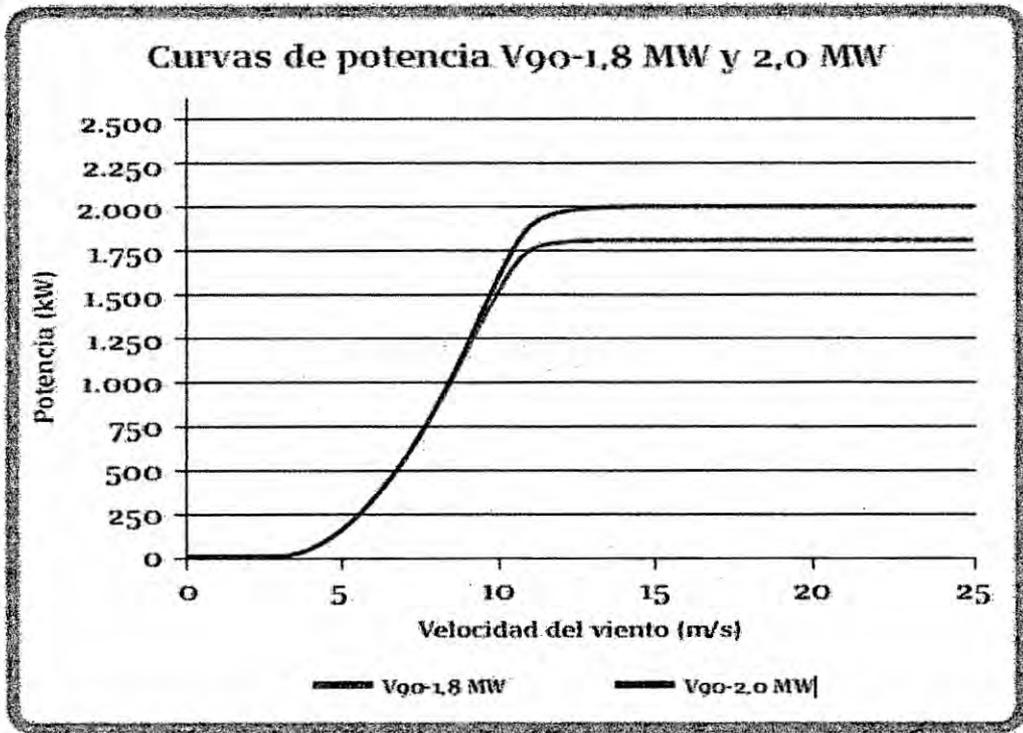
donde P_e es la potencia eléctrica entregada por el aerogenerador, ρ_{air} es la densidad del aire, A_r es el área que atraviesa, definida por el diámetro del rotor o el tamaño de las palas, y v es la velocidad de la corriente de aire. C_p es el coeficiente de potencia, que indica el grado de

aprovechamiento de la energía contenida en la corriente de aire, con un valor teórico máximo de 0.59 según el límite de Betz, aunque su valor habitual dada la tecnología actual y las pérdidas en los elementos mecánicos es bastante inferior.

La curva habitual que describe la relación entre la potencia eléctrica entregada por un aerogenerador y la velocidad del viento incidente en la turbina viene representada en (véase el gráfico N° 2.10 en la página "79"), en lo que se conoce como curva de potencia o curva P-V de un aerogenerador [23].

Para valores de velocidad de viento inferiores a unos 4 o 5 m/s, el aerogenerador no produce potencia. Ésta es la llamada velocidad de arranque del aerogenerador, a partir de la cual la potencia generada presenta un crecimiento cúbico con la velocidad del viento, hasta valores de viento de unos 12 a 15 m/s. A partir de esta velocidad del viento, la potencia eléctrica generada se mantiene en el valor nominal para el cual fue diseñada la máquina, o decrece levemente dependiendo de si se trata de un aerogenerador de velocidad variable o fija. Para valores de velocidad de viento a partir de 20 o 25 m/s, en lo que se conoce como velocidad de corte, se produce la parada del aerogenerador para protegerlo frente a vientos demasiado severos.

GRAFICO N°2.10
CURVA P-V DE UN AEROGENERADOR DE 2MW



Tiene una velocidad de arranque de 2.5m/s, en el punto de funcionamiento habitual, correspondiente a velocidades de viento entre los 5 y 12 m/s y su velocidad de corte es de 25m/s [23].

La curva de potencia presenta una fuerte pendiente, de modo que pequeñas variaciones de la velocidad del viento incidente en el rotor provocan grandes variaciones en la potencia eléctrica generada, es decir, la potencia eléctrica es muy sensible a los cambios de velocidad del viento.

Debido a esto y a que el viento es una variable que puede tener grandes cambios en escalas temporales muy cortas, del orden de apenas algunos minutos o incluso segundos, las variaciones en la

potencia entregada por un aerogenerador a lo largo del tiempo pueden ser muy pronunciadas.

Esto último da una idea de la importancia de la precisión de las previsiones de velocidad de viento proporcionadas por las agencias meteorológicas para el correcto funcionamiento de las herramientas de predicción de potencia eólica, como veremos más adelante.

a. **Potencia del viento. Máxima potencia extraíble.**

La potencia depende del área del rotor y de las velocidades de viento aguas arriba y aguas abajo del rotor. V_1 vendrá prefijada por la velocidad de viento incidente resultando la expresión siguiente [20]

$$W_{max} = \frac{8}{27} \cdot \rho \cdot A \cdot v_1^3 \quad (2.10)$$

Igualando la formula (2.3) y la formula (2.2), se obtiene el C_p máximo se expresa de la siguiente forma:

$$C_{p \max} = \frac{\frac{8}{27} \cdot \rho \cdot a \cdot v_1^3}{\frac{1}{2} \cdot \rho \cdot a \cdot v_1^3}$$

Siendo V_1 y V_2 , representadas en el siguiente dibujo:

FIGURA N°2.11
DESVIO DEL VIENTO



Llegando a la fórmula de Betz: $C_{pmax}=0.5925$, por lo tanto, como máximo, solo puede extraerse el 59% de la energía cinética que contiene una masa de aire en movimiento.

Representándose gráficamente como varia C_p en función de k , siendo k la relación de las velocidades de viento aguas arriba y aguas abajo del rotor ($k=V_1/V_2$).

GRAFICO N° 2.11
COEFICIENTE DE POTENCIA (C_p) VS K

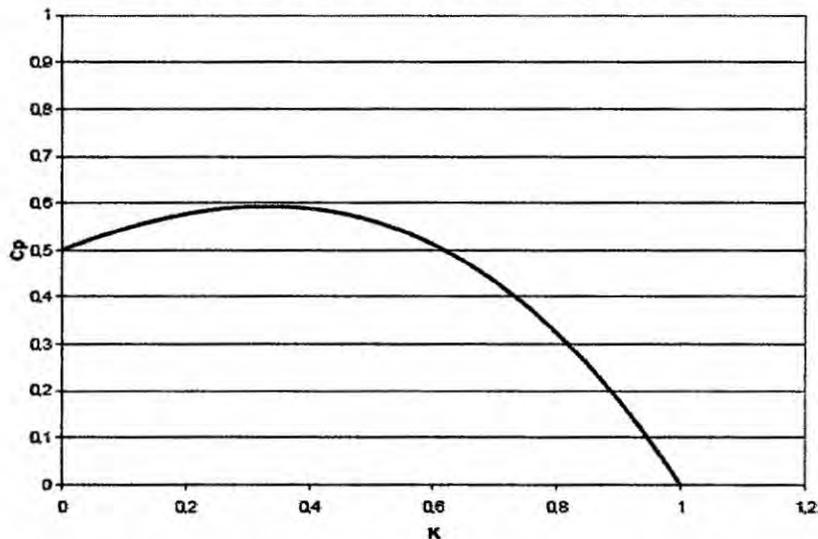
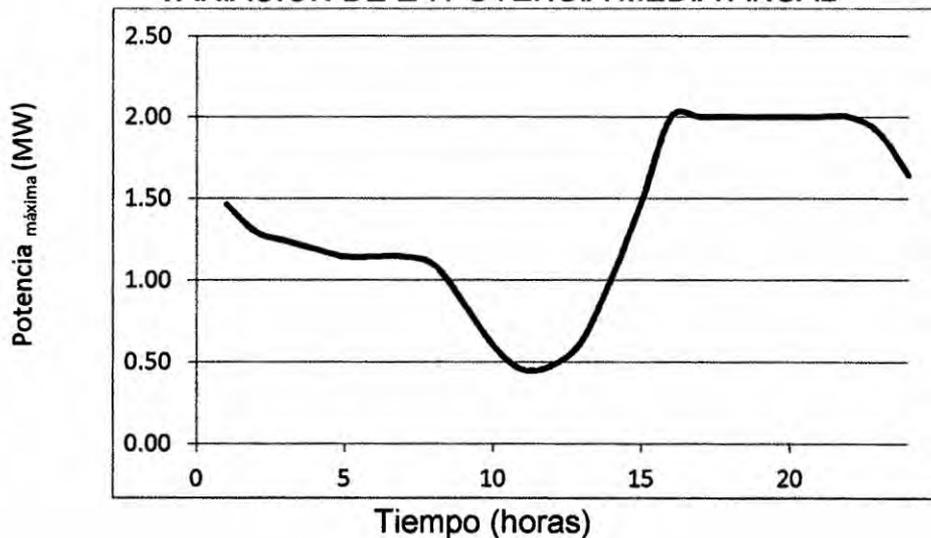


GRAFICO N° 2.12
VARIACION DE LA POTENCIA MEDIA ANUAL



b. Modelos de previsión de viento

Modelo de referencia

Al concluir con el modelo planteado y que estos valores mejoran los modelos de referencia con menor porcentaje de error, se puede concluir que el nuevo modelo va hacer de utilidad. Los modelos de referencia son el resultado de consideraciones simples. El modelo de referencia más comúnmente empleado en el marco de las herramientas de predicción eólica es el modelo persistente (que viene a ser el paseo aleatorio sin derivar). Este modelo tan sencillo establece que la predicción de la velocidad del viento es igual al valor de la última velocidad del viento medida, es decir, [6]

$$V'_{t+h,t} = V_t. \quad (2.11)$$

Pese a la extrema sencillez de este modelo, es difícil mejorarlo en los primeros horizontes de predicción (hasta 6 horas) debido a que la escala temporal de cambios en la atmósfera es muy lenta, del orden de varios días. Una generalización del modelo persistente son los predictores de medias móviles, en el cual se reemplaza el último valor de la velocidad por la media de las de las n últimas medidas disponibles.

$$V'_{t+h,t} = \frac{1}{n} \sum_{i=0}^{n-1} V_{t-i}. \quad (2.12)$$

Asintóticamente, este modelo tiende a la media global del proceso, es decir, $V'_{t+h,t} \rightarrow v$, siendo v la media del proceso. En los primeros

horizontes de predicción suele tener una precisión muy pobre, pero mejora conforme aumenta el horizonte de predicción. La exactitud de estos modelos ha sido analizada por [18], y a raíz de este estudio los autores propusieron el siguiente modelo como referencia:

$$V_{t+h,t} = a_h V_t + (1 - a_h) v_t \quad (2.13)$$

siendo a_h el coeficiente de correlación entre V_t y V_{t+h} , v_t la media del proceso hasta el instante t . Los valores V_t y a_h se determinan con la serie de datos de entrenamiento.

c. Necesidad de predicción de potencia eólica.

Debido al gran auge que están teniendo las centrales de generación de energía renovable no convencionales, se han encontrado con el problema de la predicción, que es el de conocer con anticipación la energía que va a generar las centrales eólicas, por ello la necesidad de crear herramientas de predicción para ingresar al SEIN, en condiciones confiables y eficientes sin alterar al sistema.

Lo trascendental de predecir, el mismo que forma parte de la planificación, es decir con los valores predichos se provee que no afecten al Sistema Eléctrico, siendo también una herramienta para maximizar los ingresos producto de las buenas decisiones.

La energía eléctrica generada por un parque eólico, depende fundamentalmente de las condiciones meteorológicas, por consiguiente siendo estas condiciones meteorológicas que tienen una alta variabilidad,

hacen que el problema de la predicción de energía eólica generada sea un valor complejo. Es la forma natural del viento el que origina que sea difícil tanto para largo plazo como para corto plazo, cual va hacer el valor de la velocidad.

El COES, es el operador del Sistema Eléctrico, tiene la necesidad de conocer con anterioridad cuales van hacer los valores de la energía eléctrica producida por los parques eólicos, que inyectan potencia y energía al SEIN, que ellos controlan.

Las Centrales eléctricas eólicas, producen energía cuando la velocidad del viento es superior a los 4m/s y esto puede llegar a ser muy variable e intermitente por los grandes cambios que existen en corto tiempo. Por ello, es difícil predecir con exactitud con cuanto de energía se puede contar. Por esta razón es que se hace compleja la operación del sistema y también conlleva a tener un cierto error de predicción que con los nuevos métodos y algoritmos cada vez son menores el porcentaje de error.

d. Horizonte de predicción

La elección de un modelo de predicción debe estar enfocado a que horizonte de predicción desea trabajar, es decir en qué periodo de tiempo quiere predecir, puede ser de muy corto plazo, corto plazo, mediano plazo o largo plazo.

La predicción a muy corto plazo, hasta 4 o 6 horas de horizonte, desde el punto de vista del operador del sistema tiene como objetivo la mejora de la seguridad dinámica de la red y la gestión de la regulación secundaria.

Desde el punto de vista del propietario de un parque eólico, la predicción del viento a muy corto plazo, en el orden incluso de apenas algunos minutos o segundos, resulta útil para el control de las turbinas eólicas, y por motivos de seguridad para la toma de decisiones en cuanto a paradas o arranques de los aerogeneradores. [6]

Las predicciones a corto plazo, que abarcan horizontes hasta las 48 o 72 horas, son las que tienen mayor aplicación en la determinación de necesidades de generación para los mercados diarios e intradiarios,

En las predicciones a medio y largo plazo, los horizontes pueden variar entre varios días o incluso llegar a meses o años. Las predicciones a varios días tienen aplicación en la previsión de generación de centrales eléctricas convencionales o para programar labores de mantenimiento en los aerogeneradores aquellos días en los que las expectativas de generación sean escasas o nulas. Las previsiones de potencia basadas en previsiones de viento con horizontes temporales en el plazo de años tienen utilidad en la selección de emplazamientos para ubicar parques eólicos.

En la presente tesis el horizonte de predicción es a corto plazo es decir hasta 24 horas, a la vez este horizonte es el de mayor interés para los operadores del Sistema Eléctrico.

e. Incertidumbre en la predicción.

Las predicciones meteorológicas generadas por los softwares como el Sipleolico, MM5, etc; que son modelos de predicción atmosféricas, son

generalmente propiedad de Instituto Geofísico del Perú, SENAMHI, etc., que proporcionan valores numéricos de la velocidad del viento, temperatura, presión atmosférica, etc., con estos valores numéricos se calcula la predicción de potencia eléctrica de un parque eólico ya sea por física o por cálculos matemáticos. Siendo este tipo de predicción con menor porcentaje de error comparados con modelos de predicción de energía eléctrica.

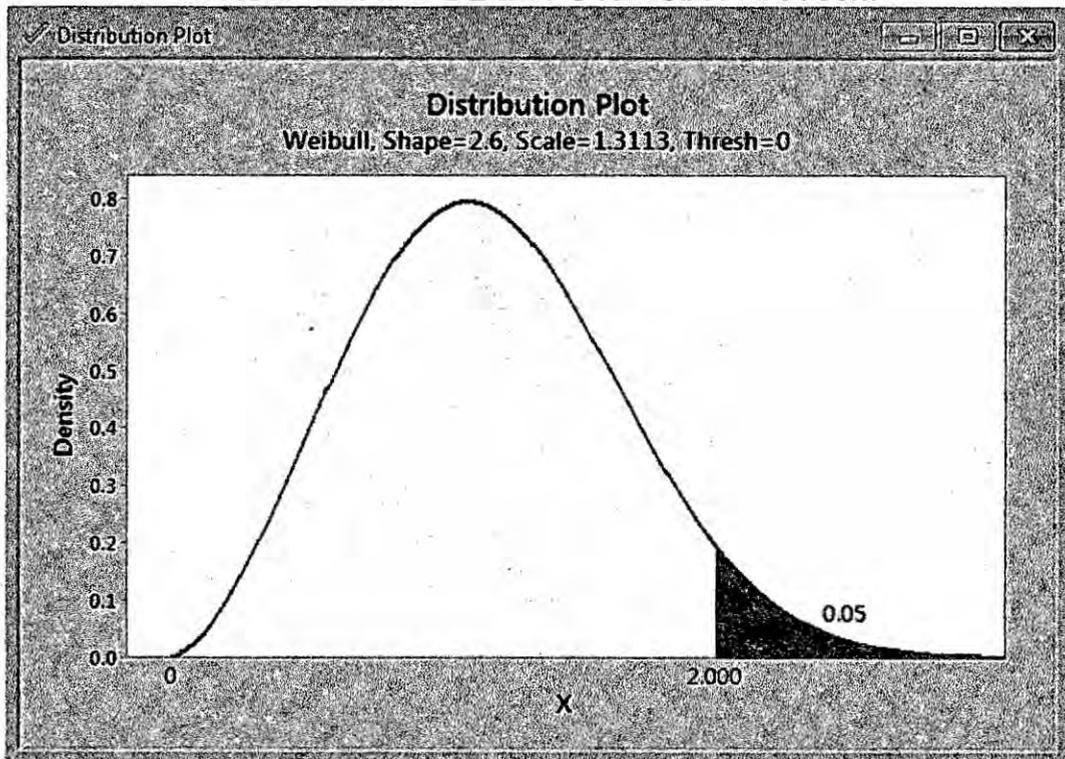
Visto el gráfico (véase el gráfico N°2.10, en la página "79"), representa la curva característica que relaciona la potencia eólica generada y la velocidad del viento (con condiciones de velocidad de viento constante). Se muestra en dicha gráfica una pendiente muy elevada en el rango de la velocidad habitual, observándose que a pequeñas variaciones del valor de la velocidad del viento induce grandes variaciones de potencia eólica generada.

La predicción por ningún método en la actualidad es exacta, tiene un porcentaje de error, pero a pesar de ello, resulta imprescindible tanto desde un punto de vista técnico como económico conocer los valores de la predicción.

f. Probabilidad de la Potencia.

Se ha analizado la distribución de probabilidad observada de velocidades medidas cada hora, y su grado de ajuste a una "distribución teórica de Weibull" (véase apéndice 2), correspondientes a 80m, para los meses de Enero a Diciembre del 2008.

GRAFICA N°2.13
PROBABILIDAD DE LA POTENCIA PARA 80m



g. Redes neuronales

1) Neuronas biológicas

Se estima que el cerebro humano contiene más de cien mil millones de neuronas y sinápsis en el sistema nervioso humano. Estudios sobre la anatomía del cerebro humano concluyen que hay más de 1000 sinápsis a la entrada y a la salida de cada neurona. Es importante notar que aunque el tiempo de conmutación de la neurona (unos pocos milisegundos) es casi un millón de veces menor que en las actuales elementos de las computadoras, ellas tienen una conectividad miles de veces superior que las actuales supercomputadoras. [24]

El objetivo principal de las redes neuronales de tipo biológico es desarrollar un elemento sintético para verificar las hipótesis que conciernen a los sistemas biológicos. Las neuronas y las conexiones entre ellas (sinápsis) constituyen la clave para el procesamiento de la información.

Desde un punto de vista funcional, las neuronas constituyen procesadores de información sencillos. Como todo sistema de este tipo, poseen un canal de entrada de información, las dendritas, un órgano de cómputo, el soma, y un canal de salida, el axón.

La neurona es la célula fundamental y básica del sistema nervioso. Es una célula alargada, especializada en conducir impulsos nerviosos. En las neuronas se pueden distinguir tres partes fundamentales, que son:

- Soma o cuerpo celular: corresponde a la parte más voluminosa de la neurona. Aquí se puede observar una estructura esférica llamada núcleo. Éste contiene la información que dirige la actividad de la neurona. Además, el soma se encuentra el citoplasma. En él se ubican otras estructuras que son importantes para el funcionamiento de la neurona.
- Dendritas: son prolongaciones cortas que se originan del soma neural. Su función es recibir impulsos de otras neuronas y enviarlas hasta el soma de la neurona.
- Axón: es una prolongación única y larga. En algunas ocasiones, puede medir hasta un metro de longitud. Su función es sacar el

impulso desde el soma neuronal y conducirlo hasta otro lugar del sistema.

Las neuronas son muy variadas en morfología y tamaño. Pueden ser estrelladas, fusiformes, piriformes. Pueden medir no más de cuatro micras o alcanzar las 130 micras. También son muy variadas en cuanto a las prolongaciones: las dendritas y el cilindroeje o axón. Las dendritas, de conducción centripeta, pueden ser únicas o múltiples.

La unión entre dos neuronas se denomina sinapsis. En el tipo de sinapsis más común no existe un contacto físico entre las neuronas, sino que éstas permanecen separadas por un pequeño vacío de unas 0.2 micras.

En relación a la sinapsis, se habla de neuronas presinápticas (la que envía las señales) y postsinápticas (la que las recibe). Las sinapsis son direccionales, es decir, la información fluye siempre en un único sentido.

Las señales nerviosas se pueden transmitir eléctrica o químicamente. La transmisión química prevalece fuera de la neurona, mientras que la eléctrica lo hace el interior. La transmisión química se basa en el intercambio de neurotransmisores mientras que la eléctrica hace uso de descargas que se producen en el cuerpo celular y que se propagan por el axón.

La forma de comunicación más habitual entre dos neuronas es de tipo químico. La neurona presináptica libera unas sustancias químicas

complejas denominadas neurotransmisores (como el glutamato o la adrenalina), que atraviesan el vacío sináptico. Si la neurona postsináptica posee en las dendritas o en el soma canales sensibles a los neurotransmisores liberados, los fijarán, y como consecuencia de ello permitirán el paso de determinados iones a través de la membrana.

Las corrientes iónicas que de esta manera se crean provocan pequeños potenciales postsinápticos excitadores (positivos) o inhibidores (negativos), que se integrarán en el soma, tanto espacial como temporalmente; éste es el origen de la existencia de sinapsis excitatoria y de sinapsis inhibitorias.

Si se ha producido un suficiente número de excitaciones, suma de los potenciales positivos generados puede elevar el potencial de la neurona por encima de los -45 mV (umbral de disparo): en ese momento se abren bruscamente los Canales de sodio, de modo que los iones Na, cuya concentración en el exterior es alta, entran masivamente al interior, provocando la despolarización brusca de neurona, que pasa de un potencial de reposo de -60 mV a unos $+50$ mV.

A continuación la neurona vuelve a la situación original de reposo de -60 mV; este proceso constituye la generación de un potencial de acción (véase la figura N°2.12, en la página "92"), que propagarse a lo largo del axón da lugar a la transmisión eléctrica de la señal nerviosa. Tras

haber sido provocado un potencial de acción, la neurona sufre un periodo refractario durante el cual no puede generarse uno nuevo.

Un hecho importante es que el pulso así generado es "digital en el sentido que existe o no existe pulso, y todos ellos son de la misma magnitud.

Por otra parte ante una estimulación más intensa disminuye el intervalo entre pulsos, por lo que la neurona se disparará a mayor frecuencia cuanto mayor sea el nivel de excitación. Es decir ante un estímulo mayor la frecuencia de respuesta aumenta hasta que se alcanza una saturación conforme se acerca a la frecuencia máxima.

La frecuencia de disparo oscila habitualmente entre 1 y 100 pulsos por segundo, aunque algunas neuronas pueden llegar a los 500 durante pequeños periodos de tiempo. Por otra parte, no todas las neuronas se disparan generando un tren de pulsos de una frecuencia aproximadamente constante, pues la presencia de otras especies iónicas hace que diferentes tipos de neuronas posean patrones de disparo distintos, en forma de trenes puros, paquetes de pulsos, o presentando patrones más complejos.

El mecanismo aquí descrito constituye la forma más común de transmisión de la señal nerviosa, pero no el único. Cuando la distancia que debe recorrer la señal es menor de 1 mm la neurona puede no codificarla en frecuencia, sino enviar una señal puramente analógica. Es decir, la evolución biológica encontró que a distancias cortas la

señal no se degradaba sustancialmente, por lo que podía enviarse tal cual, mientras que a distancias largas era preciso codificarla para evitar su degradación la consiguiente pérdida de información. La naturaleza descubrió que la codificación en forma de frecuencia de pulsos digitales proporcionaba calidad, simplicidad en la transmisión.¹

2) Modelo de una Neurona Artificial

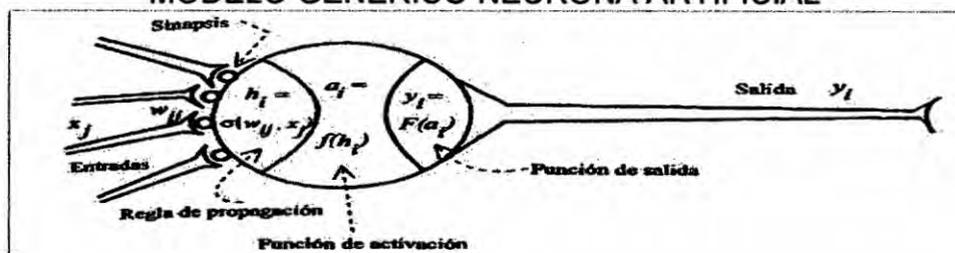
En esta sección se expone el modelo de neurona de los ANS. En primer lugar se describe la estructura de una neurona artificial muy genérica, para a continuación mostrar una versión simplificada, de amplio uso en los modelos orientados aplicaciones prácticas, que posee una estructura más próxima a la neurona tipo McCulloch-Pitts [1949].

a) Modelo general

Se denomina procesador elemental o neurona a un dispositivo simple de cálculo que, a partir de un vector de entrada procedente del exterior o de otras neuronas proporciona una única respuesta o salida.

Los elementos que constituyen la neurona de etiqueta i son los siguientes (véase la Figura 2.12):

FIGURA 2.12
MODELO GENÉRICO NEURONA ARTIFICIAL



- Conjunto de entradas, $x_j(t)$
- Pesos sinápticos de la neurona i , w_{ij} que representan la intensidad de interacción entre cada neurona presináptica j y la neurona postsináptica i .
- Regla de propagación s que proporciona el valor del potencial postsináptico $h_j(t) = s(w_{ij}, x_j(t))$ de la neurona i en función de sus pesos y entradas.
- Función de desempeño $f_i(a_i(t-1), h_i(t))$ que proporciona el estado de activación actual $a_i(t) = f_i(a_i(t-1), h_i(t))$ de la neurona i , en función de su estado anterior $a_i(t-1)$ y de su potencial postsináptico actual.
- Función de salida $F_i(a_i(t-1))$, que proporciona la salida actual $y_i(t) = F_i(a_i(t-1))$ de la neurona i en función de su estado de activación.

Este modelo de neurona formal se inspira en la operación de la biológica, en el sentido de integrar una serie de entradas y proporcionar cierta respuesta, que se propaga por el axón. A continuación se describen cada uno de los elementos.

➤ **Entradas y salidas.**

Las variables de entrada y salida pueden ser binarias (digitales) o continuas (analógicas), dependiendo del modelo y aplicación. Por ejemplo, un perceptron multicapa o MLP (Multilayer Perceptron)

admite ambos tipos de señales. Así, para tareas de clasificación poseería salidas digitales {0, +1}, mientras que para un problema de ajuste funcional de una aplicación multivariable continua, se utilizarían salidas continuas pertenecientes a un cierto intervalo.

Dependiendo del tipo de salida, las neuronas suelen recibir nombres específicos. Así, las neuronas estándar cuya salida sólo puede tomar los valores 0 o 1 se suelen denominar genéricamente neuronas de tipo McCulloch-Pitts, mientras que aquellas que únicamente pueden tener por salidas -1 o +1 se suelen denominar neuronas tipo Ising (debido al paralelismo con los modelos físicos de partículas con espín que pueden adoptar únicamente dos estados, hacia arriba y hacia abajo). Si puede adoptar diversos valores discretos en la salida (por ejemplo, -2, +1, +2), se dice que se trata de una neurona de tipo Potts. En ocasiones, el rango de los valores que una neurona de salida continua se suele limitar un intervalo definido, por ejemplo, [-1, +1] o [0, +1].

➤ **Regla de propagación**

La regla de propagación permite obtener, a partir de las entradas y los pesos el valor del potencial postsináptico h , de la neurona. La función más habitual es de tipo lineal, y se basa en la suma ponderada las entradas con los pesos sinápticos

$$h_i(t) = \sum_j w_{ij} x_j$$

que formalmente también puede interpretarse como producto escalar de los vectores de entrada y los pesos.

➤ **Función de activación**

La función de desempeño o de transferencia proporciona el estado de actual $a_i(t)$ a partir del potencial postsináptico $h_i(t)$ y del propio estado de anterior $a_i(t-1)$. Sin embargo, en muchos modelos se considera que el estado actual de neurona no depende de su estado anterior, sino únicamente del actual.

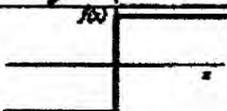
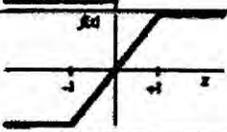
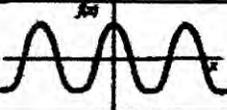
La función de desempeño se suele considerar determinista, y en la mayor parte de los modelos es monótona creciente y continua, como se observa habitualmente en las neuronas. Las funciones de activación más empleadas en muestra en (véase la figura 2.13, en la página "96").

En ocasiones los algoritmos de aprendizaje requieren que la función de desempeño cumpla la Condición de ser derivable. Las más empleadas en este sentido son las funciones de tipo sigmoideo.

Para explicar porque se utilizan estas funciones de activación se suele emplear la analogía a la aceleración de un automóvil. Cuando un auto inicia su movimiento necesita una potencia elevada para comenzar a acelerar. Pero al ir tomando velocidad, este demanda un menor incremento de dicha potencia para mantener la aceleración. Al llegar a altas velocidades, nuevamente un amplio incremento en la potencia es necesario para obtener una pequeña ganancia de

velocidad. En resumen, en ambos extremos del rango de aceleración de un automóvil se demanda una mayor potencia para la aceleración que en la mitad de dicho rango.

FIGURA 2.13
FUNCIONES DE ACTIVACIÓN MÁS USUALES

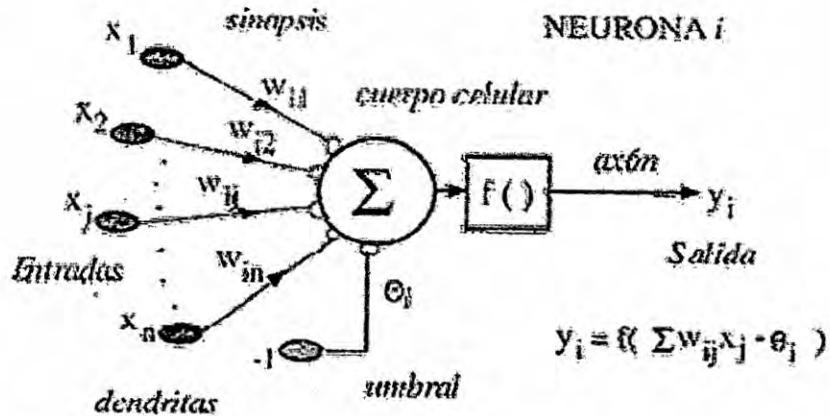
	Función	Rango	Gráfica
Identidad	$y = x$	$[-\infty, +\infty]$	
Escalón	$y = \text{sign}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$	
Lineal a tramos	$y = \begin{cases} -1, & \text{si } x < -l \\ x, & \text{si } -l \leq x \leq +l \\ +1, & \text{si } x > +l \end{cases}$	$[-1, +1]$	
Signoidea	$y = \frac{1}{1+e^{-x}}$ $y = \text{tgh}(x)$	$[0, +1]$ $[-1, +1]$	
Gaussiana	$y = Ae^{-2x^2}$	$[0, +1]$	
 sinusoidal	$y = A \text{sen}(ax + \varphi)$	$[-1, +1]$	

➤ **Función de salida**

Esta función proporciona la salida global de la neurona y en función de su estado de activación actual $a_i(t)$. Muy frecuentemente la función de salida es simplemente la identidad $f(x) = x$ de modo que el estado de activación de la neurona se considera como la propia salida.

b) Modelo estándar

FIGURA 2.14
MODELO DE NEURONA ESTÁNDAR



El modelo de neurona expuesto en la sección anterior resulta muy general la práctica suele utilizarse uno más simple, que se denomina neurona estándar, la neurona estándar consiste en

- Un conjunto de entradas $x_i(t)$ y pesos sinápticos w_{ij}
- Una regla de propagación:

$$h_i(t) = \sigma(w_{ij}, x_j(t)); \quad h_i(t) = \sum w_{ij} x_j$$

es la más común.

- Una función de desempeño que representa simultáneamente la salida de la neurona y su estado de activación.

Con frecuencia se añade al conjunto de pesos de la neurona un parámetro adicional, que se denomina umbral, que se resta del potencial postsináptico.

3) Redes neuronales y arquitectura

Las redes neuronales artificiales (ANN) son sistemas paralelos para el procesamiento de la información, inspirados en el modo en el que las redes de neuronas biológicas del cerebro procesan información.

Debido a la inspiración de las ANN en el cerebro, sus aplicaciones principales estarán centradas en campos donde la inteligencia humana no pueda ser emulada de forma satisfactoria por algoritmos aritméticos que pueden ser implementados en ordenadores. Además es de prever que dichas ANN tengan características similares a las del cerebro:

- ✓ Serán robustas y tolerantes a fallos. En el cerebro mueren todos los días gran cantidad de neuronas sin afectar sensiblemente a su funcionamiento.
- ✓ Serán flexibles. El cerebro se adapta a nuevas circunstancias mediante el aprendizaje
- ✓ Podrán trabajar con información borrosa, incompleta, probabilística, con ruido o inconsistente.
- ✓ Serán altamente paralelas. El cerebro está formado por muchas neuronas interconectadas entre si y es precisamente el comportamiento colectivo de todas ellas lo que caracteriza su forma de procesar la información.

El punto clave de las ANN es la nueva estructura de estos sistemas para el procesamiento de la información. Estos están compuestos, al igual que el cerebro, por un número muy elevado de elementos básicos

(las neuronas), altamente interconectados entre ellos y con modelo de respuesta para cada elemento en función de su entorno muy parecido al comportamiento de las neuronas biológicas. Estos modelos son simulados en ordenadores convencionales y es el comportamiento colectivo de todos los elementos lo que le confiere esas características tan peculiares para la resolución de problemas complejos. Las ANNs, como las personas, aprenden a partir de ejemplos. Aprender en sistemas biológicos involucra la modificación de la ínter conectividad entre las neuronas y esto es también cierto para las ANNs.

Las ANNs han sido aplicadas a un creciente número de problemas reales de considerable complejidad, por ejemplo reconocimiento de patrones, clasificación de datos, predicciones, etc. Su ventaja más importante está en solucionar problemas que son demasiado complejos para las técnicas convencionales: problemas que no tienen un algoritmo específico para su solución, o cuyo algoritmo es demasiado complejo para ser encontrado.

A continuación se puede ver, en la Figura 2.15, un esquema de una red neuronal:

FIGURA 2.15
EJEMPLO DE UNA RED NEURONAL TOTALMENTE CONECTADA

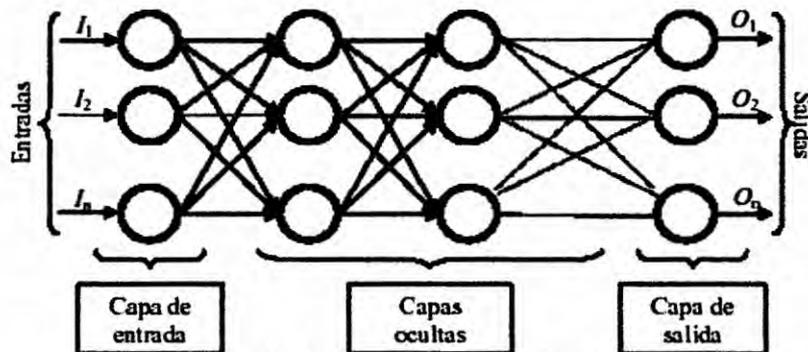


Figura 4.1: ejemplo de una red neuronal totalmente conectada.

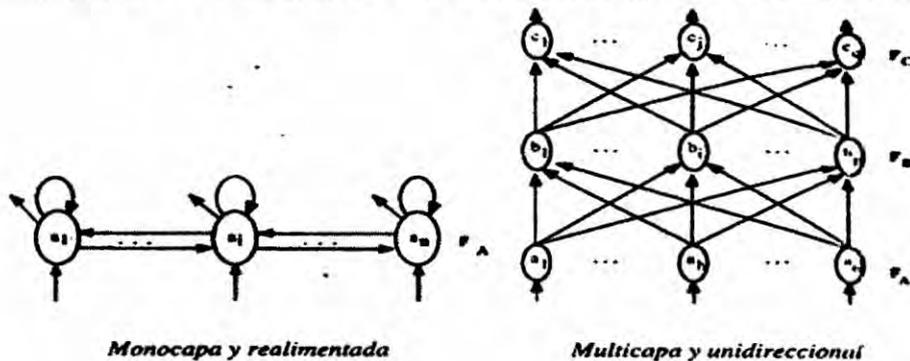
La misma está constituida por neuronas interconectadas y arregladas en tres capas (esto último puede variar). Los datos ingresan por medio de la "capa de entrada", pasan a través de la "capa oculta" y salen por la "capa de salida". Cabe mencionar que la capa oculta puede estar constituida por varias capas.

a) Tipos de arquitectura

Se denomina arquitectura a la topología, estructura o patrón de conexionado de una red neuronal.

Atendiendo a distintos conceptos, pueden establecerse diferentes tipos de arquitecturas neuronales.

FIGURA 2.16
EJEMPLOS DE ARQUITECTURAS NEURONALES



Así, en relación a su estructura en capas, se puede hablar de redes monocapa y de redes multicapa. Las redes monocapa son aquellas compuestas por una única capa de neuronas. Las redes multicapa (layered networks) son aquellas cuyas neuronas se organizan en varias capas.

Asimismo, atendiendo al flujo de datos en la red neuronal, se puede hablar de **redes unidireccionales (feedforward)** y redes recurrentes (feedback). En las redes unidireccionales **la información circula en un único sentido**, desde las neuronas de entrada hacia las de salida. En las redes recurrentes o realimentadas la información puede circular entre las capas en cualquier sentido, incluido el de salida-entrada.

Por último, también se habla de redes autoasociativas y heteroasociativas. Con frecuencia se interpreta la operación de una red neuronal como la de una memoria asociativa que ante un determinado patrón de entradas responde con un cierto patrón

Si una red se entrena para que ante la presentación de un patrón A responda otro diferente B, se dice que la red es heteroasociativa. Si una red es entrenada para que asocie un patrón A consigo mismo, se dice que es autoasociativa.

4) Aprendizaje de una red neuronal

El aprendizaje es el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se

producen durante el mismo se reducen a la destrucción, modificación y creación de conexiones entre las neuronas. Una red neuronal debe aprender a calcular la salida correcta para cada constelación (arreglo o vector) de entrada en el conjunto de ejemplos.

Este proceso de aprendizaje se denomina: Proceso de Entrenamiento o Acondicionamiento. El conjunto de datos (o conjunto de ejemplos) sobre el cual este proceso se basa es, por ende, llamado: Conjunto de datos de Entrenamiento.

Durante el proceso de aprendizaje, los pesos de las conexiones de la red sufren modificaciones, por lo tanto, se puede afirmar que este proceso ha terminado (la red ha aprendido) cuando los valores de los pesos permanecen estables, si los pesos cambian y sus valores son iguales a 0, se dice que la conexión de la red se ha destruido. De esta manera, se dice que los pesos se han adaptado, ya que sus valores son distintos de 0 y su derivada es igual a 0.

En otras palabras el aprendizaje es el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se producen durante el mismo se reducen a la destrucción, modificación y creación de conexiones entre las neuronas.

Un aspecto importante respecto al aprendizaje de las redes neuronales es el conocer cómo se modifican los valores de los pesos, es decir, cuáles son los criterios que se siguen para cambiar el valor asignado a

las conexiones cuando se pretende que la red aprenda una nueva información.

Clásicamente se distinguen dos modos de operación en los sistemas neuronales el modo recuerdo o ejecución, y el modo aprendizaje o entrenamiento.

a) Fase de aprendizaje. Convergencia

Hay dos métodos de aprendizaje importantes que pueden distinguirse:

- Aprendizaje supervisado.
- Aprendizaje no supervisado.
- **Aprendizaje supervisado**

El aprendizaje supervisado se caracteriza porque el proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo (supervisor, maestro) que determina la respuesta que debería generar la red a partir de una entrada determinada. El supervisor controla la salida de la red y en caso de que ésta no coincida con la deseada, se procederá a modificar los pesos de las conexiones, con el fin de conseguir que la salida obtenida se aproxime a la deseada. En este tipo de aprendizaje se suelen considerar, a su vez, tres formas de llevarlo a cabo, que dan lugar a los siguientes aprendizajes supervisados:

- Aprendizaje por corrección de error: Consiste en ajustar los pesos de las conexiones de la red en función de la diferencia

entre los valores deseados y los obtenidos a la salida de la red, es decir, en función del error cometido en la salida.

- Aprendizaje por refuerzo: Se trata de un aprendizaje supervisado, más lento que el anterior, que se basa en la idea de no disponer de un ejemplo completo del comportamiento deseado, es decir, de no indicar durante el entrenamiento exactamente la salida que se desea que proporcione la red ante una determinada entrada. En el aprendizaje por refuerzo la función del supervisor se reduce a indicar mediante una señal de refuerzo si la salida obtenida en la red se ajusta a la deseada (éxito = +1 o fracaso = -1), y en función de ello se ajustan los pesos basándose en un mecanismo de probabilidades. Se podría decir que en este tipo de aprendizaje la función del supervisor se asemeja más a la de un crítico (que opina sobre la respuesta de la red) que a la de un maestro (que indica a la red la respuesta concreta que debe generar), como ocurría en el caso de supervisión por corrección del error.
- Aprendizaje estocástico: Consiste básicamente en realizar cambios aleatorios en los valores de los pesos de las conexiones de la red y evaluar su efecto a partir del objetivo deseado y de distribuciones de probabilidad. En pocas palabras el aprendizaje consistiría en realizar un cambio aleatorio de los valores de los pesos y determinar la energía de

la red. Si la energía es menor después del cambio, es decir, si el comportamiento de la red se acerca al deseado, se acepta el cambio; si, por el contrario, la energía no es menor, se aceptaría el cambio en función de una determinada y preestablecida distribución de probabilidades.

- **Aprendizaje no supervisado**

Las redes con aprendizaje no supervisado (también conocido como autosupervisado) no requieren influencia externa para ajustar los pesos de las conexiones entre sus neuronas. La red no recibe ninguna información por parte del entorno que le indique si la salida generada en respuesta a una determinada entrada es o no correcta. Estas redes deben encontrar las características, regularidades, correlaciones o categorías que se puedan establecer entre los datos que se presenten en su entrada.

Existen varias posibilidades en cuanto a la interpretación de la salida de estas redes, que dependen de su estructura y del algoritmo de aprendizaje empleado. En cuanto a los algoritmos de aprendizaje no supervisado, en general se suelen considerar dos tipos, que dan lugar a los siguientes aprendizajes:

- Aprendizaje hebbiano: Esta regla de aprendizaje es la base de muchas otras, la cual pretende medir la familiaridad o extraer características de los datos de entrada. El fundamento es una suposición bastante simple: si dos neuronas N_i y N_j toman el

mismo estado simultáneamente (ambas activas o ambas inactivas), el peso de la conexión entre ambas se incrementa. Las entradas y salidas permitidas a la neurona son: $\{-1, 1\}$ o $\{0, 1\}$ (neuronas binarias). Esto puede explicarse porque la regla de aprendizaje de Hebb se originó a partir de la neurona biológica clásica, que solamente puede tener dos estados: activa o inactiva.

- Aprendizaje competitivo y comparativo: Se orienta a la clusterización o clasificación de los datos de entrada. Como característica principal del aprendizaje competitivo se puede decir que, si un patrón nuevo se determina que pertenece a una clase reconocida previamente, entonces la inclusión de este nuevo patrón a esta clase matizará la representación de la misma. Si el patrón de entrada se determinó que no pertenece a ninguna de las clases reconocidas anteriormente, entonces la estructura y los pesos de la red neuronal serán ajustados para reconocer la nueva clase.

b) Fase de recuerdo o ejecución. Estabilidad

Generalmente (aunque no en todos los modelos), una vez que el sistema ha sido entrenado, el aprendizaje se desconecta por lo que los pesos y la estructura quedan fijos, estando la red neuronal ya dispuesta para procesar datos. Este modo de operación se denomina modo recuerdo (recall) o de ejecución.

En las redes unidireccionales, ante un patrón de entrada, las neuronas responden proporcionando directamente la salida del sistema. Al no existir bucles de realimentación no existe ningún problema en relación con su estabilidad. Por el contrario, las redes con realimentación son sistemas dinámicos no lineales, que requieren ciertas condiciones para que su respuesta acabe convergiendo a un estado estable o punto fijo. Una serie de teoremas generales como el Teorema de Cohen – Grossberg – Kosko¹, indican las condiciones que aseguran la estabilidad de la respuesta en una amplia gama de redes neuronales, bajo determinadas condiciones.

Para demostrar la estabilidad del sistema, estos teoremas se basan en el método de Lyapunov², como alternativa a mucho más tedioso método directo.

El método de Lyapunov constituye una manera asequible de estudiar la estabilidad de un sistema dinámico. Es interesante observar que con la formulación matemática planteada en este método simplemente se está expresando que si se es capaz de encontrar una cierta función energía del sistema, que disminuya siempre en su operación, entonces el sistema es estable.

Una técnica similar empleó Hopfield para demostrar que su modelo

1 El enunciado completo de los teoremas así como el desarrollo y la explicación de ellos se pueden ver en <http://www.nsi.edu/users/izhikevich/publications/arbib.pdf>

2 Para más información sobre el método de Lyapunov, ver el artículo "Sobre el Método de Lyapunov" en http://www.red-mat.unam.mx/foro/volumenes/vol010/volten_2.html

de red completamente interconectada era estable en de que la matriz de pesos sinápticos fuese simétrica y de diagonal nula.

Esta técnica es también la que Cohen, Grossberg y Kosko han aplicado en los teoremas citados para demostrar la estabilidad de una amplia clase de redes neuronales realimentadas, autoasociativas y heteroasociativas. Así, el teorema de Cohen-Grossber determina las condiciones de estabilidad para redes autoasociativas no adaptativas, el de Cohen-Grossberg-Kosko, establece las condiciones de estabilidad para redes autoasociativas adaptativas; y, por último, el teorema ABAM de Kosko lo hace para redes adaptativas heteroasociativas

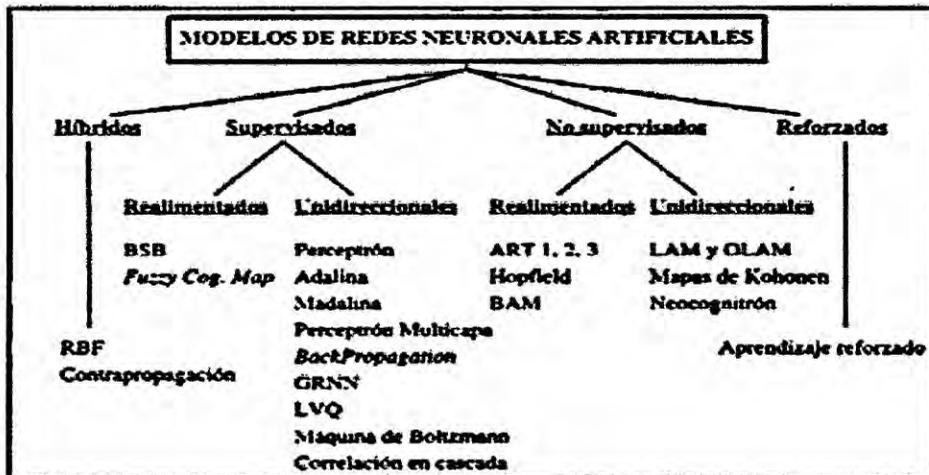
5) Clasificación de los modelos neuronales

Dependiendo del modelo concreto de neurona que se utilice, de la arquitectura o topología de la conexión, y del algoritmo de aprendizaje, surgirán distintos modelos de redes neuronales.

De la multitud de modelos y variantes que de hecho existen, unos cincuenta son medianamente conocidos, aunque tan sólo aproximadamente una quincena son utilizados con asiduidad en las aplicaciones prácticas. Por lo tanto, para llevar a cabo el estudio sistemático de los modelos se precisa algún tipo de clasificación.

Los dos conceptos que más caracterizan un modelo neuronal son el aprendizaje y la arquitectura de la red, por ello, se considera que la clasificación debe atender ambos aspectos.

FIGURA 2.17
 CLASIFICACIÓN DE LOS ANS POR EL TIPO DE APRENDIZAJE Y ARQUITECTURA



De esta manera, en primer lugar, se realiza una distinción en cuanto al tipo de aprendizaje, por lo que aparece una primera clasificación en modelos supervisados, no supervisados, de aprendizaje híbrido y modelos de aprendizaje reforzado. A su vez, y dentro de cada uno de los grandes grupos, tendremos en cuenta el tipo de topología de la red, por lo que se distinguirá además entre redes realimentadas y redes unidireccionales (no realimentadas). La clasificación que surge se muestra (véase la Figura N°2.17).

Se puede apreciar que el conjunto de **modelos de redes no realimentadas y de aprendizaje supervisado es el más numeroso**. Esta clase de modelos resulta especialmente importante por varias razones: por su interés histórico, generalidad, por ilustrar una amplia clase de aspectos que aparecen con frecuencia en todo el campo de las redes neuronales (memoria asociativa, clasificación, aproximación

funcional, etc.), y además por ser los sistemas neuronales más empleados en las aplicaciones prácticas.

6) Computabilidad neuronal

Establecidos los ANS como un estilo de procesamiento alternativo complementario al clásico basado en computadores digitales serie (tipo von Neumann), se hace necesario profundizar en sus características computacionales. Es bien sabido que un ordenador digital constituye una máquina universal de Turing, por lo que puede realizar cualquier cómputo. Además, al estar construido en base a funciones lógicas, se deduce que cualquier problema computacional puede ser resuelto con funciones booleanas.

Se ha discutido extensamente sobre las características computacionales de los ANS, demostrándose en particular que, al igual que los computadores digitales convencionales, las redes neuronales son formalmente capaces de resolver cualquier problema computacional.

Por lo tanto, los ANS, como los ordenadores convencionales, son máquinas universales, por lo que para resolver un determinado problema, cualquiera de las dos aproximaciones sería perfectamente válida, en principio.

La cuestión que entonces surge es, dado un problema, cuál de las dos alternativas, procesamiento neuronal o convencional, resulta más eficiente en su resolución.

Estudiando en el campo de las redes neuronales los aspectos relacionados con la complejidad computacional, en varios estudios se deduce que los problemas que requieren un extenso algoritmo o que precisan almacenar un gran número de datos, aprovechan mejor la estructura de una red neuronal que aquellos otros que requieren algoritmos cortos. Así, un ordenador digital resulta más eficiente en la ejecución de tareas aritméticas y lógicas, mientras que un ANS resolverá mejor problemas que deban tratar con grandes bases de datos que almacenen ingentes cantidades de información, y en los que existan muchos casos particulares, como sucede en los problemas de reconocimiento de patrones en ambiente natural.

De esta manera podemos concluir que un estilo de computación no es mejor que el otro, simplemente para cada problema particular se deberá elegir el método más adecuado, y en el caso de problemas muy complejos, éstos deberían ser separados en partes, para resolver cada una mediante el método más idóneo.

7) Realización y aplicaciones de las redes neuronales

El modo más habitual de realizar una red neuronal consiste en simularla en un ordenador convencional, como un PC o una estación de trabajo, haciendo uso de programas escritos en lenguajes de alto nivel, como C o Pascal. Aunque de esta manera se pierde su capacidad de cálculo en paralelo, las prestaciones que ofrecen los

ordenadores actuales resultan suficientes para resolver numerosos problemas prácticos, permitiendo la simulación de redes de tamaño considerable a una velocidad razonable. Esta constituye la manera más barata y directa de realizar una red neuronal. Además, no es necesario que cada diseñador confeccione sus propios simuladores, pues hay disponible comercialmente software de simulación que permite el trabajo con multitud de modelos neuronales.

En el resto de las maneras de realizar un ANS se trata de aprovechar, en mayor o menor medida, su estructura de cálculo paralelo. Un paso adelante en este sentido consiste en simular la red sobre computadores con capacidad de cálculo paralelo (sistemas multiprocesador, máquinas vectoriales, masivamente paralelas...). Una orientación diferente consiste en llevar a cabo la emulación hardware de la red neuronal, mediante el empleo de sistemas de cálculo expresamente diseñados para realizar ANS basados, o en microprocesadores de altas prestaciones (RISC DSP...), o en procesadores especialmente diseñados para el trabajo con redes neuronales.

Estas estructuras se suelen denominar placas aceleradoras neuroemuladores o neurocomputadores de propósito general. Algunos sistemas de desarrollo de redes neuronales, además de un software de simulación, incluyen dispositivos de este tipo, en forma de tarjetas conectables al bus de un PC.

La realización electrónica de redes neuronales es un campo muy activo, abordado tanto por grupos de investigación universitarios como por empresas de los sectores de la electrónica e informática. Compañías como Siemens, Philips, Hitachi, AT&T, IBM o Intel han puesto en marcha desde mediados de los años ochenta programas de investigación y desarrollo en este campo. Asimismo, se han creado diversas empresas que tratan de explotar comercialmente (con mejor o peor fortuna) estos nuevos desarrollos.

Las aplicaciones más habituales de las redes neuronales son las relacionadas con clasificación, estimación funcional y optimización; en general, el del reconocimiento de patrones suele considerarse como un denominador común. Se pueden señalar, entre otras, las siguientes áreas de aplicación de los sistemas neuronales: reconocimiento del habla, reconocimiento de caracteres, visión, robótica, control, procesamiento de señal, predicción, economía, defensa, bioingeniería, etc.

Asimismo, se están aplicando ANS para incorporar aprendizaje en los sistemas borrosos y a la confección de sistemas expertos conexionistas. Un área de intenso trabajo es el del tratamiento de la información económica, siendo uno de los grupos punteros el de A.N. Refenes, de la London Business School.

Otra de las áreas importantes es la industria. Fujitsu, Kawasaki y Nippon Steel emplean ANS en el control de procesos industriales,

como por ejemplo en plantas de producción de acero. Siemens aplica redes neuronales y sistemas borrosos en la fabricación de celulosa en laminadoras y en galvanizadoras. Citróen emplea redes neuronales en la determinación de la calidad del material utilizado en la confección de los asientos de los vehículos, Ford en reducción de contaminantes y Renault para detectar averías en el encendido de los automóviles.⁴

8) Redes neuronales supervisadas

En esta parte del capítulo se tratarán los modelos de redes más populares, así como sus principales algoritmos, se inicia con los modelos más sencillos, esto con el fin de ir introduciendo los conceptos necesarios para describir el perceptron y el algoritmo denominado backpropagation que servirán de base teórica para la experimentación. No se desarrollarán las redes neuronales no supervisadas debido a que no son redes diseñadas para la forma en que se predicen series de tiempo, principal objetivo de este trabajo.

a) Asociador lineal

Este modelo, mediante una transformación lineal, asocia un conjunto de patrones de entrada a otros de salida. El asociador lineal consta únicamente de una capa de neuronas lineales, cuyas entradas las denotamos por x y sus salidas por y , vector que constituye además la respuesta de la red neuronal.

⁴ Un índice muy completo de aplicaciones de las RNA se puede encontrar en <http://www.ipatlas.com/pub/nap/>

Asimismo, denotaremos por $W = \{w_{ij}\}$ a la matriz de pesos, cada fila de W contiene los pesos de una neurona w_i .

$$W = (w_1 \ w_2 \ \dots \ w_m)^T$$

La operación del asociador lineal es simplemente

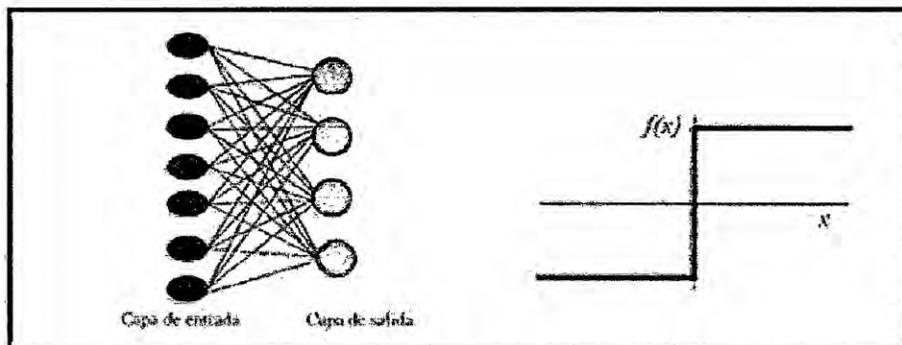
$$y = Wx$$

O bien

$$y_i = \sum_{j=1}^n w_{ij} x_j$$

Por lo tanto, cada neurona i del asociador lineal lleva a cabo la suma ponderada de las entradas con sus pesos sinápticos. Es decir esta neurona calcula el potencial postsináptico por medio de la convencional suma ponderada, cantidad a la que aplica finalmente una función activación de tipo identidad.

FIGURA 2.18
ASOCIADOR LINEAL



El problema se centra en encontrar la matriz de pesos W óptima en el sentido descrito anteriormente en este capítulo. Para ello, en el campo de las redes neuronales normalmente se hace uso de una

regla de aprendizaje, que a partir de las entradas y de las salidas deseadas (en el caso del aprendizaje supervisado), proporcione el conjunto óptimo de pesos W .

- Regla de aprendizaje de hebb

Uno de los modelos clásicos de aprendizaje de redes neuronales es el propuesto por Hebb (1949), el cual postulo un mecanismo de

aprendizaje para una neurona biológica, cuya idea básica consiste en que cuando un axón presináptico causa la activación de cierta neurona pos-sináptica, la eficacia de la sinapsis que las relaciona se refuerza.

Si bien este tipo de aprendizaje es simple y local, su importancia radica en que fue pionero tanto en neurociencias como en neurocomputación, de ahí que otros algoritmos más complejos lo tomen como punto de partida.

De manera general se denomina aprendizaje Hebbiano a un aprendizaje que involucra una modificación en los pesos, w_{ij} proporcional al producto de una entrada x_j y de una salida y_i de la neurona. Es decir, $w_{ij} = e y_i x_j$, donde a $0 < e < 1$ se le denomina ritmo de aprendizaje.

Consideremos el caso del asociador lineal. La regla de Hebb se

⁵ Algo más sobre el asociador lineal puede ser consultado en <http://www.comp.nus.edu.sg/~pris/AssociativeMemory/LinearAssociator.html>

expresa en este caso particular así:

$$\Delta w_{ij}^u = t_i^u x_j^u$$

Por lo tanto, el algoritmo de entrenamiento regla de Hebb para el asociador lineal es:

$$w_{ij}^{\text{new}} = w_{ij}^{\text{old}} + \Delta w_{ij}^u$$

Si los pesos de partida son nulos, el valor final de W para las p asociaciones será:

$$W = t^1 x^{1t} + t^2 x^{2t} + \dots + t^p x^{pt}$$

- Regla de la pseudoinversa

La regla de aprendizaje de Hebb ha sido introducida debido a su plausibilidad biológica. Sin embargo, en general se tratará de deducir los algoritmos de aprendizaje a partir de un cierto criterio a optimizar; el aprendizaje usualmente se planteará como un procedimiento para alcanzar el conjunto de pesos óptimo que resuelva un problema dado. Para ello se hace necesario definir el significado de "óptimo" en cada caso concreto, es decir, hay que proponer un criterio que mida el rendimiento de la red neuronal para encontrar una regla de actualización de pesos que lo optimice. Una forma habitual de definir el rendimiento es el error

cuadrático medio de las salidas actuales de la red respecto de las deseadas. Para el asociador lineal se tendría

$$E\{w_{ij}\} = (1/p) \sum_{u=1}^p |t^u - Wx^u|^2 = (1/p) \sum_{u=1}^p \sum_{i=1}^n (t_i^u - Wx_i^u)^2$$

De este modo, un algoritmo de aprendizaje para el asociador lineal debería obtener un conjunto de pesos que minimicen esta expresión del error. Si denominamos X a una matriz $n \times p$ que tiene por columnas los vectores de entrada x^u , $X = (x^1, x^2, \dots, x^p)$ y si llamamos Y a la matriz $m \times p$ cuyas columnas son los vectores de salida y^u , $Y = (y^1, y^2, \dots, y^p)$, la ecuación anterior se transforma en

$$E\{w_{ij}\} = (1/p) \|Y - WX\|^2$$

Con esta nomenclatura, la regla de Hebb se expresa de la forma siguiente:

$$W = YX^T$$

Una regla de aprendizaje basada en la utilización de la matriz pseudoinversa puede escribirse como:

$$W = YX^+$$

Donde X^+ denota la pseudoinversa de X. Debido a que ambas reglas son óptimas según el mismo criterio, la regla de Hebb y la de la pseudoinversa deben estar muy relacionadas. Esta circunstancia es fácil de apreciar, pues si consideramos un conjunto de vectores de entrada ortonormales, la regla de la

pseudoinversa se convierte en la de Hebb. Por otra parte, si se realiza la expansión en serie de la ecuación de la pseudoinversa, el primer término de la serie es Precisamente la ecuación de la regla de Hebb. Es decir, la regla de Hebb representa en el fondo un caso particular de la más general regla de la pseudoinversa.

b) Perceptron simple

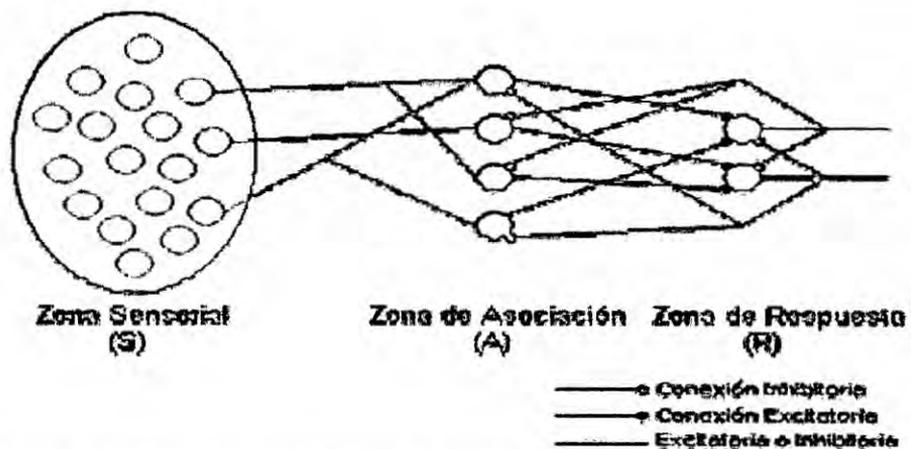
La red tipo Perceptrón fue inventada por el psicólogo Frank Rosenblatt en el año 1957. Su intención era ilustrar algunas propiedades fundamentales de los sistemas inteligentes en general, sin entrar en mayores detalles con respecto a condiciones específicas y desconocidas para organismos biológicos concretos. Rosenblatt creía que la conectividad existente en las redes biológicas tiene un elevado porcentaje de aleatoriedad, por lo que se oponía al análisis de McCulloch Pitts en el cual se empleaba lógica simbólica para analizar estructuras bastante idealizadas.

Rosenblatt opinaba que la herramienta de análisis más apropiada era la teoría de probabilidades, y esto lo llevó a una teoría de separabilidad estadística que utilizaba para caracterizar las propiedades más visibles de estas redes de interconexión ligeramente aleatorias.

El primer modelo de Perceptrón fue desarrollado en un ambiente biológico imitando el funcionamiento del ojo humano, el fotoperceptrón como se le llamo era un dispositivo que respondía a

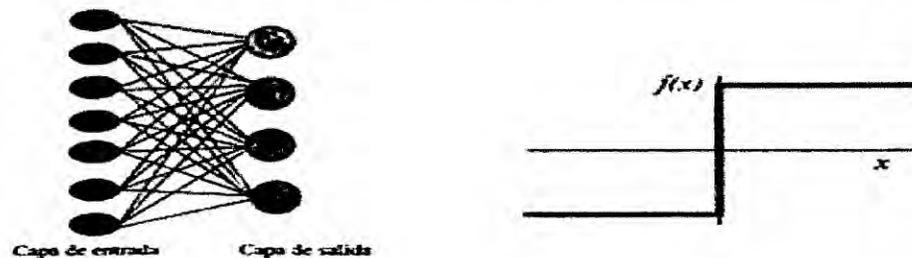
señales ópticas; como se muestra en el figura 2.18 la luz incide en los puntos sensibles (S) de la estructura de la retina, cada punto S responde en forma todo-nada a la luz entrante, los impulsos generados por los puntos S se transmiten a las unidades de asociación (A) de la capa de asociación; cada unidad A está conectada a un conjunto aleatorio de puntos S, denominados conjunto fuente de la unidad A, y las conexiones pueden ser tanto excitatorias como inhibitorias.

FIGURA 2.19
MODELO DEL FOTOPERCEPTRÓN DE ROSENBLATT



De forma similar, las unidades A están conectadas a unidades de respuesta (R) dentro de la capa de respuesta y la conectividad vuelve a ser aleatorio entre capas, pero se añaden conexiones inhibitorias de realimentación procedentes de la capa de respuesta y que llegan a la capa de asociación, también hay conexiones inhibitorias entre las unidades R.

FIGURA 2.20
ARQUITECTURA (IZQUIERDA) Y FUNCIÓN DE
TRANSFERENCIA (DERECHA) DE UN PERCEPTRÓN SIMPLE



El perceptrón simple es un modelo unidireccional, compuesto por dos capas de neuronas, una sensorial o de entradas, y otra de salida (véase la figura N°2.20). La operación de una red de este tipo, con n neuronas de entrada y m de salida, se puede expresar como.

$$y_i(t) = f\left(\sum_{j=1}^n w_{ij}x_j - \theta_i\right), \quad \forall i, 1 \leq i \leq m$$

Las neuronas de entrada no realizan ningún cómputo, únicamente envían la información (en principio consideraremos señales discretas $\{0, +1\}$) a las neuronas de salida (en el modelo original estas neuronas de entrada representaban información ya procesada, no datos directamente procedentes del exterior).

La función de desempeño las neuronas de la capa de salida es de tipo escalón. Así, la operación de un perceptrón simple puede escribirse

$$y_i = H\left(\sum_{j=1}^n w_{ij}x_j - \theta_i\right), \quad \forall i, 1 \leq i \leq m$$

con $H(\cdot)$ la función de Heaviside o escalón. El perceptrón puede utilizarse tanto como clasificador, como para la representación de funciones booleanas, pues su neurona es esencialmente de tipo MacCulloch-Pitts, de salida binaria. La importancia teórica del perceptrón radica en su carácter de dispositivo entrenable, pues el algoritmo de aprendizaje del modelo introducido por Rosenblatt, y que se describirá más adelante, permite determinar automáticamente los pesos sinápticos que clasifican conjunto de patrones a partir de un conjunto de ejemplos etiquetados.

Se mostrara a continuación que un perceptrón permite realizar tareas de clasificación. Cada neurona del perceptrón representa una determinada clase, de modo que dado un vector de entrada, una cierta neurona responde con 0 si no pertenece a la que representa, y con un 1 si pertenece. Es fácil ver que una neurona tipo perceptrón solamente permite discriminar entre dos clases linealmente separables (es decir, cuyas regiones de decisión pueden ser separadas mediante una única condición lineal o hiperplano Sea una neurona tipo perceptrón de dos entradas, x_1 y x_2 con salida y , cuya operación se define por lo tanto

$$y = H(w_1x_1 + w_2x_2 - \theta)$$

o bien

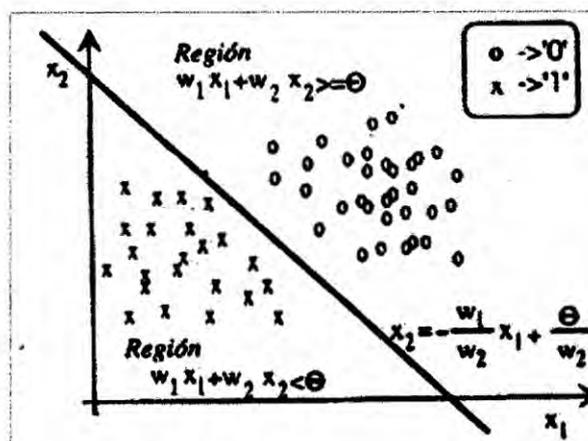
$$y = \begin{cases} 1, & \text{si } w_1x_1 + w_2x_2 \geq \theta \\ 0, & \text{si } w_1x_1 + w_2x_2 < \theta \end{cases}$$

Si consideramos x_1 y x_2 situadas sobre los ejes de abscisas y ordenadas en el plano, la condición

$$w_1x_1 + w_2x_2 - \theta \Rightarrow x_2 = -w_1x_1/w_2 + \theta/w_2$$

representa una recta (hiperplano, si trabajamos con n entradas) que divide el plano (espacio) en dos regiones, aquellas para las que la neurona proporciona una salida 0 o 1, respectivamente (véase la figura N° 2.21). Luego, efectivamente, una neurona tipo perceptrón representa un discriminador lineal, al implementar una condición lineal que separa dos regiones en el espacio, que representan dos diferentes clases de patrones.

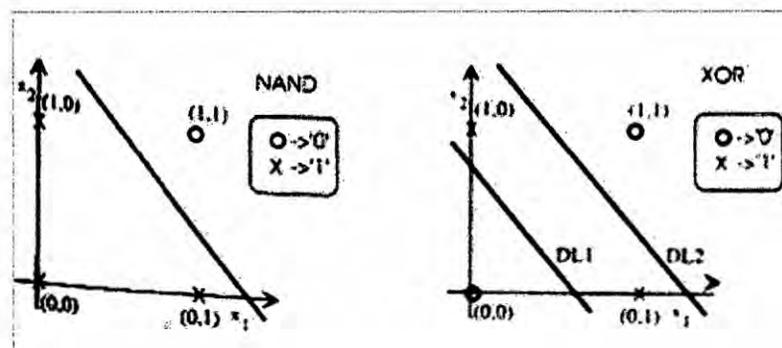
FIGURA N°2.21
REGIÓN DE DECISIÓN CORRESPONDIENTE A UN PERCEPTRÓN SIMPLE CON DOS NEURONAS DE ENTRADA



Consideremos la función lógica NAND, (AND negada de dos entradas), que representamos sobre el plano (véase la figura N° 2.22a). En este caso pueden encontrarse unos parámetros w_1 y w_2

y θ que determinen una recta que separa perfectamente las regiones correspondientes a los valores lógicos 0 y 1. Por ello, la función lógica NAND se dice separable linealmente, puesto que hemos podido encontrar una única condición lineal que divida ambas regiones. Por ejemplo, un perceptrón con los siguientes parámetros implementa la función NAND: $w_1 = w_2 = -2$ y $\theta = -3$

FIGURA 2.22:
FUNCIONES LÓGICAS NAND (a) y XOR (b)



Sin embargo consideremos la función lógica or - exclusivo o XOR (su salida es el 0 lógico si las variables de entrada son iguales y 1 si son diferentes), se representa en el plano (véase la figura N° 2.22b). En este caso podemos apreciar que no se puede encontrar una única condición lineal que separe las regiones a los valores de salida 0 y 1, por lo que se dice que la XOR no es linealmente separable. Como la neurona del perceptrón representa en el fondo un discriminador lineal, esta neurona por sí sola no puede implementar la función XOR.

Por lo tanto, concluimos con que la clase de funciones no separables linealmente no puede ser representada por un perceptrón simple.

Minsky (uno de los padres de la IA) y Papert estudiaron en profundidad el perceptrón, y en 1969 publicaron un exhaustivo trabajo en el que se subrayaba sus limitaciones, lo que resultó decisivo para que muchos de los recursos que se estaban invirtiendo en redes neuronales se desviasen hacia otros campos más prometedores entonces, como era en la época el de la inteligencia artificial. A finales de los sesenta ya se apuntaba como solución a las limitaciones del perceptrón introducir capas ocultas, pero el problema residía en que si bien se disponía de un algoritmo de aprendizaje para el perceptrón simple, el denominado algoritmo del perceptrón (se explicara el algoritmo en la siguiente sección), no se disponía de ningún procedimiento que permitiese obtener automáticamente los pesos en una multicapa, con neuronas ocultas. Este problema denominado de 'asignación de crédito' a las neuronas sin conexión directa con el exterior (consistente en cómo medir la contribución al error en la salida de la red neuronal de cada uno de los nodos ocultos que precisamente no tienen una conexión directa con ella) fue resuelto no mucho más tarde por Paul Werbos, pero fue preciso esperar hasta mediados de los años ochenta para que el grupo PDP junto con otros grupos de forma independiente) redescubriera un algoritmo similar que denominaron back-

propagation o BP, y diera a conocer a la comunidad internacional su gran potencial para la resolución de problemas prácticos.

- Algoritmo de aprendizaje del perceptrón

La importancia del perceptrón radica en su carácter de dispositivo entrenable pues el algoritmo de aprendizaje introducido por Rosenblatt permite que el perceptrón determine automáticamente los pesos sinápticos que clasifican un determinado conjunto de patrones etiquetados. El del perceptrón es un algoritmo de aprendizaje de los denominados por corrección de errores. Los algoritmos de este tipo (en el que se incluyen posteriormente también el de la adaline y el BP) ajustan los pesos en proporción a la diferencia existente entre la salida actual de la red y la salida deseada, con el objetivo de minimizar el error actual de la red.

Sea un conjunto de p patrones x^u , $u=1,\dots,p$, con sus salidas deseadas t^u . Tanto las entradas como las salidas solamente pueden tomar los valores -1 o 1 (o bien, 0 o 1 , según se definan los niveles lógicos). Se tiene una arquitectura de perceptrón simple, con pesos iniciales aleatorios, y se requiere que clasifique correctamente todos los patrones del conjunto de aprendizaje (lo cual es posible solamente si son linealmente separables).

Se decidirá del siguiente modo, ante la presentación del patrón u -ésimo, si la respuesta que proporciona el perceptrón es correcta,

no se actualizarán los pesos; si es incorrecta, según la regla de Hebb de la sección 2.2.2.g.8). Se tiene

$$\Delta w_{ij}^u(t) = \begin{cases} 2\epsilon t_i^u x_j^u, & \text{si } y_i^u \neq t_i^u \\ 0, & \text{si } y_i^u = t_i^u \end{cases}$$

que se puede reescribir del siguiente modo

$$\Delta w_{ij}^u(t) = \epsilon(t_i^u - y_i^u) x_j^u$$

que es la forma habitual de expresar la regla del perceptrón. En su utilización práctica, se debe llegar a un compromiso para el valor del ritmo de aprendizaje ϵ , puesto que un valor pequeño implica un aprendizaje lento, mientras que uno excesivamente grande puede conducir a oscilaciones en el entrenamiento, al introducir variaciones en los pesos excesivamente amplias. Al ser las entradas y las salidas discretas $\{-1, +1\}$, también lo será la actualización de los pesos, que únicamente podrá tomar los valores 0 o $\pm 2\epsilon$.

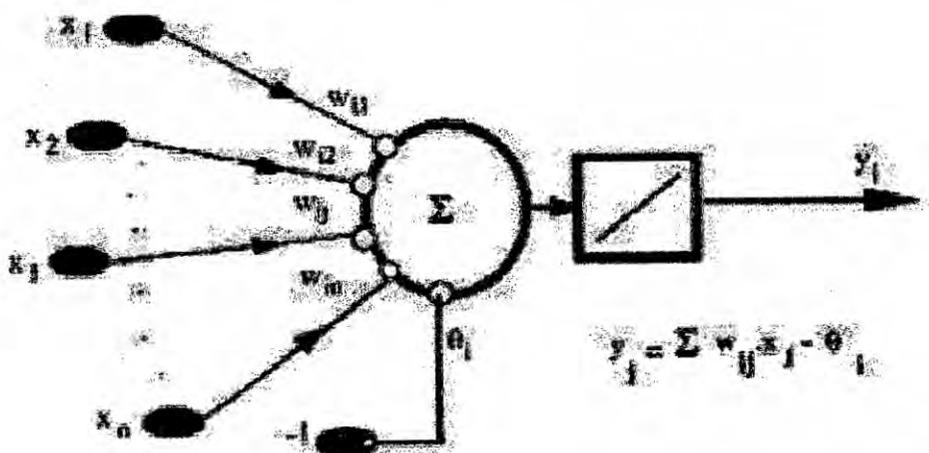
c) Adaline

Al mismo tiempo que Frank Rosenblatt trabajaba en el modelo del Perceptrón Bernard Widrow y su estudiante Marcian introdujeron el modelo de la red Adaline y su regla de aprendizaje llamada algoritmo LMS (Least Mean Square). El término Adaline es una sigla, sin embargo su significado cambió ligeramente a finales de los años

$$y_i(t) = \sum_{j=1}^n w_{ij} x_j - \theta_i$$

sesenta cuando decayó el estudio de las redes neuronales, inicialmente se llamaba Adaptive Linear Neuron (Neurona Lineal Adaptiva), para pasar después a ser Adaptive Linear Element (Elemento Lineal doptivo), este cambio se debió a que la Adaline es un dispositivo que consta de un único elemento de procesamiento, como tal no es técnicamente una red neuronal. La estructura general de la red tipo Adaline puede visualizarse en (véase la figura N°2.23). A diferencia del asociador lineal, la adaline incorpora un parámetro adicional denominado bias, el cual no debe de ser considerado como un umbral de disparo, sino como un parámetro que proporciona un grado de libertad adicional al modelo.

FIGURA 2.23
ESTRUCTURA GENERAL DE LA RED TIPO ADALINE



Teniendo en cuenta lo anterior, la ecuación de la adaline resulta ser:

con $i = 1, \dots, m$.

Otra diferencia fundamental de la adaline con respecto del asociador lineal y el perceptrón simple radica en la regla de aprendizaje. En la adaline se utiliza la regla de Widrow – Holf, también conocida como LMS (Least Mean Square) o regla de mínimos cuadrados. Esta regla permite actualizaciones de los pesos proporcionales al error cometido por la neurona.

El adaline se viene utilizando desde comienzos de los años sesenta como filtro adaptativo, por ejemplo en aplicaciones relacionadas con la reducción del ruido en la transmisión de señales. De este modo, y desde hace años, millones de modelos de todo el mundo incluyen una adaline.

d) **El perceptron multicapa (MLP)**

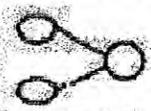
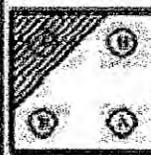
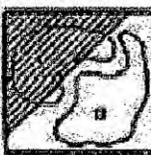
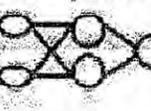
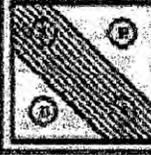
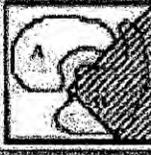
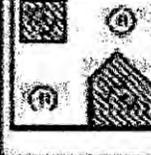
En la sección 2.2.2.g.8).b) se ha visto las limitaciones del perceptrón simple, ya que con él tan solo se puede discriminar patrones que pueden ser separados por un hiperplano (una recta en el caso de dos neuronas de entrada). Una manera de solventar estas limitaciones del perceptrón simple es por medio de la inclusión de capas ocultas, obteniendo de esta forma una red neuronal que se denomina perceptrón multicapa.

La Figura (véase la figura N°2.24) muestra las regiones de decisión que se obtienen para distintas arquitecturas de redes neuronales

considerando dos neuronas en la capa inicial. Así por ejemplo para una arquitectura de perceptrón simple la región de decisión es una recta, mientras que el perceptrón multicapa con una única capa de neuronas ocultas puede discriminar regiones convexas.

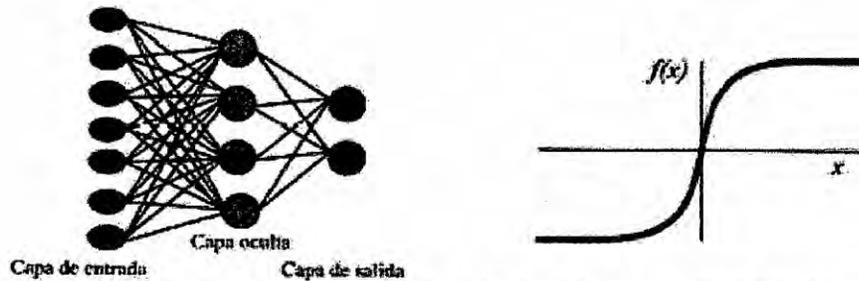
Por otra parte el perceptrón multicapa con dos capas de neuronas ocultas es capaz de discriminar regiones de forma arbitraria.

FIGURA 2.24
REGIONES DE DECISIÓN OBTENIDAS PARA EL PERCEPTRÓN SIMPLE (ARRIBA), EL PERCEPTRÓN MULTICAPA CON UNA CAPA OCULTA (EN MEDIO) Y EL PERCEPTRÓN MULTICAPA CON DOS CAPAS OCULTAS (ABAJO)

Arquitectura	Región de decisión	Ejemplo 1: XOR	Ejemplo 2: clasificación	Regiones más generales
Sin capa oculta 	Hiperplano (dos regiones)			
Una capa oculta 	Regiones polinomiales convexas			
Dos capas ocultas 	Regiones arbitrarias			

La estructura del MLP con una única capa oculta se muestra en (véase las Figuras N°2.24 y 2.25, en la página "131").

FIGURA 2.25:
ARQUITECTURA (IZQUIERDA) Y FUNCIÓN DE
DESEMPEÑO(DERECHA) PARA EL PERCEPTRÓN
MULTICAPA



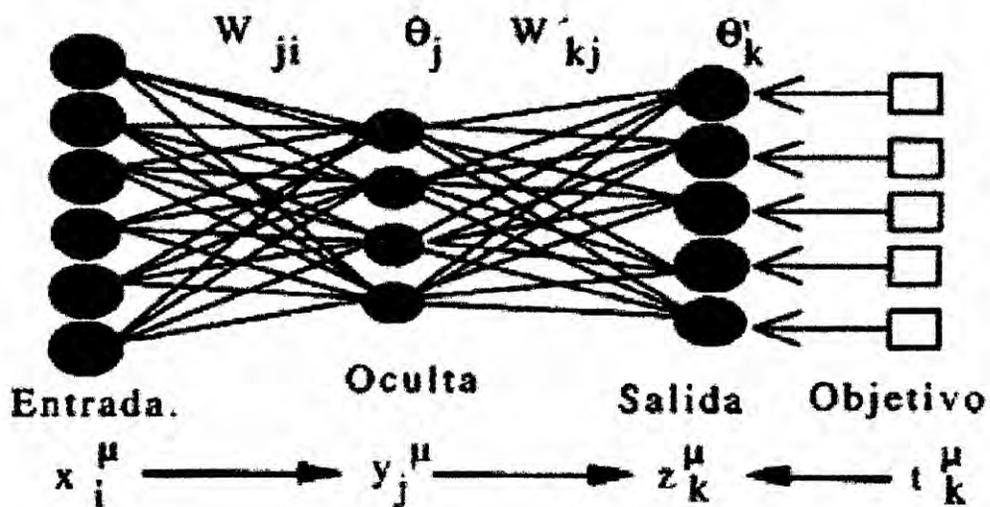
Denominaremos x_i a las entradas de la red, y_j a las salidas de la capa oculta y Z_k a las de la capa final (y globales de la red); t_k serán las salidas objetivo (target). Por otro lado, w_{ij} son los pesos de la capa oculta y j sus umbrales, w'_{kj} los pesos de la capa de salida y k sus umbrales. La operación de un MLP con una capa oculta y neuronas de salida lineal (estructura que constituye, como se verá más adelante, un aproximador universal de funciones) se expresa matemáticamente de la siguiente manera:

$$z_k = \sum_j w'_{kj} y_j - \theta'_k = \sum_j w'_{kj} f\left(\sum_i w_{ji} x_i - \theta_j\right) - \theta'_k$$

Ésta es la arquitectura más común de MLP, aunque existen numerosas variantes, como incluir neuronas no lineales en la capa de salida (solución que se adopta especialmente en problemas de clasificación), introducir más capas ocultas, emplear otras funciones de activación, limitar el número de conexiones entre una neurona y

las de la capa siguiente, introducir dependencias temporales o arquitecturas recurrentes, etc.

FIGURA 2.26:
ARQUITECTURA DEL PERCEPTRÓN MULTICAPA



- El MLP como aproximador universal de funciones

El desarrollo del MLP durante los últimos treinta años ha resultado curioso. Partiendo de un perceptrón monocapa y observando sus limitaciones computacionales, se llegó a la arquitectura perceptrón multicapa, y aplicándolo a numerosos problemas, se comprobó experimentalmente que éste era capaz de representar complejos mappings y de abordar problemas de clasificación de gran envergadura, de una manera eficaz y relativamente simple. Sin embargo, faltaba una demostración teórica que permitiese explicar sus aparentemente enormes capacidades computacionales.

Este proceso histórico comienza con McCulloch y Pitts, quienes mostraron que mediante su modelo de neurona (esencialmente un dispositivo de umbral) podría representarse cualquier función booleana; mucho más tarde, Denker y otros, demostraron que toda función booleana podía ser representada por una red unidireccional multicapa de una sola capa oculta. Por las mismas fechas, Lippmann mostró que un perceptrón con dos capas ocultas bastaba para representar regiones de decisión arbitrariamente complejas. Por otra parte, Lapedes y Farber demostraron que un perceptrón de dos capas ocultas es suficiente para representar cualquier función arbitraria (no necesariamente booleana).

Más tarde, Hecht-Nielsen aplicando el teorema de Kolmogorov demostró que una arquitectura de características similares al MLP, con una única capa oculta, resultaba ser un aproximador universal de funciones. Por fin, a finales de la década, diversos grupos propusieron casi a la par teoremas muy similares que demostraban matemáticamente que un MLP convencional, de una única capa oculta, constituía, en efecto, un aproximador universal de funciones.

Los teoremas citados resultan de vital importancia, puesto que proporcionan una sólida base teórica al campo de las redes neuronales, al incidir sobre un aspecto (la aproximación funcional

y un modelo (el MLP) centrales en la teoría de las redes neuronales artificiales. No obstante, todavía quedan muchos asuntos abiertos. Por ejemplo estos teoremas no informan sobre el número de nodos ocultos necesarios para aproximar una función determinada, simplemente se afirma que hay que colocar los necesarios para lograr el nivel de aproximación requerido.

- Aprendizaje por retropropagación de errores

Una solución al problema de entrenar los nodos de las capas ocultas pertenecientes a arquitecturas multicapa la proporciona el algoritmo de retropropagación de errores o BP (backpropagation).

Sea un MLP de tres capas, cuya arquitectura se presenta en la Figura 2.26, con [entradas, salidas, pesos y umbrales de las neuronas definidas en la introducción de la sección 2.9.4. Dado un patrón de entrada x_u , ($u=1, \dots, p$), se recuerda que la operación global de esta arquitectura se expresa del siguiente modo:

$$z_k^u = \sum_j w'_{kj} y_j^u - \theta'_k = \sum_j w'_{kj} f(\sum_i w_{ji} x_i^u - \theta_j) - \theta'_k$$

Las funciones de activación de las neuronas ocultas $f(h)$ son de tipo sigmoideo, con h el potencial postsináptico o local. La función de costo de la que se parte es el error cuadrático medio

$$E(w_{ji}, \theta_j, w'_{kj}, \theta'_k) = (1/2) \sum_{u,k} [t_k^u - f(\sum_j w'_{kj} y_j^u - \theta'_k)]^2$$

La minimización se lleva a cabo mediante descenso por el gradiente, pero en esta ocasión habrá un gradiente respecto de los pesos de la capa de salida y otro respecto de los de la oculta:

$$\delta w'_{kj} = -\varepsilon \partial E / \partial w'_{kj} \quad \delta w_{ji} = -\varepsilon \partial E / \partial w_{ji}$$

Las expresiones de actualización de los pesos se obtienen sólo con derivar, teniendo en cuenta las dependencias funcionales y aplicando adecuadamente la regla de la cadena

$$\delta w'_{kj} = \varepsilon \sum_u \Delta'_k{}^u y_j^u, \text{ con } \Delta'_k{}^u = \{ t_k^u - f(v'_k{}^u) \} \partial f(v'_k{}^u) / \partial v'_k{}^u$$

$$\delta w_{ji} = \varepsilon \sum_u \Delta_j^u x_i^u, \text{ con } \Delta_j^u = \{ \sum_k \Delta'_k{}^u w'_{kj} \} \partial f(v_j^u) / \partial v_j^u$$

La actualización de los umbrales (bias) se realiza haciendo uso de estas mismas expresiones, considerando que el umbral es un caso particular de peso sináptico, cuya entrada es una constante igual a -1.

En estas expresiones está implícito el concepto de propagación hacia atrás de los errores que da nombre al algoritmo. En primer lugar se calcula la expresión, Δ_j^u que denominaremos señal de error, por ser proporcional al error de la salida actual de la red, con el que calculamos la actualización de los pesos de la capa de

salida. A continuación se propagan hacia atrás los errores δw_{kj} , Δ_k^u, Δ_j^u a través de las sinapsis, proporcionando así las señales de error, correspondientes a las sinapsis de la capa oculta; con éstas se calcula la actualización de las sinapsis ocultas. El algoritmo puede extenderse fácilmente a arquitecturas con más de una capa oculta siguiendo el mismo esquema.

En resumen, el procedimiento a seguir para entrenar mediante BP una arquitectura MLP dada es el siguiente:

- Establecer aleatoriamente los pesos y umbrales iniciales $\{t:0\}$.
- Para cada patrón μ del conjunto de aprendizaje.
 - Llevar a cabo una fase de ejecución para obtener la respuesta de la red ante el patrón μ -ésima.
 - Calcular las señales de error asociadas Δ_k^u, Δ_j^u .
 - Calcular el incremento parcial de los pesos y umbrales debido a cada patrón μ (elemento de los sumatorios)
- Calcular el incremento total (para todos los patrones) actual de los pesos δw_{kj} y δw_{kj} . Hacer lo mismo para los umbrales.
- Actualizar los pesos y umbrales.
- Calcular el error actual. $t=t+1$. y volver a 2) si todavía no es satisfactorio.

Hacer lo mismo para los umbrales.

Se debe comenzar siempre con pesos iniciales aleatorios (normalmente números pequeños, positivos y negativos), ya que si se parte de pesos y umbrales iniciales son nulos el aprendizaje no progresará (puesto que las salidas de las neuronas y el incremento en los pesos serán siempre serán nulos).

En el esquema presentado, que surge de forma natural del proceso de descenso por el gradiente, se lleva a cabo una fase de ejecución para todos y cada uno de los patrones del conjunto de entrenamiento, se calcula la variación en los pesos debida a cada patrón, se acumulan, y solamente entonces se procede a la actualización de los pesos. Este esquema se suele denominar aprendizaje por lotes (batch).

Una variación común al algoritmo consiste en actualizar los pesos sinápticos tras la presentación de cada patrón (en vez de presentarlos todos y luego actualizar), esquema denominado aprendizaje en serie (on - line). Aunque el verdadero BP es el que se ejecuta por lotes, el aprendizaje en serie es habitualmente empleado en aquellos problemas en los que se dispone de un muy numeroso conjunto de patrones de entrenamiento (compuesto por cientos o miles de patrones), en el que habrá mucha redundancia en los datos. Si se emplease en este caso el modo por lotes, el tener que procesar todos los patrones antes de actualizar los pesos demoraría considerablemente el

entrenamiento (además de precisar el almacenamiento de numerosos resultados parciales). Por ejemplo, podemos imaginar un conjunto de entrenamiento compuesto por 10.000 patrones, en el que cada patrón aparece repetido cien veces, entrenando por lotes el aprendizaje duraría cien veces más que en modo serie.

En el aprendizaje en serie se debe tener presente que el orden en la presentación de los patrones debe ser aleatorio, puesto que si siempre se siguiese un mismo orden el entrenamiento estaría viciado en favor del último patrón del conjunto de entrenamiento, cuya actualización, por ser la última, siempre predominaría sobre las anteriores. Además, esta aleatoriedad presenta una importante ventaja, puesto que puede permitir escapar de mínimos locales en determinadas ocasiones, por lo que al final del proceso puede alcanzarse un mínimo más profundo.

EL BP tiene como ventaja principal es que se puede aplicar a multitud de problemas diferentes, proporcionando con frecuencia buenas soluciones con no demasiado tiempo de desarrollo. No obstante, si se requiere una solución realmente excelente, habrá que dedicar más tiempo al desarrollo del sistema neuronal, teniendo en cuenta diferentes cuestiones adicionales, como desventaja se encuentra, entre otras, su lentitud de convergencia, uno los precios que hay que pagar por disponer de un método

general de ajuste funcional que no requiere (en principio) información apriorística.

Sin embargo, se debe tener cuenta que el BP no requiere tanto esfuerzo computacional como el que sería necesario si se tratasen de obtener los pesos de la red mediante la evaluación directa de las derivadas; en ese sentido se ha comparado el BP con la transformada rápida Fourier, que permite calcular la transformada de Fourier con un muy inferior esfuerzo computacional. Otro problema del BP es que puede incurrir en el denominado sobreaprendizaje (o sobreajuste), fenómeno directamente relacionado con la capacidad de generalización de la red a partir de los ejemplos presentados. Por otra parte, debe tenerse en cuenta que el algoritmo BP no garantiza alcanzar el mínimo global de la función error, sólo un mínimo local, por lo que el proceso de aprendizaje puede quedarse estancado en uno de estos mínimos locales.⁷

9) Capacidad de generalización de la red

Uno de los aspectos fundamentales de los ANS es su capacidad de generalizar a partir de ejemplos, lo que constituye el problema de la memorización frente a generalización. Por generalización se entiende

⁷ Mas información sobre pequeñas variaciones al algoritmo de backpropagation en la dirección <http://www.iti.upv.es/~fcn/students/ma/Index.html>

la capacidad de la red de dar una respuesta correcta ante patrones que no han sido empleados en su entrenamiento.

Una red neuronal correctamente entrenada generalizará, lo que significa que ha aprendido adecuadamente el mapping no sólo los ejemplos concretos presentados, por lo que responderá correctamente ante patrones nunca vistos con anterioridad.

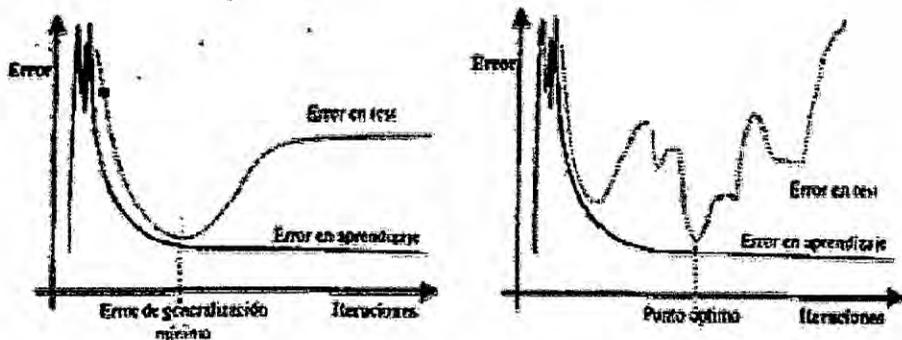
a) Validación cruzada (cross-validation)

En un proceso de entrenamiento se debe considerar, por una parte, un error de aprendizaje, que se suele calcular como el error cuadrático medio de los resultados proporcionados por la red para el conjunto de patrones de aprendizaje. Con una red suficientemente grande, puede reducirse tanto como se quiera sólo con llevar a cabo más iteraciones. Por otra parte, existe un error de generalización, que se puede medir empleando un conjunto representativo de patrones diferentes a los utilizados en el entrenamiento. De esta manera, podemos entrenar una red neuronal haciendo uso de un conjunto de aprendizaje, y comprobar su eficiencia real, o error de generalización, mediante un conjunto de test.

Un hecho experimental, fácilmente observable con cualquier simulador, es que si se entrena una red hasta alcanzar un muy pequeño error en aprendizaje (por ejemplo, inferior a un 1%), la eficacia real del sistema o generalización (medido como error en test) se degrada.

Si representamos a la vez el error en aprendizaje y el error en test durante el transcurso del aprendizaje, se obtiene una gráfica como la representada en (véase la Figura N°2.27, izquierda): tras una fase inicial, en la que pueden aparecer oscilaciones en el valor del error, el de aprendizaje tiende a disminuir monótonamente, mientras que el error de generalización a partir de cierto punto comienza a incrementarse, lo que indica una degradación progresiva del aprendizaje.

FIGURA 2.27
EVOLUCIÓN DEL ERROR DE APRENDIZAJE Y DEL ERROR DE GENERALIZACIÓN. A LA IZQUIERDA, SITUACIÓN IDEALIZADA, A LA DERECHA SITUACIÓN REAL.



La explicación de este fenómeno es la siguiente. Al principio la red se adapta progresivamente al conjunto de aprendizaje, acomodándose al problema y mejorando la generalización. Sin embargo, en un momento dado el sistema se ajusta demasiado a las particularidades de los patrones empleados en el entrenamiento, aprendiendo incluso el ruido en ellos presente, por lo que crece el

error que comete ante patrones diferentes a los empleados en el entrenamiento (error de generalización). En este momento la red no ajusta correctamente el mapping, sino que simplemente está memorizando los patrones del conjunto de aprendizaje, lo que técnicamente se denomina sobreaprendizaje o sobreajuste (overtraining o overfitting), pues la red está aprendiendo demasiado (incluso el ruido presente en los patrones-ejemplo). Idealmente, dada una arquitectura de red neuronal, ésta debería entrenarse hasta un Punto óptimo (Figura 2.27, izquierda) en el que el error de generalización es mínimo. El procedimiento consistente en entrenar y validar a la vez para detenerse en el punto óptimo se denomina validación cruzada (cross validation), y es ampliamente utilizado en la fase de desarrollo de una red neuronal supervisada (como el MLP).

No obstante, la situación descrita ha sido en cierta medida idealizada; una situación más realista sería la de la parte derecha de la Figura 2.27: en realidad pueden presentarse varios mínimos para el conjunto de test, debiendo detener el aprendizaje en el punto óptimo de mínimo error de generalización, y no quedarnos en el primer mínimo en test que aparezca.

Muchas veces basta con dejar que el aprendizaje discurra hasta una cota razonable de error (0.5%, 0.1%, 0.01%..., depende el problema), guardando periódicamente las distintas configuraciones

intermedias de pesos para luego quedarnos con la que proporcionó un error en test mínimo.

La clave de este asunto está en que las redes neuronales son estimadores no lineales poderosos, capaces de modelar situaciones muy complejas. En las herramientas lineales (por ejemplo, en ajuste mediante polinomios) la complejidad del modelo viene dada simplemente por el número de parámetros libres a ajustar (coeficientes), mientras que en los sistemas no lineales (como son las redes neuronales), la complejidad del modelo depende tanto del número de parámetros como de su valor actual. Los modelos que implementan las redes neuronales son de elevada complejidad, por lo que pueden aprender (memorizar) casi cualquier cosa, motivo por el cual incurren fácilmente en sobre aprendizaje.

Cuando se entrena una red unidireccional supervisada debe tenerse muy en cuenta el tema del sobreaprendizaje, y la técnica de validación cruzada suele ser un buen remedio (aunque, como se verá a continuación, no es el único disponible); usualmente, de todo el conjunto de entrenamiento se emplea aproximadamente un 80% de los patrones para entrenar, reservándose un 20% como conjunto de test.

b) Número de ejemplos de entrenamiento

En definitiva, la capacidad de generalización de la red la determinan en buena medida las siguientes tres circunstancias: 1) la arquitectura

de la red, 2) el número de ejemplos de entrenamiento y 3) la complejidad del problema. Los tres puntos están muy relacionados; en términos generales, cuanto más complejo sea el problema a modelar más grande deberá ser la red (con más parámetros a ajustar) y, por lo tanto, más ejemplos se necesitarán para entrenarla (ejemplos que deberán cubrir todo el espacio de entrada, contemplando todas las situaciones posibles).

A menudo el número de patrones-ejemplo disponibles es limitado (y reducido), y en proporción el número de parámetros efectivos de la red elegida (grados de libertad) suele ser muy grande. Así, si se quiere que la red alcance un error de generalización de, por ejemplo, $\epsilon = 0.1$ (un 10%), el número de patrones de aprendizaje necesarios p será del orden de $p = 10w$, expresión que se suele dar como indicativa del número aproximado de patrones que serán necesarios para entrenar adecuadamente una red neuronal de w pesos.

Por ejemplo para una red 10-5-1 (10 neuronas de entrada, 5 ocultas y 1 de salida.), que dispone de 61 parámetros, entre pesos y umbrales, el número de patrones necesarios para alcanzar un error del 10% será de unos 610, lo que representa una cifra de patrones muy alta, no disponible en muchas aplicaciones prácticas. Ello ilustra de nuevo la facilidad de incurrir en sobre aprendizaje al entrenar una red neuronal.

c) Reducción del tamaño de la arquitectura de una red

Además, hay que tener presente la llamada maldición de la dimensionalidad (curse of dimensionality) que consiste en que el número de datos necesarios para especificar un mapping, en general crece exponencialmente con la dimensión del espacio de entrada, lo que agrava en los problemas de dimensión de entrada elevada es el de disponer de un número de patrones para el aprendizaje escaso.

Disminuyendo el número de parámetros de la red (tamaño) se tendrá una relación $p = w/?$ más favorable. Una forma de reducirlo consiste en limitar el número de las entradas de la red, pues ello implica la disminución drástica del número de pesos.

Por ejemplo, una red con 200 entradas, 100 neuronas ocultas y 3 salidas, contendrá del orden de 20.000 pesos, con lo que se necesitarían unos 200.000 patrones para entrenarla adecuadamente. Si reducimos el número de entradas a 10 (por ejemplo, realizando un análisis de componentes principales a las variables de entrada, empleando ratios, etc.), el número de pesos se reduce a 143, con lo que se precisarían únicamente unos 1400 patrones de aprendizaje.

Otras técnicas empleadas en la reducción del número de parámetros de la red se relacionan con eliminar algunos de sus pesos; algunas

bien conocidas son las de compartir pesos (weight sharing), podado de la red (pruning) o decaimiento de pesos (weight decay).

En la primera de las citadas, diversas neuronas comparten sus pesos, de modo que el número total disminuye. En el proceso de podado la red es entrenada hasta un cierto nivel, para luego eliminar aquellos pesos que no aportan prácticamente nada a su operación. El decaimiento es un caso especial del podado; durante el aprendizaje se deja a los pesos tender poco a poco a cero, para que aquellos que no sean actualizados periódicamente, se anulen y desaparezcan.

h. Medida de error

De los estudios de la medida de error de las tesis de referencia el que presenta una mejora en los modelos de referencia es de la referencia bibliográfica [4].

El criterio fundamental para evaluar el grado de precisión de un modelo de predicción se basa en examinar las diferencias entre las producciones predichas y las reales. Para cada horizonte de predicción h se define el error de predicción como

$$e_{t+h,t} = P_{t+h} - P_{t+h,t} \quad (2.14)$$

Para independizar esta medida del tamaño del parque eólico, se normaliza como porcentaje de la potencia nominal del mismo si los datos no han sido previamente normalizados.

$$Q_{t+h,t} = \frac{P_{t+h} - P_{t+h,t}}{P_{\text{instalada}}} \quad (2.15)$$

Una vez definidos los errores de predicción, existen diversos estadísticos que se pueden utilizar para caracterizar el error para cada horizonte de predicción h . Los que se van a utilizar en esta tesis y que son los comúnmente empleados en predicción eólica son:

- Sesgo: Es una medida del error sistemático y se define como

$$\text{sesgo}_{t+h,t} = \bar{e}_{t+h} = \bar{e}(h) = \frac{1}{N_h} \sum_{t=1}^{N_h} e_{t+h,t} \quad (2.16)$$

Siendo N_h el número de datos de error.

- Sesgo normalizado

$$\text{sesgo_normalizado}_{t+h,t} = \bar{q}_{t+h} = \bar{q}(h) = \frac{1}{N_h} \sum_{t=1}^{N_h} q_{t+h,t} \quad (2.17)$$

- Error Medio Absoluto (EMA)

$$\text{EMA}_{t+h,t} = \frac{\sum_{t=1}^{N_h} |e_{t+h,t}|}{N_h} \quad (2.18)$$

- Error Medio Absoluto Normalizado (EMAN) o Error de Estado

$$\text{EMAN}_{t+h,t} = \frac{\sum_{t=1}^{N_h} |q_{t+h,t}|}{N_h} \quad (2.19)$$

- Error Cuadrático Medio (ECM)

$$\text{ECM}_{t+h,t} = \frac{\sum_{t=1}^{N_h} (e_{t+h,t})^2}{N_h} \quad (2.20)$$

- Error Cuadrático Medio Normalizado (ECMN)

$$\text{ECMN}_{t+h,t} = \frac{\sum_{t=1}^{N_h} (q_{t+h,t})^2}{N_h} \quad (2.21)$$

- Raíz del Error Cuadrático Medio (RECM)

$$\text{RECM}_{t+h,t} = \sqrt{\frac{\sum_{t=1}^{N_h} (e_{t+h,t})^2}{N_h}} \quad (2.22)$$

- Raíz del Error Cuadrático Medio Normalizado (RECMN)

$$\text{RECMN}_{t+h,t} = \sqrt{\frac{\sum_{t=1}^{N_h} (Q_{t+h,t})^2}{N_h}} \quad (2.23)$$

Estadísticamente, el sesgo está asociado con el momento de orden uno del error de predicción, por lo que es una medida relacionada directamente con la energía producida.

El error cuadrático medio está asociado con el momento de orden dos, y por lo tanto con la varianza del error de predicción. En este último estadístico, los mayores errores de predicción tienen mayor influencia.

Es interesante poder comparar la mejora que representa uno de los modelos desarrollados frente a los modelos que se emplean como referencia. La ganancia, medida como mejora respecto al modelo de referencia, se denota como

$$\text{mejora}_{\text{ref,CE}}(h) = \frac{\text{CE}_{\text{ref}}(h) - \text{CE}(h)}{\text{CE}_{\text{ref}}(h)} \quad (2.24)$$

donde CE denota Criterio de Evaluación, que puede ser EMAN, ECMN, o bien, RECMN

2.3 Definición de términos básicos

2.3.1 **Aceleración de la gravedad,** es la aceleración que produce la fuerza gravitatoria terrestre; su valor en la superficie de la Tierra es,

aproximadamente, de $9,8 \text{ m/s}^2$. En la superficie de la Tierra el valor de esta aceleración, que se indica con la letra g, varía ligeramente de un lugar a otro.

En el ecuador, la aceleración de la gravedad es de 9,7799 metros por segundo cada segundo.

2.3.2 Densidad del aire, En física y química, la **densidad**(símbolo **ρ**) es una magnitud escalar referida a la cantidad de masa contenida en un determinado volumen de aire. La **densidad media** es la razón entre la masa de un cuerpo y el volumen que ocupa.

2.3.3 Humedad, La cantidad de vapor de agua presente en el aire. Se puede expresar de forma absoluta mediante la humedad absoluta, o de forma relativa mediante la humedad relativa o grado de humedad. La humedad relativa es la relación porcentual entre la cantidad de vapor de agua real que contiene el aire y la que necesitaría contener para saturarse a idéntica temperatura.,

2.3.4 Nivel de rugosidad del terreno, define clase de rugosidad a partir de la longitud de rugosidad medida en metros, z_0 , es decir, la altura sobre el nivel del suelo donde la velocidad del viento es teóricamente cero.

2.3.5 Potencia eléctrica, La **potencia eléctrica** es la relación de paso de energía de un flujo por unidad de tiempo; es decir, la cantidad de energía entregada o absorbida por un elemento en un tiempo

determinado. La unidad en el Sistema Internacional de Unidades es el vatio (watt).

2.3.6 Presión atmosférica, La **presión atmosférica** es la fuerza por unidad de superficie que ejerce el aire sobre la superficie terrestre.

2.3.7 Temperatura, La **temperatura** es una magnitud referida a las nociones comunes de caliente, tibio o frío que puede ser medida con un termómetro. En física, se define como una magnitud escalar relacionada con la energía interna de un sistema termodinámico. Más específicamente, está relacionada directamente con la parte de la energía interna conocida como «energía cinética». A medida de que sea mayor la energía cinética de un sistema, se observa que éste se encuentra más «caliente»; es decir, que su temperatura es mayor.

2.3.8 Viento, El viento es el movimiento del aire que está presente en la atmósfera, especialmente, en la troposfera.

2.3.9 Velocidad del viento, el viento viene definido por dos parámetros esenciales que son, su dirección y su velocidad. El **viento** es el flujo de gases a gran escala. En la Tierra, el viento es el movimiento en masa del aire en la atmósfera en movimiento horizontal. Günter D. Roth lo define como «la compensación de las diferencias de presión atmosférica entre dos puntos».

2.3.10 Potencia nominal, Potencia nominal es aquella potencia para la cual está diseñada la máquina, teniendo en cuenta su temperatura

y el trabajo continuo durante su tiempo de uso (funcionamiento); la Potencia Nominal del Parque eólico de Piura tendrá una Potencia Nominal de 60MW.

CAPITULO III

VARIABLES E HIPÓTEISIS

3.1 Definición de las variables

Relacionando las variables relevantes que intervienen en el presente problema objeto de estudio, que conllevarán a la explicación, demostración y probación de la formulada hipótesis, se han identificado las siguientes variables:

- a. Variable X =Parque eólico de 60 MW en Piura.
- b. Variable Y=Modelamiento de la potencia de inyección.

3.2 Operacionalización de las variables

El grado más elevado de los referentes empíricos la determinamos operacionalizando, las variables que se simbolizan con sus propiedades concretas con las letras X y Y, de igual manera, se obtiene los siguientes indicadores:

- a. **Variable X** = Parque eólico de 60 MW en Piura.

Indicadores:

- Potencia nominal X₁
- Temperatura X₂
- Densidad del aire X₃
- Velocidad del viento X₄
- Nivel de rugosidad del terreno X₅
- Presión a nivel del mar X₆

- Aceleración de la gravedad X₇
- Área del rotor eólico X₈
- Humedad X₉

b. **Variable Y** = Modelamiento de la potencia de inyección

Indicadores: Componentes

- Temperatura variable por la altura Y₁
- Densidad del viento variante con la altura Y₂
- Velocidad del viento variante con la rugosidad Y₃
- Predicciones de la velocidad del viento Y₄
- Potencia de salida del aerogenerador Y₅
- Potencia de salida del parque eólico Y₆

TABLA N° 3.1
VARIABLES DIMENSIONES E INDICADORES

VARIABLE	DIMENSIONES	INDICADORES
Variable independiente: Parque eólico de 60 MW en Piura	Conceptos fundamentales para el modelamiento de la potencia de inyección parque eólico.	- Potencia nominal(MW) X₁ - Temperatura (°C) X₂ - Densidad del aire X₃ - Velocidad del viento(m/s) X₄ - Nivel de rugosidad del terreno X₅ - Presión a nivel del mar X₆ - Aceleración de la gravedad X₇ - Área del rotor eólico X₈ - Humedad X₉
	programas que se emplean para simular la potencia de salida de un parque eólico	✓ Potencia de salida 1 utilizando RETScreen software 1. ✓ Potencia de salida por cálculos matemáticos y formulas en Excel.

VARIABLE	DIMENSIONES	INDICADORES
Variable dependiente: Modelamiento de la potencia de inyección	Revisar los modelos lineales y no lineales que se emplean para la predicción de la potencia de un parque eólico con intervalos de confianza	<ul style="list-style-type: none"> ✓ % de error de la velocidad del viento de los modelos lineales y no lineales. ✓ Coeficiente de mejora de la predicción de la velocidad del viento.
	Revisar los modelos de predicción de la potencia en países que lideran la producción de energía eólica.	<ul style="list-style-type: none"> ✓ Predicción de la potencia de España.

3.3 Hipótesis general y hipótesis específica

3.3.1 Hipótesis general: Es posible el modelamiento de la potencia de inyección de un parque eólico de 60MW ubicado en la ciudad de Piura.

3.3.2 Hipótesis específicas:

- a. Es posible la investigación de los modelos de predicción de la potencia en países que lideran la producción de energía eólica.
- b. Es posible la identificación de los conceptos fundamentales para el modelamiento de la potencia de inyección de un parque eólico.
- c. Es posible la identificación de los programas que se emplean en la simular la potencia de salida de un parque eólico.
- d. Es posible la revisión de los modelos lineales y no lineales que se emplean para la predicción de la potencia de un parque eólico con intervalos de confianza.

CAPITULO IV

METODOLOGÍA

4.1 Tipo de investigación

El tipo de investigación es Descriptiva prospectiva, porque está basado a la recolección de datos y características de los fenómenos meteorológicos para poder determinar la predicción de la velocidad del viento. Es una investigación correlacional porque permite encontrar la relación entre la velocidad del viento y la potencia de generación eólica.

4.2 Diseño de la investigación

Las etapas para el desarrollo de la presente investigación son las siguientes:

- a. Recolección de datos de la velocidad de los vientos en Piura, se tomara como fuente de información de la empresa el mismo **que es confidencial**.
- b. Comparación de datos de las velocidad de los vientos en Piura.
- c. Método matemático sipreolico para la predicción del viento, apoyados por la página web del SENAMHI y el IGP.
- d. Software para la predicción del viento. Se utilizara RNA, para la predicción de la velocidad del viento.
- e. Método matemático (combinación de RNA con el modelo de predictores de medias móviles), mejorando los errores de

predicción de los dos métodos anteriores para la predicción del viento.

- f. Software (Retscreen) y cálculo matemático para determinar la potencia de generación eólica.

4.3 Población y muestra

Por la naturaleza de esta investigación no se requiere técnica de población y muestra.

4.4 Técnicas e instrumentos de recolección de datos

Por la naturaleza de esta investigación no se requiere técnica de Técnicas e instrumentos de recolección de datos.

4.5 Procedimientos de recolección de datos.

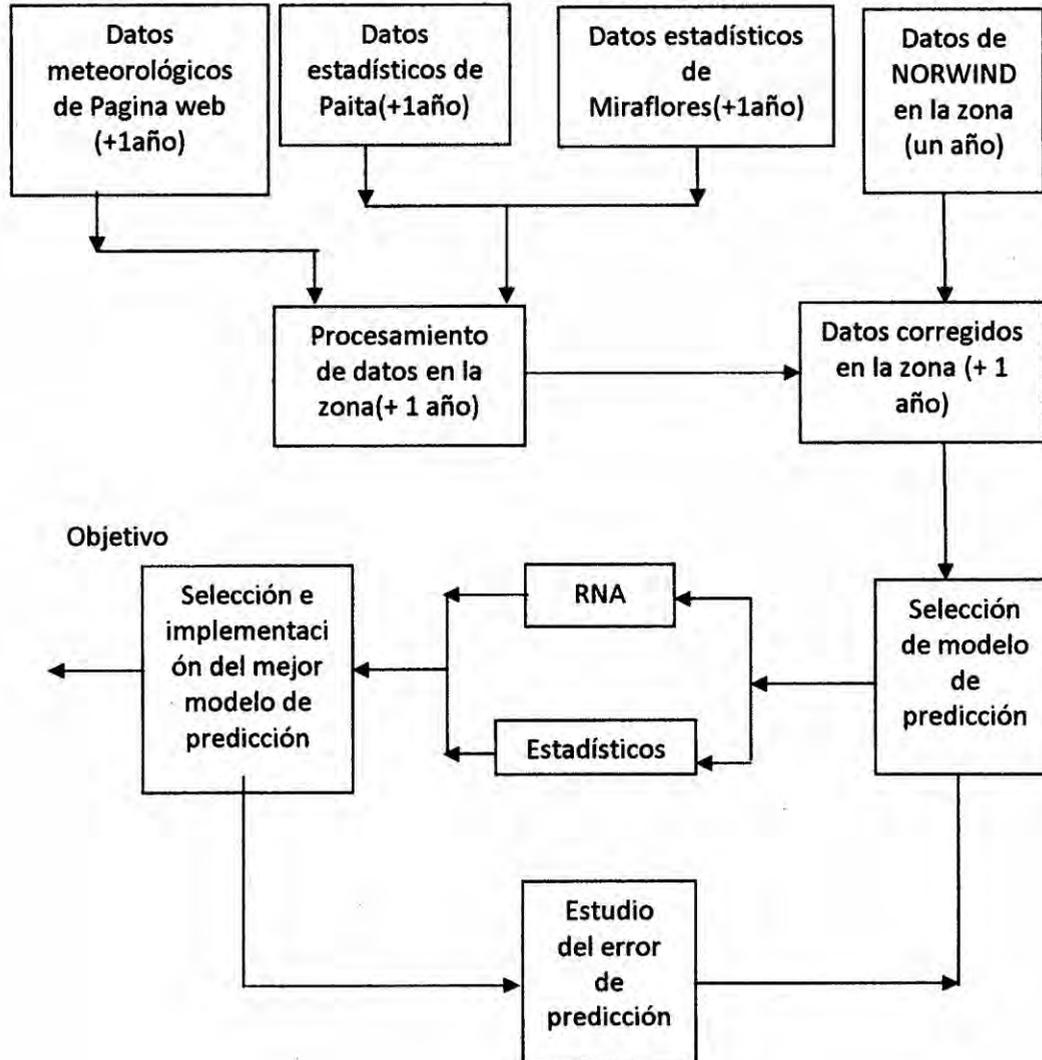
La recolección de datos se realizó mediante coordinaciones con el Instituto Geofísico del Perú (IGP), el SENAMHI y la empresa NORWIND.

4.6 Procesamiento estadístico y análisis de datos.

El procesamiento estadístico se realizara de acuerdo al primer paso de la estructura de la metodología y el análisis de datos se realizara de acuerdo al segundo paso de la estructura de la metodología.

4.7 Metodología de trabajo

La estructura de la metodología que se propone es la siguiente



4.8 Matlab Toolbox

4.8.1 Introducción

El objetivo de este capítulo es explicar el uso del toolbox de matlab para entrenar redes neuronales con el algoritmo backpropagation de tipo feedforward, para resolver problemas específicos. Generalmente se tienen cuatro pasos en el proceso de dicho entrenamiento:[24]

- a. Configurar los datos de entrenamiento
- b. Crear la red
- c. Entrenar la red
- d. Simular la respuesta de la red con nuevas entradas

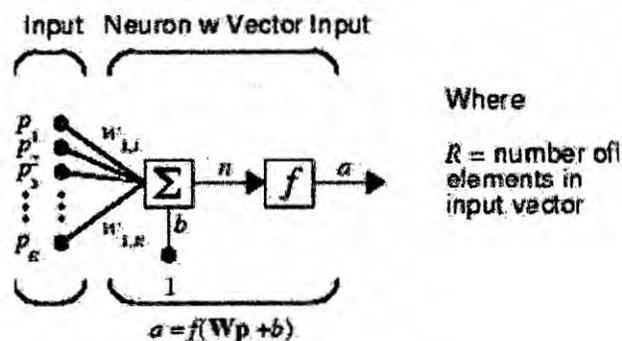
Este capítulo se muestran varias de las funciones de entrenamiento, y para cada función se deben seguir estos cuatro pasos.

4.8.2 Backpropagation

La arquitectura de la red que normalmente es la más usada con el algoritmo del backpropagation es feedforward de múltiples capas (multilayer).

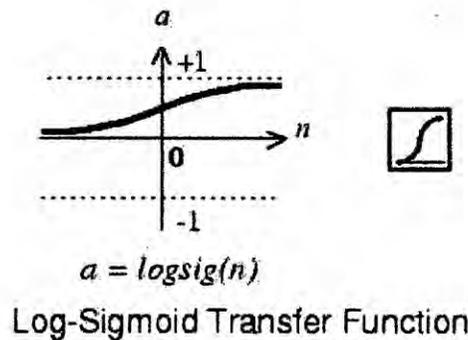
Una neurona elemental con R entradas se muestra en (véase el grafico N°4.1). Cada entrada es pesada con su w correspondiente. La suma de las entradas y sus pesos forman la entrada de a la función de transferencia f. las Neuronas pueden usar diferentes funciones de transferencia f para generar una salida.

FIGURA N° 4.1
NEURONA CON R ENTRADAS



En una red de múltiples capas (multilayer) a menudo se usa el log-sigmoid o función de transferencia logarítmica (véase la figura N°4.2).

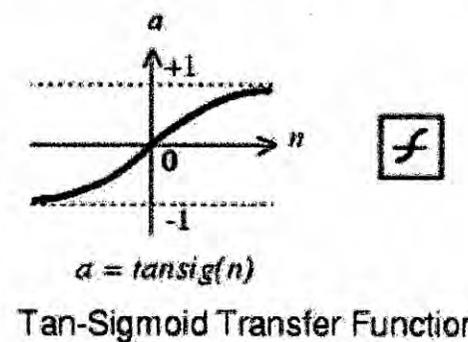
FIGURA 4.2
FUNCIÓN DE TRANSFERENCIA LOGARÍTMICA



La función logarítmica genera salidas entre 0 y 1 y la entrada de la neurona puede ir desde infinitamente negativo a infinitamente positivo.

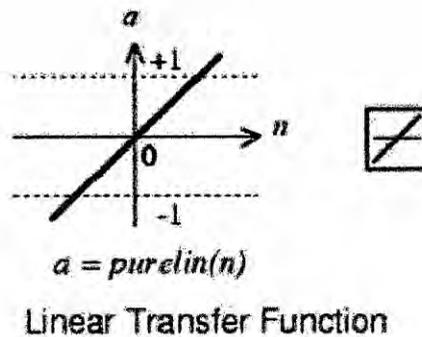
Alternativamente, las redes multicapa pueden usar el tan-sigmoid o función de transferencia tangencial. (véase la figura N°4.3)

FIGURA N°4.3
FUNCIÓN DE TRANSFERENCIA TANGENCIAL



De vez en cuando, el purelin o función de traslado lineal se usa en redes backpropagation. (véase la figura N°4.4)

FIGURA 4.4
FUNCION DE TRASLADO LINEAL(PURELIN)



Si la última capa de una red multicapa tiene neuronas sigmoideas (-sigmoid), entonces se escalan las salidas de la red en un rango pequeño.

Si se usan neuronas de salida lineal, las salidas de la red pueden tomar cualquier valor.

En backpropagation es importante poder calcular el derivado de cualquier función de transferencia utilizada. Cada una de las funciones anteriores, tangencial, logarítmica, y purelin, tienen una función derivativa que les corresponde: dtansig, dlogsig, y dpurelin. Para obtener el nombre de la función derivativa asociada a cada función de transferencia se debe llamar la función de transferencia y la función de matlab 'deriv'.

Ejemplo:

- tansig('deriv')

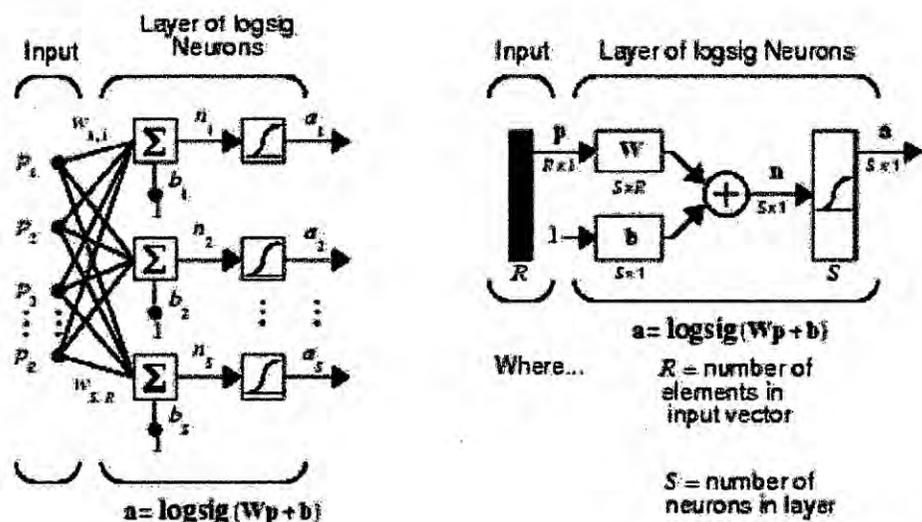
- ans = dtansig

Las tres funciones de transferencia descritas anteriormente son las usadas normalmente para backpropagation, pero pueden crearse otras funciones diferentes y usarse con backpropagation si así se desea.

4.8.3 Red FEEDFORWARD

Una red neuronal de una capa con S neuronas logsig y con R entradas se muestra en la figura 4.5, se encuentra de forma general en el lado izquierdo y en el lado derecho un diagrama específico de la capa.

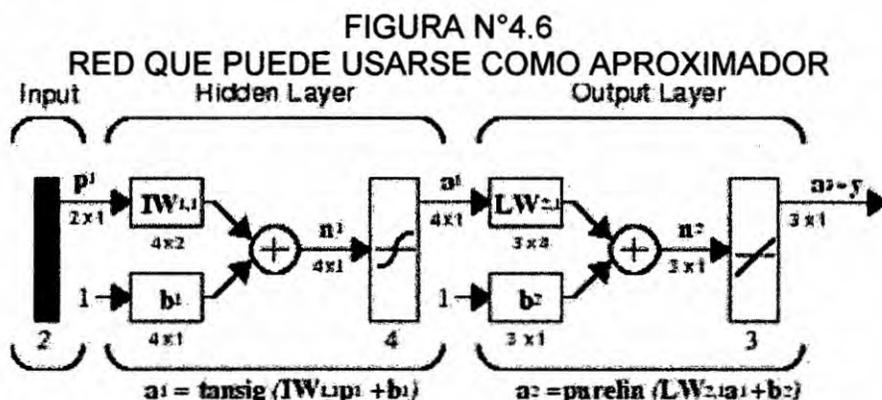
FIGURA 4.5
RED NEURONAL FEEDFORWARD



Una red Feedforward a menudo tiene una o más capas ocultas de neuronas de tipo sigmoideas, seguidas por una capa de salida lineal. Las capas múltiples de neuronas con funciones de transferencia no lineal permiten a la red aprender relaciones lineales y no lineales entre la

entrada y la salida. La capa de salida lineal permite a la red producir el umbral fuera del rango entre -1 y +1.

Por otro lado, si se quiere reprimir las salidas de una red (entre 0 y 1), entonces la capa de salida debe usar una función de transferencia sigmoidea (como logsig). Para las redes de múltiples capas se acostumbra que el número de las capas determine el exponente en la matriz de pesos.



Esta red puede usarse como un aproximador de una función en general. Puede aproximar cualquier función con un número finito de discontinuidades arbitrariamente bien, con la cantidad adecuada de neuronas en la capa oculta.

4.8.4 Creando una red (newff).

El primer paso de entrenamiento de una red feedforward es crear la red. La función newff crea una red feedforward. Esta función requiere cuatro parámetros de entrada. La primera es un R que consta de una matriz del mínimo y máximo valor para cada uno de los elementos de R en el vector de entrada. La segunda entrada es un arreglo que contiene los tamaños

de cada capa (la cantidad de neuronas por capa). La tercera entrada es un arreglo que contiene los nombres de las funciones de entrenamiento que serán usadas en cada capa. La última contiene el nombre de la función de entrenamiento que se usara.

Por ejemplo, lo siguiente crea una red de la dos capas. Hay un vector de la entrada con dos elementos. Los valores para el primer elemento del vector de entrada son - 1 y 2, los valores del segundo elemento del vector de entrada son 0 y 5. Hay tres neuronas en la primera capa y una neurona en la segunda capa (capa de salida).

La función de transferencia en la primera capa es tan-sigmoid, y en la capa de salida la función de transferencia es lineal. La función de entrenamiento es *traingd* (qué se describirá mas adelante).

- `net=newff([-1 2; 0 5],[3, 1],{'tansig','purelin'},'traingd');`

Esta orden crea la red y también inicializa los pesos y el umbral de la red; por consiguiente la red está lista para ser entrenada.

Inicializando pesos (*init*). Antes de entrenar una red feedforward, los pesos y el umbral debe inicializarse. El método *newff* inicializará los pesos automáticamente, pero si se quieren reinicializar, esto puede hacerse con *init*. Esta función toma una red como entrada y reestablece todos los pesos y el umbral de la red. Así es inicializa una red (o reinicializada):

- `net = init(net);`

4.8.5 Simulación (sim)

La función `sim` simula una red. `sim` toma la entrada de la red `p`, y la red como tal, y esta retorna las salidas de la red. Así se puede simular la red que se creó anteriormente para un solo vector de la entrada:

- `p = [1;2];`
- `a = sim(net,p)`
- `a = -0.1011`

(Si prueba estos datos, su resultado puede ser diferente, dependiendo del estado del generador de números aleatorios que tenga cuando la red fue inicializada.).

4.8.6 Entrenamiento

Una vez se han inicializado los pesos de la red y los el umbral, la red está lista para ser entrenada. La red puede entrenarse para: la aproximación de una función (regresión no lineal), la asociación del modelo, o la clasificación del modelo. El proceso de entrenamiento requiere de los patrones de conducta apropiados para la red, las entradas de la red `p` y las salidas en blanco `t`. Durante el entrenamiento los pesos y el umbral de la red son iterativamente ajustados para minimizar la función de desempeño de la red `net.performFcn`. La función de desempeño predefinida para las redes `feedforward` es `MSE`, el promedio cuadrado del error entre los rendimientos de la red y los rendimientos designados `t`.

Todos los algoritmos usan la función de desempeño para determinar cómo ajustar los pesos y minimizar `performance`.

Hay dos maneras diferentes en las que este algoritmo de descenso de gradiente puede llevarse a cabo: modo incremental y modo del lote. En el modo de incremental, el gradiente se computa y los pesos se actualizan después de cada entrada que se aplica a la red. En el modo del lote todas las entradas se aplican a la red antes que los pesos se actualicen. A continuación se describe el modo de entrenamiento por lote.

a. Entrenamiento por lote (train)

En modo por lote se actualizan sólo los pesos y el umbral de la red después que el juego de entrenamiento entero se ha aplicado a la red. Se suman los gradientes calculados en cada ejemplo de entrenamiento para determinar el cambio en los pesos y el umbral.

b. Entrenamiento por incremental (traingd)

El entrenamiento de disminución de gradiente o por pasos se ejecuta con la función `traingd`. En la que se actualizan los pesos y el umbral en dirección negativa del gradiente de la función de desempeño. Si se quiere entrenar una red que use incremental por pasos de lote, se debe anteponer `trainFcn` al `traingd`, y entonces llama al entrenamiento de la función.

Hay siete parámetros de entrenamiento asociados con `traingd`: `epochs`, `show`, `goal`, `time`, `min_grad`, `max_fail`, y `lr`. El `lr` da la proporción de aprendizaje se obtiene multiplicando tiempos por negativo del gradiente para determinar los cambios a los pesos y el umbral. Si la proporción de aprendizaje se hace demasiado grande, el algoritmo se vuelve

inestable. Si la proporción de aprendizaje se fija demasiado pequeña, el algoritmo toma un largo tiempo para converger.

El estado de entrenamiento se muestra para cada iteración. (Si `show` se inicializa `NaN`, entonces el estado de entrenamiento nunca se muestra.) Los otros parámetros determinan cuando para el entrenamiento. El entrenamiento se detiene si el número de iteraciones excede `epochs`, si los decrementos de la función de desempeño están por debajo de `goal`, si la magnitud del gradiente está en menos del `mingrad`, o si el tiempo de entrenamiento es más largo que `time`. Y `max_fail` que es asociado con la técnica de detención temprana.

El siguiente código crea un entrenamiento de entradas `p` y objetivos `t`. se entrenara por lotes, todos los vectores de la entrada se ponen en una matriz.

- `p = [-1 -1 2 2;0 5 0 5];`
- `t = [-1 -1 1 1];`

Luego, se crea la red feedforward.

- `net=newff(minmax(p),[3,1],{'tansig','purelin'},'traingd');`

Si se quiere modificar algunos de los parámetros de entrenamiento predefinidos se hace así.

- `net.trainParam.show = 50;`
- `net.trainParam.lr = 0.05;`
- `net.trainParam.epochs = 300;`
- `net.trainParam.goal = 1e-5;`

Si se quiere usar los parámetros de entrenamiento predefinidos, las órdenes anteriores no son necesarias.

Ahora entrenar la red.

- `[net,tr]=train(net,p,t);`
- `traingd, Performance goal met.`

Los entrenamientos graban el `tr` que contiene la información sobre el progreso de entrenamiento.

Ahora la red especializada puede simularse para obtener su respuesta a las entradas en el entrenamiento.

- `a = sim(net,p)`

c. Descenso de gradiente por lotes con momentumn (traingdm).

Además del `traingd`, hay otro algoritmo de descenso de gradiente de lotes para redes feedforward que proporciona una convergencia más rápida, el `traingdm`, que incluye un descenso con velocidad adquirida. el `momentumn` permite que la red no sólo responda a el gradiente local sino también a las tendencias del error. Actuando como un filtro del bajo-paso, el `momentumn` permite a la red ignorar rasgos pequeños en el error. Sin el `momentumn` una red puede estancarse en un mínimo local poco profundo.

El `momentumn` puede agregarse al aprendizaje del `backpropagation` haciendo cambios de peso iguales a la suma de un fragmento del último cambio de peso y el nuevo cambio sugerido por la regla del `backpropagation`. La magnitud del efecto que el último cambio de peso se permite tener es mediada por una constante de velocidad adquirida,

mc que puede ser cualquier número entre 0 y 1. Cuando la constante de momentum es 0, el cambio de peso es solamente basado en el gradiente. Cuando la constante de momentum es 1, el nuevo cambio de peso se pone para igualar el último cambio de peso y el gradiente simplemente se ignora. El gradiente es computada sumando los gradientes calculadas a cada dato de entrenamiento y sólo se actualizan los pesos y el umbral después de todos los datos del entrenamiento que se ha presentado.

Si la nueva función de desempeño en una iteración dada excede la función de desempeño en una iteración anterior por más de un `max_perf_inc` de proporción de predefinida (típicamente 1.04), se desechan los nuevos pesos, el umbral, y el mc coeficiente de momentum se pone en cero.

La forma de entrenamiento de descenso de gradiente con momentum se invoca con la función `traingdm`. Esta función usa los mismos pasos mostrados que el `traingd` sólo que el mc, lr y `max_perf_inc` aprenden todos parámetros.

En el código siguiente se utilizara la red creada anteriormente pero será entrenada con la función `traingd`. Los parámetros serán iguales con la inclusión del mc factor de momentum y el máximo performance aumento `max_perf_inc`. (Los parámetros de entrenamiento se restablecen a los valores predefinidos siempre que `net.trainFcn` se ponga al `traingdm`).

- `net=newff(minmax(p),[3,1],{'tansig','purelin'},'traingdm');`
- `net.trainParam.show = 50;`
- `net.trainParam.lr = 0.05;`
- `net.trainParam.mc = 0.9;`
- `net.trainParam.epochs = 300;`
- `net.trainParam.goal = 1e-5;`
- `[net,tr]=train(net,p,t);`
- `a = sim(net,p)`
- `a = -1.0026 -1.0044 0.9969 0.9992`

4.8.7 Entrenamiento rápido

Los dos algoritmos de entrenamiento backpropagation: descenso de gradiente, y descenso de gradiente con velocidad adquirida. Son métodos lentos y para resolver problemas prácticos. Existen algoritmos que pueden converger de diez a cien veces más rápidamente que estos.

Estos algoritmos más rápidos entran en dos categorías principales. La primera categoría usa técnicas heurísticas que se desarrollaron de un análisis para la optimización del algoritmo de descenso de gradiente normal. Una de estas modificaciones heurísticas es la técnica de velocidad adquirida. Dos de estas técnicas heurísticas son: el de variable de tasa de aprendizaje `traingda`; y el de rezago `trainrp`.

a. Taza de aprendizaje (`traingda`, `traingdx`)

Con el algoritmo de descenso normal, la proporción de aprendizaje se mantiene constante a lo largo del entrenamiento. La activación del algoritmo es muy sensible a la tasa de aprendizaje. Si a esta se fija demasiado alta, el algoritmo puede oscilar y volverse inestable. Si la tasa de aprendizaje es demasiado pequeña, el algoritmo tardara en converger. No es práctico determinar el valor óptimo para la tasa de aprendizaje

antes de entrenar. El rendimiento del algoritmo de descenso puede mejorarse si se permite que la tasa de aprendizaje pueda cambiar durante el proceso de entrenamiento.

Además requiere algunos cambios en el procedimiento de entrenamiento usado por `traingd`. Primero, se calculan el rendimiento de la red inicial y el error. A cada epochs se le calculan nuevos pesos, el umbral usando y la tasa aprendizaje actual. Se calculan las nuevas salidas y los errores.

Como con velocidad adquirida, si el nuevo error excede el error viejo por más de un `max_perf_inc` predefinido (típicamente 1.04), se desechan los nuevos pesos y el umbral. Además, la proporción de aprendizaje se disminuye (típicamente multiplicando por `lr_dec = 0.7`). Por otra parte, se guardan los nuevos pesos, etc. Si el nuevo error es menor del error anterior, la tasa de aprendizaje se aumenta (típicamente multiplicando por `lr_inc = 1.05`).

Este procedimiento aumenta la tasa de aprendizaje, pero sólo a la magnitud que la red puede aprender sin aumentos grandes de error. Así, se obtiene una tasa de aprendizaje aproximada a la óptima. Cuando la tasa de aprendizaje es muy alta el aprendizaje es estable, la tasa de aprendizaje se aumenta. Para garantizar una disminución en el error.

El entrenamiento con tasa de aprendizaje adaptable se ejecuta con la función `traingda`, se invoca como el `traingd`, salvo el `max_perf_inc`, `lr_dec`, y `lr_inc` se deben incluir en los parámetros así.

- `net=newff(minmax(p),[3,1],{'tansig','purelin'},'traingda');`
- `net.trainParam.show = 50;`

- `net.trainParam.lr = 0.05;`
- `net.trainParam.lr_inc = 1.05;`
- `net.trainParam.epochs = 300;`
- `net.trainParam.goal = 1e-5;`
- `[net,tr]=train(net,p,t);`
- `a = sim(net,p)`

La función `traingdx` combina tasa de aprendizaje adaptable con el entrenamiento de velocidad adquirida. Se invoca de la misma manera como `traingda`, sólo que tiene el `mc` de coeficiente de momentum como un parámetro de entrenamiento adicional.

b. Rezago (trainrp)

En las redes multicapa se usan funciones de transferencia sigmoideas en la capa oculta. Estas funciones comprimen un rango de la entrada infinito en uno finito. Las funciones sigmoideas son caracterizadas por que tienden a acercarse a ceros cuando la entrada es más grande. Esto es un problema cuando se usan algoritmos de descenso para entrenar una red multicapa con funciones sigmoideas, porque el gradiente puede tener una magnitud muy pequeña y por consiguiente, causar cambios pequeños en los pesos y el umbral, aunque los pesos y el umbral están lejos de sus valores óptimos.

El propósito del algoritmo de entrenamiento de rezago (Rprop) es eliminar los efectos de las magnitudes de los derivados parciales. La señal del derivado se usa para determinar la dirección de la actualización del peso; la magnitud del derivado no tiene efecto en la actualización del peso. El tamaño del cambio de peso es determinado por un valor de actualización aparte. El valor de actualización para cada peso y el umbral es

aumentado por un factor `delt_inc` siempre que el derivado de la función de desempeño con respecto a ese peso tenga la misma señal para dos iteraciones sucesivas. El valor de actualización es disminuido por un factor `delt_dec` siempre que el derivado con respecto que el peso cambie la señal de la iteración anterior. Si el derivado es cero, entonces los restos de valor de actualización son los mismos. Siempre que los pesos estén oscilando que el cambio de peso se reducirá. Si el peso continúa cambiando en la misma dirección para varias iteraciones, entonces la magnitud del cambio de peso se aumentará.

Los parámetros de entrenamiento para el `trainrp` son `epochs`, `show`, `goal`, `time`, `min_grad`, `max_fail`, `delt_inc`, `delt_dec`, `delta0`, `deltamax`. Los últimos dos son el tamaño del paso inicial y el máximo tamaño, respectivamente. El algoritmo `Rprop` generalmente converge más rápido que los algoritmos anteriores.

- `net=newff(minmax(p),[3,1],{'tansig','purelin'},'trainrp');`

`Rprop` generalmente es más rápido que el algoritmo de descenso normal. También requiere sólo un espacio modesto en memoria.

4.8.8 Algoritmos de gradiente conjugado

El algoritmo del `backpropagation` básico ajusta los pesos en la dirección de descenso (negativo del gradiente). Ésta es la dirección en la que la función de desempeño disminuye rápidamente. Aunque la función

disminuya rápidamente a lo largo del negativo del gradiente, esto necesariamente no produce una convergencia más rápida. En los algoritmos de gradiente conjugado una búsqueda se realiza a lo largo de direcciones conjugadas que producen generalmente una convergencia más rápida que las direcciones de descenso.

En la mayoría de los algoritmos de gradiente conjugado, el tamaño del paso se ajusta en cada iteración. Una búsqueda se hace a lo largo de la dirección de gradiente conjugado para determinar el tamaño del paso que minimiza la función de desempeño. Hay cinco funciones de la búsqueda diferentes incluidas en el toolbox y cuatro variaciones diferentes de algoritmos de este tipo.

a. Actualización de fletcher-reeves (traincgf)

Todos los algoritmos de gradiente conjugado empiezan investigando en la dirección de descenso (negativo de gradiente) en la primera iteración.

$$P_0 = -g_0$$

Se realiza una búsqueda de la línea para determinar la distancia óptima para determinar la dirección de la búsqueda actual entonces:

$$X_{k+1} = X_k + \alpha_k P_k$$

Entonces la próxima dirección de búsqueda es determinada para que sea conjugada a las direcciones de la búsqueda anteriores. El

procedimiento determina la nueva dirección de la búsqueda al combinar la nueva dirección de descenso con la dirección de la búsqueda anterior:

$$\mathbf{p}_k = -\mathbf{g}_k + \beta_k \mathbf{p}_{k-1}$$

Las diferentes versiones de gradiente conjugado son distinguidas por la manera en la que la constante β_k se computa. Para el Fletcher-Reeves el procedimiento es:

$$\beta_k = \frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}}$$

$$\beta_k = \frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}}$$

En el siguiente código se incluyen los parámetros de entrenamiento para el `traincgf` son `epochs`, `epochs`, `show`, `goal`, `time`, `min_grad`, `max_fail`, `srchFcn`, `scal_tol`, `alfa`, `beta`, `delta`, `gama`, `low_lim`, `up_lim`, `maxstep`, `minstep`, `bmax`.. El parámetro `srchFcn` es el nombre de la función de búsqueda lineal. Puede ser cualquiera de las funciones que existen (o una función usuario). Los parámetros restantes son asociados con rutinas de búsqueda lineal específicas. La búsqueda lineal predefinida el `srchcha` se usa en este ejemplo. El `traincgf`

generalmente converge en menos iteraciones que el `trainrp` (aunque hay más cómputo requerido en cada iteración).

- `net=newff(minmax(p),[3,1],{'tansig','purelin'},'traincgf');`
- `net.trainParam.show = 5;`
- `net.trainParam.epochs = 300;`
- `net.trainParam.goal = 1e-5;`
- `[net,tr]=train(net,p,t);`
- `a = sim(net,p)`

b.Actualización de polak-ribière (`traincgp`)

Otra versión del algoritmo conjugado fue propuesta por Polak y Ribière. Como algoritmo de Fletcher-Reeves, la dirección de la búsqueda a cada iteración es determinada por:

$$p_k = -g_k + \beta_k p_{k-1}$$

Para Polak-Ribière la constante β_k se calcula con:

$$\beta_k = \frac{\Delta g_{k-1}^T g_k}{g_{k-1}^T g_{k-1}}$$

$$\beta_k = \frac{\Delta g_{k-1}^T g_k}{g_{k-1}^T g_{k-1}}$$

En el siguiente código, se tomara el mismo ejemplo anterior y se entrenara con versión del algoritmo de Polak-Ribière. Los parámetros de entrenamiento para el `traincgp` son iguales que para el `traincgf`. La búsqueda lineal predefinida el `srchcha` se usa en este ejemplo. Los parámetros `show` y `epochs` tienen a los mismos valores.

- `net=newff(minmax(p),[3,1],{'tansig','purelin'},'traincgp');`
- `net.trainParam.show = 5;`

- net.trainParam.epochs = 300;
- net.trainParam.goal = 1e-5;
- [net,tr]=train(net,p,t);
- a = sim(net,p)

c. Restablecimiento de powell-beale (traincgb)

Para todos los algoritmos de gradiente conjugado, la dirección de la búsqueda se restablecerá periódicamente al negativo del gradiente. La norma de restablecer ocurre cuando el número de iteraciones es igual al número de parámetros de la red (pesos y el umbral), pero hay otros métodos de restablecimiento que pueden mejorar la eficacia de entrenamiento. Esto se prueba con la siguiente desigualdad:

$$\left| \mathbf{g}_{k-1}^T \mathbf{g}_k \right| \geq 0.2 \|\mathbf{g}_k\|^2$$

Si esta condición está satisfecha, la dirección de la búsqueda se restablece al negativo de la pendiente.

En el siguiente código, se crea nuevamente la red anterior y se entrena con la versión del algoritmo de Powell-Beal. Los parámetros de entrenamiento para el traincgb son los mismos que para el traincgf. La búsqueda lineal predefinida el srchcha se usa en este ejemplo. Los parámetros show y epochs tienen los mismos valores.

- net=newff(minmax(p),[3,1],{'tansig','purelin'},'traincgb');
- net.trainParam.show = 5;
- net.trainParam.epochs = 300;
- net.trainParam.goal = 1e-5;
- [net,tr]=train(net,p,t);.
- a = sim(net,p)

d. Gradiente conjugado escalado (trainscg)

Hasta ahora cada uno de los algoritmos de gradiente conjugado requiere una búsqueda lineal en cada iteración. Esta búsqueda lineal requiere un alto nivel computacional, ya que requiere que la respuesta de la red a todas las entradas de entrenamiento sea computada en varios tiempos por cada búsqueda. El algoritmo de gradiente conjugado escalado (SCG) fue diseñado para evitar la búsqueda lineal.

Este algoritmo es demasiado complejo para explicar en unas líneas, pero la idea básica es combinar el la aproximación de región de modelo-confianza con el gradiente conjugado.

En el siguiente código, se crea nuevamente la red anterior pero ahora utilizando el algoritmo de entrenamiento de gradiente escalado. Los parámetros de entrenamiento para el trainscg son epochs, show, goal, time, min_grad, max_fail, sigma, lambda. El parámetro sigma determina el cambio en el peso para la segunda aproximación derivativa. El parámetro lambda regula la indecisión. Los parámetros show y epochs son 10 y 300, respectivamente.

```
▪ net=newff(minmax(p),[3,1],{'tansig','purelin'},'trainscg');
▪ net.trainParam.show = 10;
▪ net.trainParam.epochs = 300;
▪ net.trainParam.goal = 1e-5;
▪ [net,tr]=train(net,p,t);
▪ a = sim(net,p)
```

4.8.9 Rutinas de búsqueda lineal

Algunos de los algoritmos de entrenamiento de gradiente conjugado y cuasi-Newton requieren una búsqueda lineal. Toolbox ofrece cinco

búsquedas lineales diferentes que se puede usar con el parámetro `srchFcn` de la función `entrenamiento` y dando a este el valor de la función de búsqueda deseada.

a. Búsqueda de sección dorada (`srchgol`)

La búsqueda de sección dorada `srchgol`, es una búsqueda lineal que no requiere el cálculo de la pendiente. Esta rutina empieza localizando un intervalo en el que el mínimo de la activación ocurre. Esto se hace evaluando la activación a una sucesión de puntos, empezando a una distancia de `delta` y doblando en esta distancia cada paso, a lo largo de la dirección de la búsqueda. Cuando la activación aumenta entre dos iteraciones sucesivas, un mínimo se une. El próximo paso es reducir el tamaño del intervalo que contiene el mínimo. Se localizan dos nuevos puntos dentro del intervalo inicial. Los valores de la activación a estos dos puntos determinan una sección del intervalo que puede desecharse y un nuevo punto interior se pone dentro del nuevo intervalo. Este procedimiento es seguido hasta que el intervalo de incertidumbre se reduzca al tamaño que es igual al `delta/scale_tol`.

b. Búsqueda de brent (`srchbre`)

La búsqueda de Brent es una búsqueda lineal y una combinación híbrida de la búsqueda de la sección dorada y una interpolación cuadrática. Los métodos de comparación de función, como la

búsqueda de la sección dorada, tienen una proporción del primer orden de convergencia, mientras los métodos como el de interpolación polinómica tienen una proporción que es más rápida que la súper lineal. Por otro lado, la proporción de convergencia para las salidas de búsqueda de sección doradas cuando el algoritmo se inicializa es considerando que la conducta proporcional para los métodos de la interpolación polinómica puede tomar muchas iteraciones para solucionarse. La búsqueda de Brent intenta combinar los mejores atributos de ambos acercamientos.

Para la búsqueda de Brent se empieza con el mismo intervalo de incertidumbre que se usó con la búsqueda de la sección dorada, pero algunos puntos adicionales se computan. Una función cuadrática se ajusta entonces a estos puntos y el mínimo de la función cuadrática se computa. Si este mínimo está dentro del intervalo apropiado de incertidumbre, se usa en la próxima fase de la búsqueda y una nueva aproximación cuadrática se realiza. Si las caídas mínimas están fuera del intervalo conocido de incertidumbre, entonces un paso de la búsqueda de la sección dorada se realizara.

c. Búsqueda bisección-cúbica híbrida (srchhyb)

La búsqueda de Brent, srchhyb son un algoritmo híbrido. Es una combinación de bisección y la interpolación cúbica. Para el algoritmo de la bisección, un punto se localiza en el intervalo de incertidumbre, la activación y su derivado se computan. Basados en esta información, la

mitad del intervalo de incertidumbre se desecha. En el algoritmo híbrido, una interpolación cúbica de la función es obtenida usando el valor de la activación y su derivado a los dos puntos del extremo. Si el mínimo de la interpolación cúbica cae dentro del intervalo conocido de incertidumbre, entonces se usa para reducir el intervalo de incertidumbre. De lo contrario un paso del algoritmo de bisección se usa.

d. Búsqueda de carambola (srchcha)

El método de carambola srchcha fue diseñado para ser usado en combinación con un algoritmo de gradiente conjugado para el entrenamiento de red neuronal. Como los dos métodos anteriores, es una búsqueda híbrida. Usa una interpolación cúbica, junto con un tipo de seccionamiento.

e. Backtracking (srchbac)

La búsqueda backtracking srchbac se satisface mejor para usar con los algoritmos de optimización cuasi-Newton. Empieza con un multiplicador del paso de 1 y entonces se reduce hasta un valor aceptable en la activación. En el primer paso usa el valor de activación al punto actual y a un multiplicador del paso de 1. También usa el valor del derivado de activación al punto actual, obtener una aproximación cuadrática a la función de la activación a lo largo de la dirección de la búsqueda. El mínimo de la aproximación cuadrática se vuelve un punto óptimo provisional (bajo ciertas condiciones) y la activación se prueba. Si la

activación no se reduce suficientemente, una interpolación cúbica se obtiene y el mínimo de la interpolación cúbica se vuelve el nuevo punto óptimo provisional. Este proceso es continuado hasta lograr reducción suficiente en la activación.

4.8.10 Algoritmos Cuasi-Newton

a. Algoritmo de BFGS (trainbgf)

El método de Newton es una alternativa a los métodos de gradiente conjugado para la optimización rápida. El paso básico del método de Newton es:

$$X_{k+1} = X_k - A_k^{-1} g_k$$

Donde A_k es la matriz de Hessianos de la segunda derivada del índice de la activación, los valores actuales de los pesos y el umbral. El método de Newton converge a menudo más rápidamente que los métodos de gradiente conjugado. Desgraciadamente, es complejo y requiere alta computabilidad. Hay una clase de algoritmos que son basados en el método de Newton, pero que no requiere el cálculo de la segunda derivada. Éstos se llaman Cuasi-Newton (o secantes). Ellos actualizan una matriz de Hessianos aproximada a cada iteración del algoritmo. La actualización se computa como una función de gradiente. El método Cuasi-Newton más exitoso en estudios publicados es el Broyden, Fletcher, Goldfarb, y Shanno (BFGS) la actualización. Este algoritmo se logra con la función trainbgf.

Ahora en el ejemplo se reiniciara la red pero ahora entrenándola con el algoritmo BFGS Cuasi-Newton. Los parámetros de entrenamiento para el `trainbfg` serán igual que aquéllos para el `traincgf`. La búsqueda lineal predefinida es el `srchbac`. Los parámetros `show` y `epochs` corresponden a 5 y 300, respectivamente.

- `et=newff(minmax(p),[3,1],{'tansig','purelin'},'trainbfg');`
- `net.trainParam.show = 5;`
- `net.trainParam.epochs = 300;`
- `net.trainParam.goal = 1e-5;`
- `[net,tr]=train(net,p,t);`
- `a = sim(net,p)`

b. Algoritmo secante de un paso (`trainoss`)

Desde que el algoritmo BFGS requiere más almacenamiento y cómputo en cada iteración que los algoritmos de gradiente conjugado, surge la necesidad de una aproximación secante con almacenamiento más pequeño y menos requisitos de cómputo. El método secante de un paso (OSS) es un esfuerzo cubrir el espacio entre los algoritmos de gradiente conjugado y los algoritmos Cuasi-Newton (secante). Este algoritmo no guarda la matriz de Hessianos completa; asume en cada iteración que el Hessiano anterior era la matriz de identidad. Esto tiene la ventaja adicional que la nueva dirección de la búsqueda puede calcularse sin computar el inverso de la matriz.

Nuevamente es creado nuestro habitual ejemplo reemplazando la red y entrenándola mediante el algoritmo de secante de un paso. Los parámetros de entrenamiento para el `trainoss` son iguales que para el

traincgf. La búsqueda de la línea predefinida es srchbac. Los parámetros show y epochs son ajustados a 5 y 300, respectivamente.

- net=newff(minmax(p),[3,1],{'tansig','purelin'},'trainoss');
- net.trainParam.show = 5;
- net.trainParam.epochs = 300;
- net.trainParam.goal = 1e-5;
- [net,tr]=train(net,p,t);
- a = sim(net,p)

c. Levenberg-marquardt (trainlm)

Como los métodos Cuasi-Newton, el algoritmo de Levenberg-Marquardt fue diseñado para acercarse en segundo orden que entrena a gran velocidad sin tener que computar la matriz de Hesianos. Cuando la función de la activación tiene la forma de una suma de cuadrados (como es típico entrenando feedforward), entonces la matriz de Hesianos puede aproximarse como:

$$\mathbf{H}=\mathbf{J}^T\mathbf{J}$$

Y el gradiente puede computarse como:

$$\mathbf{g}=\mathbf{J}^T\mathbf{e}$$

Donde \mathbf{J} es la matriz de Jacobianos que contiene la derivada de los errores de la red primero con respecto a los pesos y el umbral \mathbf{e} es el vector de errores de la red. La matriz de Jacobianos puede computarse a través de backpropagation normal que es mucho menos complejo de computar que la matriz de Hesianos.

El algoritmo de Levenberg-Marquardt hace esta aproximación de la matriz de Hessianos como la siguiente actualización Newton:

$$\mathbf{X}_{k+1} = \mathbf{X}_k - [\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}]^{-1} \mathbf{J}^T \mathbf{e}$$

Cuando el escalar μ es cero, se comporta como el método de Newton usando la matriz de Hessianos aproximada. Cuando μ es alto, se comporta de descenso de gradiente pero con un paso pequeño. El método de Newton es más rápido y más exacto cerca de un mínimo del error, así que el objetivo es cambiar hacia el método de Newton tan rápidamente como sea posible. Así, μ se disminuye después de cada paso exitoso (reducción en función de la activación) y sólo se aumenta en un paso provisional. De esta manera, la función de desempeño se reducirá siempre a cada iteración del algoritmo.

Los parámetros de entrenamiento para `trainlm` son `epochs`, `show`, `goal`, `time`, `min_grad`, `max_fail`, `mu`, `mu_dec`, `mu_inc`, `mu_max`, `mem_reduc`. El parámetro `mu` es el valor inicial de μ . Este valor es multiplicado por `mu_dec` siempre que la función de desempeño se reduzca un paso. Es multiplicado por `mu_inc` siempre que aumentara la función de desempeño en un paso. Si `mu` se vuelve mayor que el `mu_max`, el algoritmo se detiene. El parámetro `mem_reduc` se usa para controlar la cantidad de memoria usada por el algoritmo. Los parámetros `show` y `epochs` se inicializan en 5 y 300, respectivamente.

```
· net=newff(minmax(p),[3,1],{'tansig','purelin'},'trainlm');  
· net.trainParam.show = 5;  
· net.trainParam.epochs = 300;
```

- net.trainParam.goal = 1e-5;
- [net,tr]=train(net,p,t);
- a = sim(net,p)

d. Levenberg-marquardt de memoria reducida (trainlm)

El inconveniente principal del algoritmo de Levenberg-Marquardt es que requiere hacer el almacenamiento de las matrices que con toda seguridad es bastante grande. El tamaño de la matriz de Jacobianos es, $Q \times n$, donde Q es que el número de juegos de entrenamiento y n es el número de pesos y el umbral en la red. Resulta que esta matriz no tiene que ser computada y guarda en conjunto. Por ejemplo, si se quiere dividir el Jacobiano en dos submatrices iguales que se pueden computar como la matriz aproximada de Hesianos así:

$$H = J^T J = \begin{bmatrix} J_1^T & J_2^T \end{bmatrix} \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} = J_1^T J_1 + J_2^T J_2$$

El Hesiano aproximado puede ser computado sumando una serie de subterminos. Una vez un subtermino se ha computado pueden aclararse las submatrices correspondientes del Jacobiano.

Cuando se usa la función de entrenamiento trainlm, el parámetro mem_reduc se usa para determinar cuántas filas del Jacobiano serán computadas en cada submatriz. Si el mem_reduc toma el valor 1, entonces el Jacobiano lleno se computa y ninguna reducción de memoria se logra. Si el mem_reduc se pone a 2, entonces sólo la mitad del Jacobiano se computará. Esto ahorra es la mitad de la memoria usada por el cálculo del Jacobiano lleno.

4.8.11 Regularización

El primer método para mejorar la generalización se llama regularización. Donde se va modificando la función de desempeño que normalmente se es la suma de cuadrados de los errores de la red en entrenamiento.

a. Función de desempeño modificada

La función de desempeño típica que se usa para entrenar las redes feedforward es la suma de cuadrados de los errores de la red.

$$F = \text{mse} = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - a_i)^2$$

Es posible mejorar generalización si se modifica la función de desempeño agregando un término que consiste en la suma de cuadrados de los pesos de la red y el umbral.

$$\text{msereg} = \gamma \text{mse} + (1 - \gamma) \text{msw}$$

Donde γ es la proporción de activación, y

$$\text{msw} = \frac{1}{n} \sum_{j=1}^n w_j^2$$

Usando esta función de desempeño causarán la red tenga pesos y umbrales más pequeños y esto obligará a que las salidas la red sea más plana.

En el siguiente ejemplo se reinicializara la red y se volverá a entrenar usando el algoritmo de BFGS con la función de desempeño regularizada. La proporción de actuación será 0.5 y el peso igual a los errores cuadrados de los pesos, pesos cuadrados malos.

- `net=newff(minmax(p),[3,1],{'tansig','purelin'},'trainbfg');`
- `net.performFcn = 'msereg';`
- `net.trainParam.ratio = 0.5;`
- `net.trainParam.show = 5;`
- `net.trainParam.epochs = 300;`
- `net.trainParam.goal = 1e-5;`
- `[net,tr]=train(net,p,t);`

El problema con regularización es que es difícil de determinar el valor óptimo por el parámetro de proporción de activación. Si se hace este parámetro demasiado grande, nosotros podemos conseguir más aproximado. Si la proporción es demasiado pequeña, la red no encajará los datos de entrenamiento de forma adecuada.

b. Regularización automatizada (trainbr)

Es deseable determinar los parámetros de regularización óptimos en una moda automatizada. Un acercamiento a este proceso Bayesiano de David MacKay. En este armazón, los pesos y el umbral de la red se asumen como variables al azar con distribuciones especifica. Los parámetros de regularización se relacionan a las variaciones desconocidas y asociadas con estas distribuciones.

La regularización Bayesiana se hace con la función trainbr. El siguiente código muestra cómo se puede entrenar una 1-20-1 red que usa esta función para aproximar.

- `p = [-1:.05:1];`
- `t = sin(2*pi*p)+0.1*randn(size(p));`
- `net=newff(minmax(p),[20,1],{'tansig','purelin'},'trainbr');`
- `net.trainParam.show = 10;`
- `net.trainParam.epochs = 50;`
- `randn('seed',192736547);`
- `net = init(net);`
- `[net,tr]=train(net,p,t);`

Un rasgo de este algoritmo proporciona una medida de cuántos parámetros de la red (pesos y umbral) es usando eficazmente por esta.

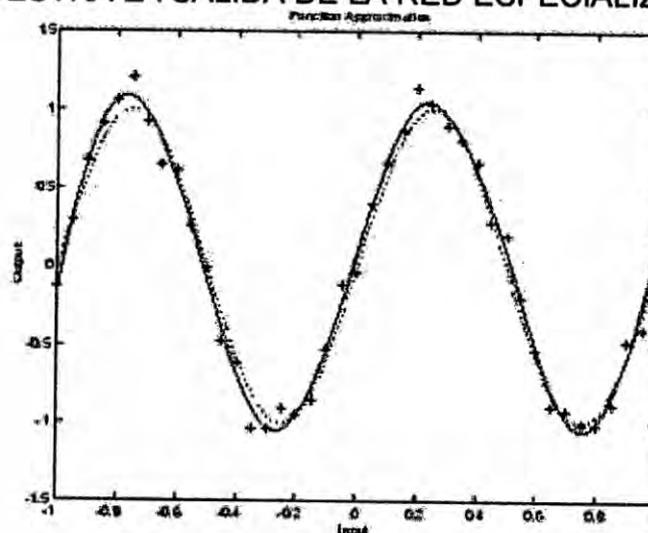
En este caso, la red especializada final usa aproximadamente 12 parámetros (indicado por #Par) fuera del total de 61 pesos y umbral en la red 1-20-1. Este número eficaz de parámetros debe permanecer aproximadamente igual, no importa que tan grande sea el número total de parámetros en la red. (Esto asume que la red ha estado especializada para un número suficiente de iteraciones y asegura su convergencia.)

El algoritmo del `trainbr` generalmente trabaja mejor cuando se descascaran las entradas de la red y objetivos para que ellos se escalen aproximadamente $[-1,1]$. Si sus entradas y objetivos no se escalan en este rango, se puede usar la función `premnmx`, o `prestd`.

La grafico N°4.7 muestra la salida de la red especializada. Se ve que la respuesta de la red está muy cerca de la función seno subyacente (línea punteada), y por consiguiente, la red generalizará bien a las nuevas entradas.

Al usar trainbr, es importante dejar correr el algoritmo hasta el número justo de parámetros que han convergido. Los entrenamientos se detienen con el mensaje "Maximum MU reached.". Esto indica que el algoritmo ha convergido de verdad. También puede decirse que el algoritmo ha convergido si la suma cuadrada del error (SSE) y la suma cuadrada de los pesos (SSW) son relativamente constantes en varias iteraciones.

GRAFICO N° 4.1
MUESTRA LA SALIDA DE LA RED ESPECIALIZADA



4.8.12 Detención temprana

Otro método para mejorar la generalización es el de detención temprana. En esta técnica los datos disponibles son divididos tres subconjuntos. El primer subconjunto es el entrenamiento puesto que se usa para computar el gradiente y actualizar los pesos de la red y el umbral. El segundo subconjunto es el juego de aprobación. El error en el juego de aprobación

se supervisa durante el proceso de entrenamiento. El error de aprobación normalmente disminuirá durante la fase inicial de entrenamiento. Sin embargo, cuando la red empieza validar los datos, el error en el juego de aprobación empezará a subir típicamente. Cuando el error de aprobación aumenta para un número especificado de iteraciones, el entrenamiento se detiene y se vuelven los pesos y el umbral al mínimo del error de aprobación.

Primero se crea un problema de una prueba simple. Para este entrenamiento generando una ola de la función seno que oscila entre -1 a 1 con pasos de 0.05.

- `p = [-1:0.05:1];`
- `t = sin(2*pi*p)+0.1*randn(size(p));`

Luego se genera el juego de aprobación. Las entradas van de -1 a 1, como en la prueba puesta. Para hacer el problema más realista, se agrega también una sucesión del ruido diferente a la función seno subyacente.

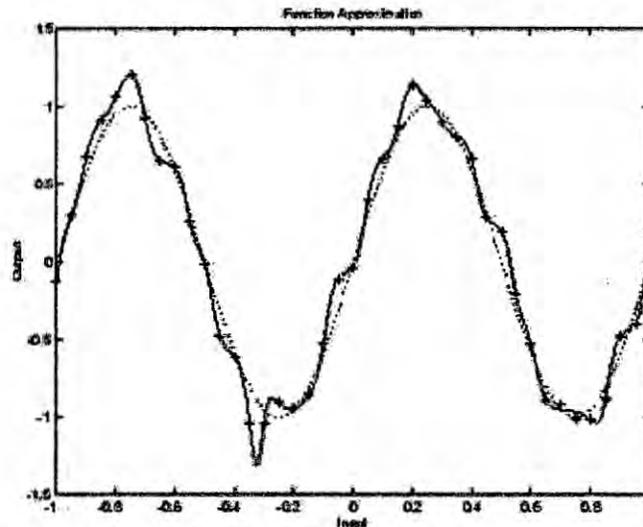
- `val.P = [-0.975:.05:0.975];`
- `val.T = sin(2*pi*v.P)+0.1*randn(size(v.P));`

Para este ejemplo se usa la función de entrenamiento `traingdx`.

- `net=newff([-1 1],[20,1],{'tansig','purelin'},'traingdx');`
- `net.trainParam.show = 25;`
- `net.trainParam.epochs = 300;`
- `net = init(net);`
- `[net,tr]=train(net,p,t,[],[],val); 6`
- `TRAININGDX, Validation stop.`

La figura siguiente muestra un gráfico de la respuesta de la red.

GRAFICO N° 4.2
RESPUESTA DE LA RED



4.8.13 Preproceso y postproceso

Puede hacerse una red neuronal más eficaz si se realizan ciertos pasos en el preprocesando las entradas y los objetivos de la red.

Min y Max (premnmx, postmnmx, trammx)

Antes de entrenar, es útil escalar las entradas y objetivos para que siempre caigan dentro de un rango específico. La función premnmx puede usarse para escalar las entradas y objetivos para que queden en el rango [-1,1]. El código siguiente ilustra el uso de esta función.

- `[pn,minp,maxp,tn,mint,maxt] = premnmx(p,t);`
- `net=train(net,pn,tn);`

Se dan las entradas y objetivos originales de la red en matrices p y t. pn y tn que devuelven todos a caer en el intervalo [-1,1]. El minp de los vectores y maxp que contienen el mínimo y el máximo valor de las entradas originales y los vectores, los maxt que contienen el mínimo y el

máximo valor de los objetivos originales. Después que la red ha estado especializada, estos vectores deben usarse para transformar cualquier entrada futura que se aplica a la red. Ellos se vuelven una parte de la red. Si el `premnmx` se usa para escalar las entradas y objetivos, entonces la salida de la red se entrenará para producir respuestas en el rango $[-1, 1]$. Si se quiere convertir estos resultados en las mismas unidades que se usaron para los objetivos originales, entonces se debe usar el `postmnmx`. En el código siguiente, simula la red, y convierte las salidas de la red a las unidades originales.

- `an = sim(net,pn);`
- `a = postmnmx(an,mint,maxt);`

El rendimiento de la red corresponde al `tn` de los objetivos normalizados. El rendimiento de la red in-normalizado está en las mismas unidades como los objetivos originales `t`.

Siempre que la red especializada se use con nuevas entradas estas deben ser preprocesadas con el mínimo y máximo que se computaron para el juego de entrenamiento. Esto puede lograrse con el `tramnmx`. En el código siguiente, se aplica un nuevo juego de entradas a la red que ya se ha entrenado.

- `pnewn = tramnmx(pnew,minp,maxp);`
- `anewn = sim(net,pnewn);`
- `anew = postmnmx(anewn,mint,maxt);`

Normalización Estándar Dev.(`prestd`, `poststd`, `trastd`)

Otro procedimiento para las entradas y los objetivos de la red es la normalización, la desviación estándar y normal del juego de entrenamiento. Este procedimiento se lleva a cabo con la función `prestd`. Normaliza las entradas y objetivos. El código siguiente ilustra el uso de `prestd`.

- `[pn,meanp,stdp,tn,meant,stdt] = prestd(p,t);`

Se dan las entradas y objetivos de la red originales en matrices `p` y `t`. Las entradas normalizadas y objetivos, `pn` y `tn` que tendrán ceros intermedios y la unidad de desviación normal. El `meanp` de los vectores y `stdp` que contienen las desviaciones estándar y normales de las entradas originales, y los vectores `meant` y `stdt` contienen los medios y las desviaciones normales de los objetivos originales. Después que la red ha estado especializada, estos vectores deben usarse para transformar cualquier entrada futura que se aplica a la red.

En el siguiente código se simula la red especializada en el código anterior y convierte las salidas de la red atrás en las unidades originales.

- `an = sim(net,pn);`
- `a = poststd(an,meant,stdt);`

Las salidas de la red corresponden al `tn` de los objetivos normalizados. La salida de la red in-normalizada está en las mismas unidades como los objetivos originales `t`.

Esto puede lograrse con el `trastd`. En el siguiente código, se aplica un nuevo juego de entradas a la red ya hemos entrenado.

- `pnewn = trastd(pnew,meanp,stdp);`
- `anewn = sim(net,pnewn);`
- `anew = poststd(anewn,meant,stdt);`

4.9 Comparación de resultados estadísticos y Matlab

En la comparación de los resultados obtenidos por el método estadístico y Matlab se observó claramente que los resultados por Matlab son los resultados de predicción con más acercamiento a lo real; para este caso se trabajó con el parámetro principal de la potencia eólica, la velocidad del viento. Se trabajó con los datos del mes de abril prediciendo para el día 30 de abril, teniendo como antecedente los datos del mes de abril del 1ro al 29vo día de abril.

GRAFICO N° 4.3
COMPARACION DE LA VELOCIDAD DE PREDICION ESTADISTICA Y VELOCIDAD REAL

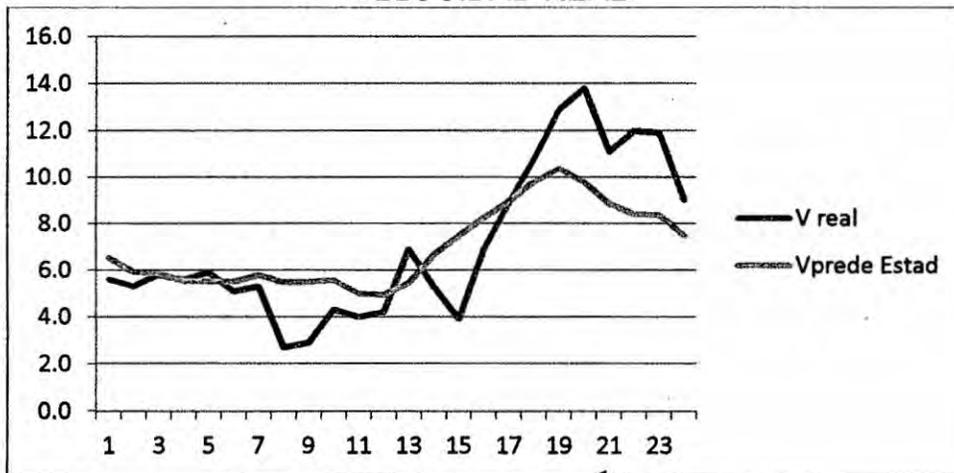
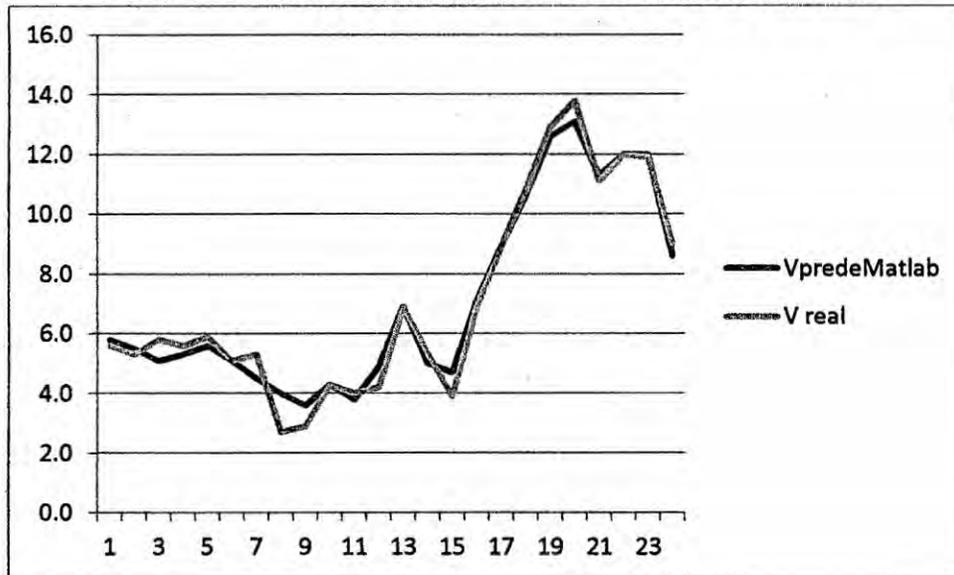


GRAFICO N° 4.4
COMPARACION DE LA VELOCIDAD DE PREDICCIÓN DEL VIENTO
POR REDES NEURONALES Y LA VELOCIDAD REAL



4.10 Elección del modelo de predicción de potencia

Como se ha planteado al inicio de este trabajo, el objetivo final consiste en la elección de un modelo que permita predecir los valores de la potencia antelación de 24horas (un día).

La técnica de los ingresos de datos adecuadamente es la clave para un buen resultado no dejando de lado también que el inicio es lo fundamental me refiero a la selección de datos, un solo dato de un determinado día de una determinada hora errado, conllevaría a una predicción errada.

Se aplicó redes neuronales para la predicción del viento con una antelación de 24 horas, obteniéndose los datos, luego se calculó con fórmulas establecidas la potencia de salida y también se realizó el cálculo de la potencia de salida mediante RetsCreen donde se obtuvo los resultados satisfactoriamente, se detalla en el capítulo siguiente .

CAPITULO V

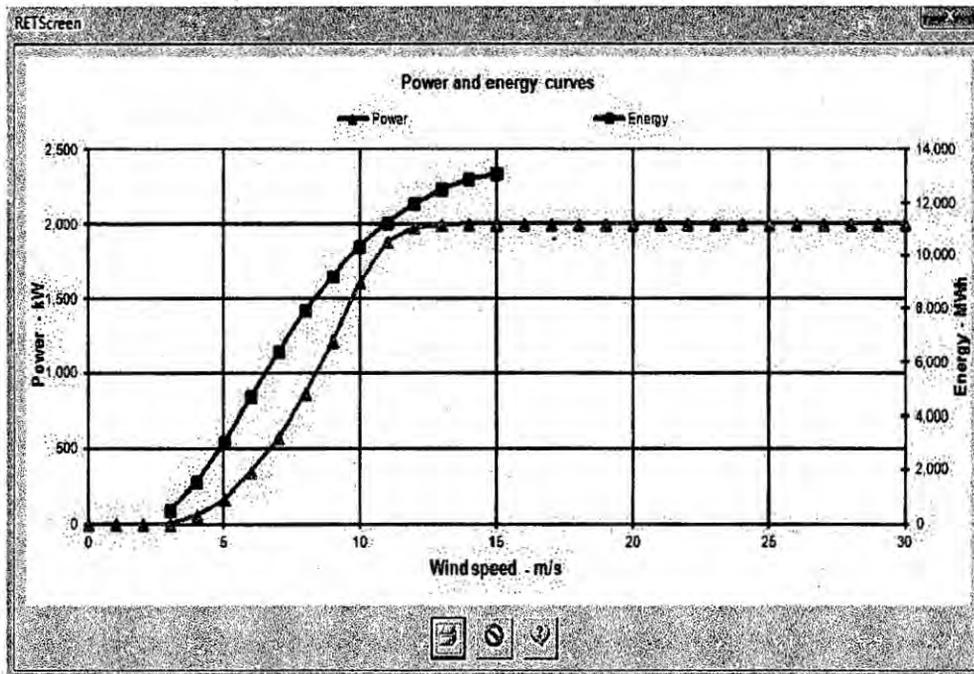
RESULTADOS

5.1 Resultados finales.

5.1.1 Determinación de la potencia de generación eólica con RETScreen.

Teniendo los datos de la predicción de la velocidad del viento, aplicando redes neuronales que son los que se obtuvieron mejores resultados e ingresando dichos datos al software RETScreen, se determinó la potencia salida de cada aerogenerador.

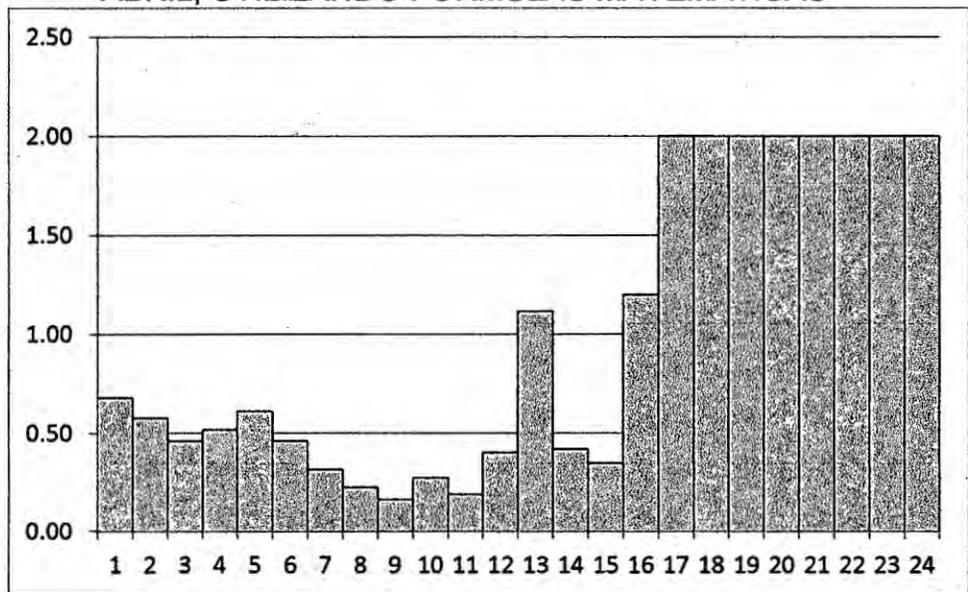
GRAFICA N°5.1
PREDICION DE LA POTENCIA DE SALIDA PARA EL 30 DE ABRIL;
UTILIZANDO RETScreen.



5.1.2 Determinación de la potencia de generación eólica con fórmulas matemáticas.

Teniendo los datos de la predicción de la velocidad del viento, aplicando redes neuronales que son los que se obtuvieron mejores resultados, se determinó la potencia de salida del aerogenerador por fórmulas matemáticas por lo que se obtuvo los siguientes resultados.

GRAFICA N° 5.2
PREDICION DE LA POTENCIA DE SALIDA PARA EL 30 DE ABRIL, UTILIZANDO FORMULAS MATEMATICAS



CAPITULO VI

DISCUSIÓN DE RESULTADOS

6.1 Contratación de hipótesis con los resultados

- a. A la hipótesis de: es posible la investigación de los modelos de predicción de la potencia en países que lideran la producción de energía eólica. Se analizó los modelos de predicción de España y Colombia, las mismas que fueron de base para la presente tesis.
- b. A la hipótesis de: Es posible la identificación de los conceptos fundamentales para el modelamiento de la potencia de inyección de un parque eólico. Se estudió los conceptos de todas las variables de la fórmula para hallar la potencia y a su variación con la altura.
- c. A la hipótesis de: Es posible la identificación de los programas que se emplean en la simular la potencia de salida de un parque eólico. Para hallar la frecuencia se utilizó Excel, para hallar la probabilidad se utilizó Weibull en Minitab17, para hallar la predicción por redes neuronales se utilizó Matlab y para hallar la potencia se utilizó RETScreen.
- d. A la hipótesis de: Es posible la revisión de los modelos lineales y no lineales que se emplean para la predicción de la potencia de un parque eólico con intervalos de confianza. Para hallar la potencia en intervalos de confianza se revisó y comparo los dos

modelos tanto la predicción estadística y como la predicción con redes neuronales.

6.2 Contrastación de resultados con otros estudios similares.

No existen otros estudios similares en la zona de estudio, ni por tiempo ni por espacio (lugar).

VII CONCLUSIONES

Las conclusiones más relevantes que se han obtenido en el presente trabajo son las siguientes:

1. Las características fundamentales del viento superficial en el Cerro Chocan se pueden definir en función de la situación meteorológica favorable que son acompañadas de una gran estabilidad atmosférica en la zona.
2. Se calculó a 80m de altura, el nuevo valor de, la temperatura, velocidad del viento, densidad del viento, teniendo como base los datos a 24m de altura.
3. Como el sentido y la intensidad del viento superficial dependen de las variaciones del gradiente de presión, se tiene que el viento predominante es de SSW a NNE y de Sur a Norte.
4. La variación en el día de la velocidad promedio anual se observa una ventaja satisfactoria ya que los mejores vientos están a partir de las 17 a las 23 horas.
5. La variación mensual de la velocidad del viento para los meses de abril a febrero se tiene buenos vientos y el mes de marzo sufre una leve caída hasta casi 4m/s como promedio.
6. Sobre la frecuencia de las velocidades, es predominante las velocidades de 5m/s a 9m/s y la frecuencia de la velocidad de 3m/s y 2m/s son los que menos se repite, en consecuencia se tiene que la

inyección de potencia eólica va a hacer satisfactoria en el año. Todo ello se corrobora con la gráfica de la probabilidad en Weibull

7. La velocidad media anual de los vientos en el Cerro Chocan es 8.58 m/s a 80m de altura, siendo máxima en Verano (13.9 m/s) y mínima en Invierno (2.42 m/s).
8. Las distribuciones de probabilidad observada de la velocidad se ajustan casi perfectamente a las teóricas de Weibull.
9. Debido a la ubicación siempre se aprecian velocidades de viento superiores a las de las áreas circundantes, por tal motivo los vientos son favorables.
10. La variación de la temperatura en el día promedio anual a 24m, se observó un mínimo de 17°C y un máximo de 28°C y las temperaturas más bajas ocurre de las 23 horas hasta las 09 horas.
11. Se comparó la predicción de la velocidad real y la velocidad por estadística, encontrándose diferencias significativas; en la comparación de la velocidad real y la predicción de la velocidad por redes neuronales (Matlab) se puede decir que es el mejor método ya que las curvas se sobreponen casi su totalidad para las 24 horas del día 30 abril.
12. Existe diferencias en el cálculo de la potencia por RETScreen y el cálculo de la potencia por formulas de la teoría; una vez que el parque eólico empiece a funcionar se verificara cuál de los dos se aproxima a lo real.

Se concluye que es posible la aplicación de redes neuronales para la predicción operativa a escala local del módulo de viento en el Cerro Chocan a 24 horas, a partir de datos meteorológicos anteriores, con unos resultados bastante aceptables.

VIII. RECOMENDACIONES

Dentro de las recomendaciones de la presente tesis, siempre se desea que haya una mejora continua del mismo; por lo que se recomienda a futuros estudiantes que tengan interés en esta tesis, la complementación con los datos de la potencia real de salida una vez que empiece a inyectar potencia al sistema y hacer comparaciones entre los resultados obtenidos. Otra recomendación sería incluir más modelos de predicción por lo que con el transcurrir con los años aparecen nuevos modelos y estos comparar con los resultados obtenidos en la presente tesis.

IX. REFERENCIAS BIBLIOGRAFICAS

- [1] MINISTERIO DE ENERGÍA Y MINAS, Atlas eólico del Perú, Disponible en: <http://dger.minem.gob.pe/AtlasEolico/atlaseolicolibro/presentacion.pdf> . articulo web. Consultada el 22 de Noviembre del 2013.
- [2]. JIMENEZ MEJIA Miguel, **Pronostico de la velocidad del viento utilizando modelos ARIMA y Redes Neuronales Artificiales.** Tesis pos grado. Morelia Michoacán de México. Universidad Michoacana de San Nicolás de Hidalgo. 2011.
- [3]. FERNANDEZ JIMENEZ Luis Alfredo, **Modelos Avanzados para la Predicción a Corto Plazo de la Producción Eléctrica en Parques Eólicos.** Tesis Doctoral. Logroño. Universidad de Rioja. 2007
- [4]. PONCELA BLANCO Marta, **Nuevos modelos de predicción eólica basados en series temporales.** Tesis Doctoral. Valladolid. Universidad de Valladolid. 2012
- [5]. PALOMARES LOSADA Ana María, **Caracterización del régimen de viento y desarrollo de un modelo de predicción eólica a escala local en el estrecho de Gibraltar.** Tesis Doctoral. Madrid. Universidad Complutense de Madrid. 2002.
- [6] GARCIA LOBO Miguel, **Métodos de predicción de la generación agregada de energía eólica.** Tesis Doctoral. Leganés. Universidad Carlos III de Madrid. 2010.
- [7] FIERAS DE LA INGENIERIA. Purchasing from manufacturers in China. Disponible en: <http://www.fierasdelaingenieria.com/los-10-mayores-parques-eolicos-del-mundo/> articulo web. Consultada el 10 de Febrero del 2014
- [8] OSINERGMIN, Central Eólica de Marcona. Disponible en <http://www.osinerg.gob.pe/newweb/uploads/GFE/SupervisionContratos/sup6/61/CE%20Marcona.pdf?3>. articulo web. Consultada el 22 de Noviembre del 2013
- [9] MINISTERIO DE ENERGÍA Y MINAS. Central Eólica de Marcona ,Disponible en: <http://www.minem.gob.pe/minem/archivos/file/DGGAE/ARCHIVOS/e-studios/EIAS%20-%20electricidad/EIA/EIA%20PARQUE%20EOLICO%20MARCONA/4.1%20LBF.pdf> . articulo web. Consultada el 22 de Noviembre del 2013.
- [10] MINISTERIO DE ENERGÍA Y MINAS. Central Eólica de Cupisnique, Disponible en: <http://www.minem.gob.pe/minem/archivos/file/DGGAE/ARCHIVOS/e-studios/EIAS%20-%20electricidad/EIA/EIA%20PARQUE%20EOLICO%20CUPISNIQUE/4.1%20LBF.pdf> . articulo web. Consultada el 22 de Noviembre del 2013.

- [11] MINISTERIO DE ENERGÍA Y MINAS. Central Eólica de Talara, Disponible en: http://www.minem.gob.pe/minem/archivos/file/DGGAE/ARCHIVOS/e_studios/EIAS%20-%20electricidad/EIA/EIA%20PARQUE%20EOLICO%20TALARA/4.1%20LBF.pdf . articulo web. Consultada el 22 de Noviembre del 2013
- [12] MINISTERIO DE ENERGÍA Y MINAS. Central Eólica de Tres hermanas, Disponible en: http://www.minem.gob.pe/minem/archivos/file/DGGAE/ARCHIVOS/e_studios/EIAS%20-%20electricidad/EIA/EIA%20PARQUE%20EOLICO%20TRESHERMANAS/4.1%20LBF.pdf . articulo web. Consultada el 22 de Noviembre del 2013
- [13] DELGADO MARÍN José Pablo, **Energía eólica en Murcia**. Disponible en: www.argem.regionmurcia.net articulo web. Consultada el 18 Octubre del 2013.
- [14] Página docente de Antonio J. Barbero García Profesor Titular Física Aplicada. Disponible en: http://www.uclm.es/profesorado/ajbarbero/FAA/EEOLICA_Febrero_2012_G9.pdf . articulo web. Consultada el 15 Noviembre del 2013.
- [15] Proyecto Naylamp (El niño Anual Y las Anomalías Medidas en el Pacifico). Disponible en: <http://www.naylamp.dhn.mil.pe/>. Artículo web. Consultada el 10 de Octubre del 2014.
- [16] GUIA TÉCNICA DE APLICACIÓN PARA INSTALACIONES DE ENERGÍAS RENOVABLES. Disponible en: www.agenergia.org/files/.../1234272735_GuiaEolica_GobCan.pdf . articulo web. Consultada el 18 Octubre del 2013.
- [17] VENTANAS AL UNIVERSO presentado por la asociación nacional de maestros de ciencias. Disponible en: http://www.windows2universe.org/earth/Atmosphere/wind_speeds.html&lang=sp Artículo web. Consultada el 10 de Octubre del 2014.
- [18] PEREZ RAMIREZ Daniel **Caracterización del aerosol atmosférico en la ciudad de Granada**. Tesis doctoral. Granada. Universidad de Granada. 2010
- [19] Wind Energy Explained Theory, Design and Application. Manwel J.F., McGowan J.G. y Rogers A.L. Amherst: John Wiley& Sons Ltd. 2002. Pags 21-81
- [20] MUR AMADA Joaquín, Curso de energía eólica. España. Departamento de Ingeniería de la Universidad de Zaragoza.2001.
- [21] Laboratorio Costarricense de Metrología. Mediciones de humedad. Disponible en: http://www.lacommet.go.cr/index.php?option=com_content&view=article&id=214:mediciones&catid=86:humedad&Itemid=268. LACOMET 14ABR14 Email: metrologia@lacommet.go.cr articulo web. Consultada el 5 de Enero del 2014.

- [22] FUENTES TOLEDO Amilcar, **Generación eoelectrica con almacenamiento de hidrogeno**, tesis para optar el grado de Maestro en Ingeniería, Universidad nacional Autónoma de México, 2007.
- [23] Generadores Vesta, especificaciones técnicas. Disponible en:
http://www.vestas.com/Files%2FFiler%2FES%2FBrochures%2FV90_2_ES.pdf . Artículo web. Consultada el 10 de Noviembre del 2014.
- [24] SOCHA GARZON Diego Fernando, ORTIZ HERRADA Gilberto Antonio. Aplicación de redes neuronales MLP a la predicción de un paso en series de tiempo. Fundación Universitaria Konrad Lorenz, Facultad de ingeniería de Sistemas. Bogotá 2005.

• ANEXO 1: MATRIZ DE CONSISTENCIA

TÍTULO DEL PROYECTO: Modelamiento de la potencia de inyección, de un parque eólico de 60MW en Piura.

PROBLEMAS	OBJETIVOS	HIPÓTESIS	VARIABLES	MÉTODOLÓGIA
<p>General</p> <p>¿Cómo se modela la potencia de inyección de un parque eólico de 60MW ubicado en la ciudad de Piura?</p> <p>Específico</p> <p>¿Cómo se realiza la predicción de la potencia en países que lideran la producción de energía eólica?</p> <p>¿Cuáles son los conceptos fundamentales para el modelamiento de la potencia de inyección de un parque eólico?</p> <p>¿Cuáles son los programas que se emplean en la predicción de la potencia de un parque eólico?</p> <p>¿Qué modelos lineales y no lineales se pueden emplear para la predicción de la potencia de un parque eólico con intervalos de confianza?</p>	<p>General</p> <p>Modelar la potencia de inyección de un parque eólico de 60MW ubicado en la ciudad de Piura</p> <p>Específicos</p> <p>Identificar los conceptos fundamentales para el modelamiento de la potencia de inyección de un parque eólico.</p> <p>Identificar los programas que se emplean para simular la potencia de salida de un parque eólico.</p> <p>Revisar los modelos lineales y no lineales que se emplean para la predicción de la potencia de un parque eólico con intervalos de confianza.</p>	<p>Hipótesis general: Es posible el modelamiento de la potencia de inyección de un parque eólico de 60MW ubicado en la ciudad de Piura.</p> <p>Hipótesis específicas:</p> <p>a) Es posible la investigación de los modelos de predicción de la potencia en países que lideran la producción de energía eólica.</p> <p>b) Es posible la identificación de los conceptos fundamentales para el modelamiento de la potencia de inyección de un parque eólico.</p> <p>c) Es posible la identificación de los programas que se emplean en la simulación de la potencia de salida de un parque eólico.</p> <p>d) Es posible la revisión de los modelos lineales y no lineales que se emplean para la predicción de la potencia de un parque eólico con intervalos de confianza.</p>	<p>Variable: X</p> <p>Parque eólico de 60 MW en Piura.</p> <p>Dimensión:</p> <p>Conceptos fundamentales para el modelamiento de la potencia de inyección de un parque eólico.</p> <p>Indicadores:</p> <ul style="list-style-type: none"> - Potencia nominal(MW) X_1 - Temperatura (°C) X_2 - Densidad del aire X_3 - Velocidad del viento(m/s) X_4 - Nivel de rugosidad terreno X_5 - Presión a nivel del mar X_6 - Aceleración de la gravedad X_7 - Área del rotor eólico X_8 - Humedad X_9 <p>Variable: Y</p> <p>Modelamiento de la potencia de inyección.</p> <p>Dimensión:</p> <p>Revisar los modelos lineales y no lineales que se emplean para la predicción de la potencia de un parque eólico con intervalos de confianza</p> <p>Indicadores:</p> <ul style="list-style-type: none"> ✓ % de error de la velocidad del viento de los modelos lineales y no lineales ✓ Coeficiente de mejora de la predicción de la velocidad del viento. <p>Y_1 Y_2</p>	<p>General</p> <p>El tipo de investigación es Descriptiva prospectiva, porque está basado a la recolección de datos y características de los fenómenos meteorológicos para poder determinar la predicción de la velocidad del viento. Es una investigación correlacional porque permite encontrar la relación entre la velocidad del viento y la potencia de generación eólica.</p> <p>Específico</p> <p>Las etapas para el desarrollo de la presente investigación son los siguientes:</p> <ol style="list-style-type: none"> a. Recolección de datos de la velocidad de los vientos en Piura, se tomara como fuente de información del Instituto Geofísico del Perú y/o información de la empresa NORWING (confidencial). b. Corrección de datos a la zona del proyecto. c. Método estadístico para la predicción del viento, considerando años, estaciones y meses típicos. d. Software para la predicción del viento. e. Método matemático por determinar mejorando los errores de predicción de los dos métodos anteriores para la predicción del viento. f. Software ya establecido para determinar la potencia de generación eólica.

ANEXO 2: FUNCIÓN DE DISTRIBUCIÓN DE WEIBULL

La expresión analítica más utilizada, dentro del campo de la energía eólica, para representar la probabilidad de velocidades de viento (función densidad de probabilidad), es la distribución de Weibull, que se expresa mediante la fórmula:

$$p(v) = a c v^{(c-1)} \exp (-a v^c)$$

en la que $v \geq 0$, $a > 0$ y $c > 0$, siendo c el parámetro de forma (adimensional) y $A = a^{-1/c}$ el factor de escala, que tiene unidades de velocidad, cuyo valor se acerca a la velocidad media.

El parámetro C nos da el sesgo de la distribución. Para $C > 1$ la función tiene un máximo lejos del origen. Para $C=1$ la función es exponencial. Para $C=2$ es una distribución de Rayleigh y para $C=3,5$ es una distribución normal. El valor más usual de C es de $2 \pm 0,5$.

La distribución de frecuencias acumuladas de Weibull $P(v)$, que da la probabilidad de ocurrencia de una velocidad de viento menor o igual que " v " se obtiene integrando la ecuación anterior, resultando:

$$P(v) = 1 - \exp (-a v^c)$$

Los parámetros estadísticos de la distribución, en función de los parámetros de Weibull son:

$$V = a^{-1/c} \Gamma(1+1/c)$$

$$\sigma^2 = a^{-2/c} [\Gamma(1+2/c) - \Gamma^2(1+1/c)]$$

$$\langle v^2 \rangle = a^{-2/c} \Gamma(1+2/c)$$

$$\text{Mediana} = a^{-1/c} (\ln 2)^{1/c}$$

$$\text{Moda} = a^{-1/c} [(c-1)/c]^{1/c}$$

Siendo:

\bar{v} = velocidad media observada

σ^2 = varianza observada

$\langle v^2 \rangle$ = promedio de la velocidad cuadrática

$\Gamma(x)$ = función gamma, definida:

$$\Gamma(x) = \int_0^{\infty} e^{-x} x^{n-1} dx; \quad \Gamma(x+1) = x\Gamma(x) \text{ y } \Gamma(n) = (n-1)!$$

Donde x es un número real y n entero positivo.

El método de cálculo seguido para la determinación de los parámetros de

Weibull es el siguiente:

Se definen dos funciones:

$$x_1(c) = \Gamma\left(1 + \frac{2}{c}\right) \Gamma^{-2}\left(1 + \frac{1}{c}\right)$$

$$x_2(c) = \Gamma\left(1 + \frac{1}{c}\right) \Gamma^{-1}\left(1 + \frac{2}{c}\right)$$

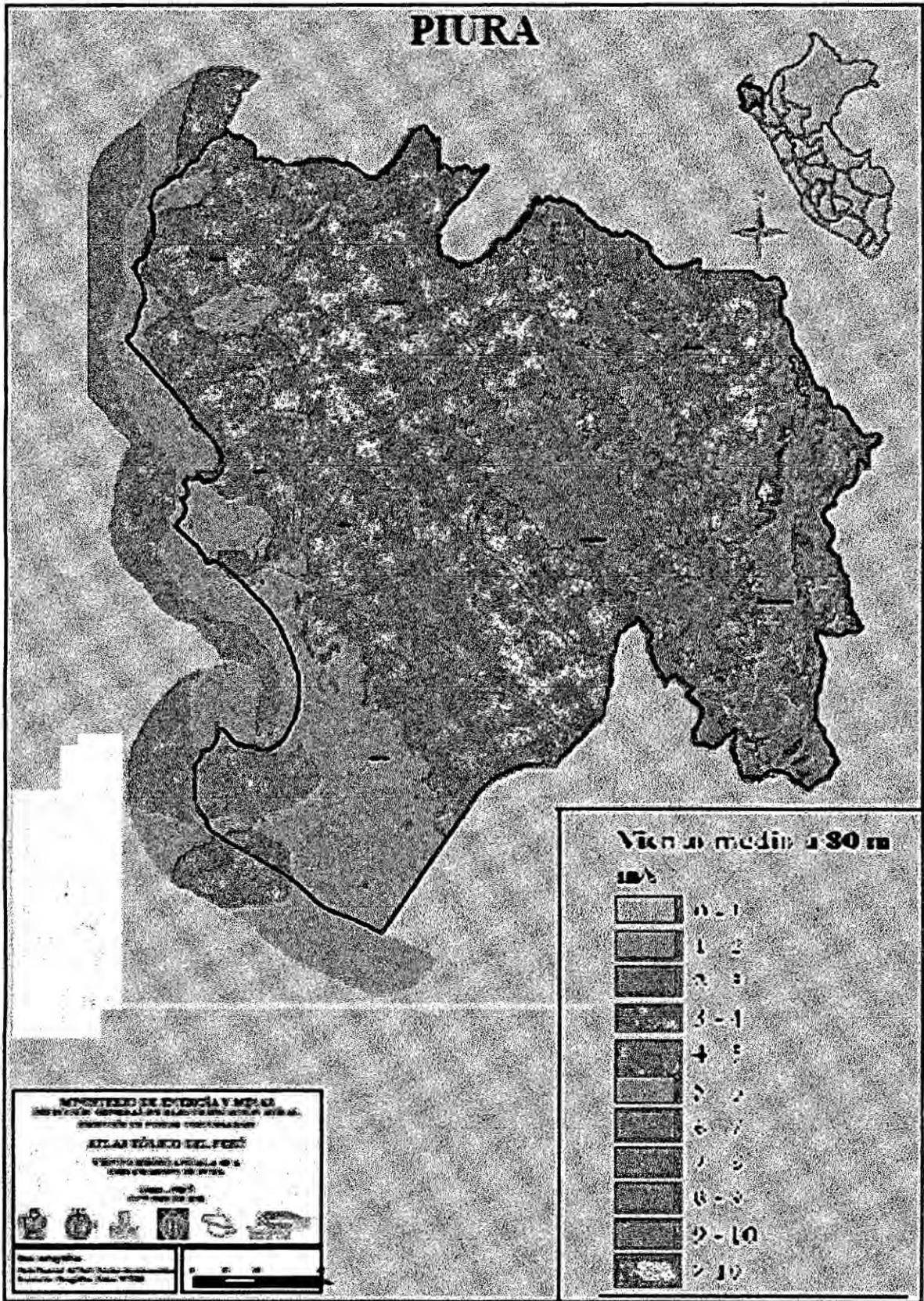
Llegándose a las siguientes expresiones para "a" y "c"

$$x_1(c) = \frac{\langle v^2 \rangle}{\langle v \rangle^2}$$

$$A = a^{-1/c} = \frac{\langle v^2 \rangle}{\langle v \rangle} x_2(c)$$

Este sistema de ecuaciones se resuelve por iteración, mediante aproximaciones sucesivas, suponiendo un valor de "C" y comprobando si el valor de "a" es el apropiado. El proceso de iteración termina cuando el cambio relativo entre aproximaciones sucesivas en los parámetros es menor del 1%.

**ANEXO 3:
ATLAS EÓLICO DE PIURA**



**ANEXO 4:
CROQUIS DE LA ZONA DE ESTUDIO**

