



Instituto Politécnico Nacional

Centro de Investigación en Computación



**ESTIMACIÓN DEL CRECIMIENTO DE UN BOSQUE DE PINO
(PINUS COOPERÍ BLANCO) POR MEDIO DE UNA RED NEURONAL
BACKPROPAGATION ENTRENADA CON ÍNDICES
DE COMPETENCIA INDEPENDIENTES DE LA DISTANCIA**

T E S I S
QUE PARA OBTENER EL GRADO DE
DOCTOR EN CIENCIAS DE LA COMPUTACIÓN
P R E S E N T A
JESÚS CELIS PORRAS

DIRECTOR DE TESIS: DR. JUAN LUIS DÍAZ DE LEÓN SANTIAGO
CO - DIRECTOR : DR. JOSÉ ALBERTO GALLEGOS INFANTE

MÉXICO, D.F.

SEPTIEMBRE 2006



INSTITUTO POLITECNICO NACIONAL
COORDINACION GENERAL DE POSGRADO E INVESTIGACION

ACTA DE REVISIÓN DE TESIS

En la Ciudad de México, D. F. siendo las 16:30 horas del día 26 del mes de Octubre de 2005 se reunieron los miembros de la Comisión Revisora de Tesis designada por el Colegio de Profesores de Estudios de Posgrado e Investigación del:

Centro de Investigación en Computación

para examinar la tesis de grado titulada:

" ESTIMACIÓN DEL CRECIMIENTO DE UN BOSQUE DE PINO (PINUS COOPERÍ BLANCO) POR MEDIO DE UNA RED NEURONAL BACKPROPAGATION ENTRENADA CON ÍNDICES DE COMPETENCIA INDEPENDIENTES DE LA DISTANCIA "

Presentada por el alumno:

CELIS

PORRAS

JESÚS

Apellido paterno

materno

nombre(s)

Con registro:

A	9	7	1	6	2	9
---	---	---	---	---	---	---

aspirante al grado de: **DOCTOR EN CIENCIAS DE LA COMPUTACIÓN**

Después de intercambiar opiniones los miembros de la Comisión manifestaron **SU APROBACIÓN DE LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA

Presidente

DR. LUIS PASTOR SÁNCHEZ FERNÁNDEZ

Secretario

DR. OLEKSIY POGREBANYAK

Primer vocal

DR. JUAN LUIS DÍAZ DE LEÓN SANTIAGO

Segundo vocal

DR. JOSÉ ALBERTO GALLEGOS INFANTE

Tercer vocal

DR. CORNELIO YÁÑEZ MÁRQUEZ

EL PRESIDENTE DEL COLEGIO



DR. HUGO CÉSAR COYOTE ESTRADA

DIRECCION

*“Dichoso el hombre que medita
la sabiduría y atiende a la
inteligencia.”*

Eclesiástico 14:21,22

Dedicatoria:

*A ti, ¡Oh Dios!, se te debe la alabanza en
Sión, Y a ti el cumplimiento de los votos.*

Salmos 65: 2

A mi Esposa:

Elsa

y nuestra

Pequeñita

Ana Paulina.

*Como muestra de gratitud a su apoyo y
comprensión*

Agradecimientos:

Al Instituto Politécnico Nacional y al Centro de Investigación en Computación.

Por la formación que me proporciono dándome la oportunidad de ser una persona útil a la comunidad.

A mis Padres:

A Quienes gracias a su ejemplo y por apoyo y dedicación este logro.

Lic. Luis Celis M.

Sra. Imelda Porras de C.

A mis hermanos:

Por su apoyo que siempre me han brindado.

Lic. Luis Celis Porras.

Sra. Ma. Guadalupe Celis Porras.

Dr. Jorge Armando Celis Porras.

Lic. Imelda Celis Porras.

A mis Directores de tesis:

Por sus consejos y orientación.

Dr. José Alberto Gallegos Infante Dr. Juan Luis Díaz de León.

A mis Maestros del Centro de investigación.

Por su guía y enseñanza.

A mis amigos y compañeros que siempre me alentaron a continuar el esfuerzo en la culminación de la tarea fijada.

Dra. Elsa Rubio Espino.

M.C. Alejandro Rubio Martinez.

M.C. Norma Angélica Noguez de M.

ESTIMACIÓN DEL CRECIMIENTO DE UN BOSQUE DE PINO (*PINUS COOPERÍ* BLANCO) POR MEDIO DE UNA RED NEURONAL BACKPROPAGATION ENTRENADA CON ÍNDICES DE COMPETENCIA INDEPENDIENTES DE LA DISTANCIA.

RESUMEN

El desarrollo de las técnicas de manejo de los bosques es similar en muchos lugares del mundo. Sin embargo; pocas teorías silvícolas pueden ser generalizadas para una región forestal específica, que permitan describir los procesos en el desarrollo futuro de los ecosistemas forestales; debido a que las condiciones climáticas, suelos y especies son muy diferentes. En estos casos se necesita conocer la dinámica de crecimiento en cada localidad, con base en las interacciones suelo-clima-vegetación y así conocer sus efectos en el desarrollo y crecimiento de los bosques naturales.

El tipo de bosques que caracterizan a la mayoría de los bosques en México, así como a la región boscosa del Estado de Durango es del tipo llamado irregular los cuales presentan una diversidad en edad y en estructura y/o tamaño, o bien en su distribución espacial en el sitio.

En estos bosques complejos, el conocer la producción presente y predecir el rendimiento futuro es difícil debido a la heterogeneidad, por lo cual se requiere más información silvícola y modelos basados en metodologías que puedan abordar de forma más adecuada la complejidad de un sistema.

En este trabajo se estudia el desempeño de las redes neuronales entrenadas con índices de competencia independientes de la distancia en la predicción del crecimiento en diámetro y altura de un bosque irregular de pino de la especie *Pinus Cooperí* Blanco.

El crecimiento de un bosque de coníferas integrado por individuos de diferente edad y tamaño, suele ser caracterizado por índices de competencia independientes de la distancia, los cuales son construidos en base a las dimensiones del individuo tales como altura, diámetro cuadrático, área basal, amplitud de copa, longitud copa y son relacionadas a las respectivas dimensiones de la comunidad donde se encuentran.

Los métodos de análisis estadístico pueden ser utilizados para entender un sistema biológico de nuestro interés el cual es demasiado complejo para ser modelado analíticamente.

La información del comportamiento del sistema es recolectada de la medición de variables seleccionadas. Esas observaciones son entonces analizadas aplicando métodos estadísticos o infiriendo conocimiento del sistema. Los resultados del análisis dependerán del conocimiento a priori acerca del sistema y de los métodos empleados durante el análisis. Esencialmente, todas las opciones consideradas durante el análisis, empezando desde la recolección de la información, y terminando con el análisis de resultados obtenidos de un modelo estadístico tienen efecto en la inferencia de conocimiento.

Las redes neuronales constituyen una clase de modelos estadísticos, los cuales pueden ser utilizados en una amplia gama de aplicaciones, desde análisis exploratorio

de datos y visualización hasta problemas de clasificación y regresión. En el modelado con redes neuronales, la meta es construir un modelo basado en datos de observaciones de un sistema, los cuales representen las propiedades esenciales del sistema de nuestro interés.

La red neuronal backpropagation entrenada con índices competencia independientes de la distancia pronosticó la velocidad de crecimiento del diámetro y altura del *Pinus Cooperi* Blanco con 90% y 94,5% de aproximación dentro de un intervalo de confianza del 90%.

ÍNDICE	Página
ÍNDICE DE FIGURAS	iv
ÍNDICE DE GRÁFICAS	iv
ÍNDICE DE TABLAS	vi
INTRODUCCIÓN	1
RESUMEN	8
ABSTRACT	10
CAPITULO I	
1. ANTECEDENTES, JUSTIFICACIÓN, MOTIVACIÓN, HIPÓTESIS, Y OBJETIVO	
1.1. ANTECEDENTES	11
1.1.1. Definición del problema	11
1.1.1.1 Crecimiento de los árboles	11
1.1.1.2 Un bosque como un sistema.	18
1.1.2. Estado del Arte	19
1.1.2.1. ANN en Predicción y Clasificación.	26
1.1.2.2. ANN en el modelado de la dinámica de crecimiento de un bosque.	26
1.2. JUSTIFICACIÓN	28
1.3. MOTIVACIÓN	29
1.4. HIPÓTESIS	30
1.5. OBJETIVO	31
CAPITULO II	
2. METODOLOGÍA, MATERIALES Y RESULTADOS	
2.1 MATERIALES	32
2.1.1. Descripción del Sitio	32
2.1.2. Selección de la muestra de validación	33
2.1.3. Determinación del tamaño de muestra.	37
2.2. METODOLOGÍA	38
2.2.1. Red Neuronal Backpropagation	38
2.2.1.1 Algoritmo de Aprendizaje	40
2.2.1.2. Aplicaciones de las ANN	41
2.2.1.3. Utilización de modelos FANN.	42
2.2.1.4. Algoritmo de aprendizaje. Método de optimización.	42
2.2.1.5. Ventajas del uso de redes neuronales.	43
2.2.1.6. Limitaciones del uso de redes neuronales.	44
2.2.2. Sistemas neuro-difusos (ANFIS)	44
2.2.2.1. ANFIS, estructura y funcionamiento.	45
2.2.2.2. Entrenamiento de ANFIS	48

	Página
2.2.2.3. Sistemas híbridos neuro difusos.	49
2.2.2.4. Taxonomía de los sistemas neuro-difusos (NF)	50
2.2.2.5 Ventajas de los sistemas neuro-difusos	50
2.2.2.6. Limitaciones de los sistemas neuro-difusos	51
2.2.3. Parámetros para entrenamiento de la red neuronal.	52
2.2.4 Software utilizado.	56
2.2.5. Procedimiento de Solución	56
2.2.5.1 Selección de Variables.	56
2.2.5.2 Pre- tratamiento de datos de entrenamiento.	56
2.2.5.3. Selección de Arquitectura de la ANN.	58
2.2.5.4. Evaluación de la capacidad de generalización de la ANN.	64
2.3 RESULTADOS	66
2.3.1. Parámetros utilizados	66
2.3.2. Estimación del crecimiento en diámetro por medio de una red neuronal backpropagation	67
2.3.2.1. Predicción de crecimiento en Diámetro	67
2.3.2.2. Predicción de crecimiento en Diámetro	69
2.3.2.3. Predicción de crecimiento en Diámetro	71
2.3.2.4. Predicción de crecimiento en Diámetro	73
2.3.2.5. Predicción de crecimiento en Diámetro	75
2.3.2.6. Predicción de crecimiento en Diámetro	77
2.3.2.7. Predicción de crecimiento en Diámetro	79
2.3.3. Estimación del crecimiento en altura por medio de una red neuronal backpropagation	81
2.3.3.1. Predicción de crecimiento en altura	81
2.3.3.2. Predicción de crecimiento en altura	83
2.3.3.3. Predicción de crecimiento en altura	85
2.3.3.4. Predicción de crecimiento en altura	87
2.3.3.5. Predicción de crecimiento en altura	89
2.3.3.6. Predicción de crecimiento en altura	91
2.3.4. Estimación del crecimiento en diámetro utilizando ANFIS	93
2.3.4.1. Predicción de velocidad de crecimiento en diámetro.	93
2.3.4.2. Predicción de velocidad de crecimiento en diámetro.	94
2.3.4.3. Predicción de velocidad de crecimiento en diámetro.	95
2.3.4.4. Predicción de velocidad de crecimiento en diámetro.	96
2.3.4.5. Predicción de velocidad de crecimiento en diámetro.	97
2.3.4.6. Predicción de velocidad de crecimiento en diámetro.	98
2.3.4.7. Predicción de velocidad de crecimiento en diámetro.	99
2.3.4.8. Predicción de velocidad de crecimiento en diámetro.	100
2.3.4.9. Predicción de velocidad de crecimiento en diámetro.	101
2.3.4.10. Predicción de velocidad de crecimiento en diámetro.	102
2.3.4.11. Predicción de velocidad de crecimiento en diámetro.	103
2.3.4.12. Predicción de velocidad de crecimiento en diámetro.	104
2.3.4.13. Predicción de velocidad de crecimiento en diámetro.	105

	Página
2.3.4.14. Predicción de velocidad de crecimiento en diámetro.	106
2.3.5. Estimación del crecimiento en altura utilizando ANFIS	107
2.3.5.1. Predicción de velocidad de crecimiento en altura.	107
2.3.5.2. Predicción de velocidad de crecimiento en altura.	108
2.3.5.3. Predicción de velocidad de crecimiento en altura.	109
2.3.5.4. Predicción de velocidad de crecimiento en altura.	110
2.3.5.5. Predicción de velocidad de crecimiento en altura.	111
2.3.5.6. Predicción de velocidad de crecimiento en altura.	112
2.3.5.7. Predicción de velocidad de crecimiento en altura.	113
2.3.5.8. Predicción de velocidad de crecimiento en altura.	114
CAPITULO III	
3. ANÁLISIS DE RESULTADOS, CONCLUSIONES, RECOMENDACIONES Y TRABAJO A FUTURO	
3.1 ANÁLISIS DE RESULTADOS	115
3.2. CONCLUSIONES	116
3.3 RECOMENDACIONES	117
3.4 TRABAJO A FUTURO	118
BIBLIOGRAFÍA	119
ANEXO	125

ÍNDICE DE FIGURAS

Figura		Página
1	La región sombreada muestra parte del Estado de Durango, México poblada por la especie endémica <i>Pinus Cooperi</i> Blanco	2
2	Distribución de la región boscosa en México	11
3	Partes que forman el tronco de un árbol	12
4	Curva de crecimiento arbóreo	14
5	Suelo de un bosque	15
6	Agujas (Hojas) de pino	18
7	Neurona Artificial tipo McCulloch-Pitts	22
8	Clasificación de Redes	23
9	Bosque antiguo. Sierra Madre Occidental del estado de Durango	29
10	<i>Pinus Cooperi</i> Blanco	31
11	Precipitación media anual de lluvia	32
12	Sitio experimental “cielito azul” de estudio del <i>Pinus Cooperi</i> Blanco	33
13	Se ha tratado de modelar el funcionamiento del cerebro por medio de las redes neuronales artificiales	38
14	Una ANN feedforward de para aproximar $h(.)$	39
15	Razonamiento borroso tipo Takagi-Sugeno	46
16	Estructura de una red ANFIS	47
17	$bell(x)$	48

ÍNDICE DE GRÁFICAS

Gráfica		Página
1	Histograma de frecuencias de la velocidad de crecimiento en diámetro de la población de un bosque de <i>Pinus Cooperi</i> Blanco	34
2	Histograma de frecuencias de la velocidad de crecimiento en altura de la población de un bosque de <i>Pinus Cooperi</i> Blanco	36
3	Escalamiento $[-1 +1]$ de patrones de entrenamiento	57
4	Escalamiento media 0 varianza 1 de patrones de entrenamiento	58
5	Entrenamiento con función sigmoideal logística	59
6	Entrenamiento con función tangencial	60
7	Método Gradiente Conjugado	61
8	Método RPROP	62
9	Red una capa oculta	63
10	Red tres capas ocultas	64
11	Entrenamiento de la red (ejemplo 2.3.2.1.)	68
12	Predicción de crecimiento en diámetro (ejemplo 2.3.2.1.)	68
13	Entrenamiento de la red (ejemplo 2.3.2.2)	69
14	Predicción de crecimiento en diámetro (ejemplo 2.3.2.2)	70
15	Entrenamiento de la red (ejemplo 2.3.2.3)	71
16	Predicción de crecimiento en diámetro (ejemplo 2.3.2.3)	72
17	Entrenamiento de la red (ejemplo 2.3.2.4)	73
18	Predicción de crecimiento en diámetro (ejemplo 2.3.2.4)	74
19	Entrenamiento de la red (ejemplo 2.3.2.5)	75
20	Predicción de crecimiento en diámetro (ejemplo 2.3.2.5)	76
21	Entrenamiento de la red (ejemplo 2.3.2.6)	77
22	Predicción de crecimiento en diámetro (ejemplo 2.3.2.6)	78

Gráfica		Página
23	Entrenamiento de la red (ejemplo 2.3.2.7)	79
24	Predicción de crecimiento en diámetro (ejemplo 2.3.2.7)	80
25	Entrenamiento de la red (ejemplo 2.3.3.1)	81
26	Predicción de crecimiento en altura (ejemplo 2.3.3.1)	82
27	Entrenamiento de la red (ejemplo 2.3.3.2)	83
28	Predicción de crecimiento en diámetro (ejemplo 2.3.3.2)	84
29	Entrenamiento de la red (ejemplo 2.3.3.3)	85
30	Predicción de crecimiento en altura (ejemplo 2.3.3.3)	86
31	Entrenamiento de la red (ejemplo 2.3.3.4)	87
32	Predicción de crecimiento en altura (ejemplo 2.3.3.4)	88
33	Entrenamiento de la red (ejemplo 2.3.3.5)	89
34	Predicción de crecimiento en altura (ejemplo 2.3.3.5)	90
35	Entrenamiento de la red (ejemplo 2.3.3.6)	91
36	Predicción de crecimiento en altura (ejemplo 2.3.3.6)	92
37	Predicción de crecimiento en diámetro (ejemplo 2.3.4.1)	93
38	Predicción de crecimiento en diámetro (ejemplo 2.3.4.2)	94
39	Predicción de crecimiento en diámetro (ejemplo 2.3.4.3)	95
40	Predicción de crecimiento en diámetro (ejemplo 2.3.4.4)	96
41	Predicción de crecimiento en diámetro (ejemplo 2.3.4.5)	97
42	Predicción de crecimiento en diámetro (ejemplo 2.3.4.6)	98
43	Predicción de crecimiento en diámetro (ejemplo 2.3.4.7)	99
44	Predicción de crecimiento en diámetro (ejemplo 2.3.4.8)	100
45	Predicción de crecimiento en diámetro (ejemplo 2.3.4.9)	101
46	Predicción de crecimiento en diámetro (ejemplo 2.3.4.10)	102
47	Predicción de crecimiento en diámetro (ejemplo 2.3.4.11)	103
48	Predicción de crecimiento en diámetro (ejemplo 2.3.4.12)	104
49	Predicción de crecimiento en diámetro (ejemplo 2.3.4.13)	105
50	Predicción de crecimiento en diámetro (ejemplo 2.3.4.14)	106
51	Predicción de crecimiento en altura (ejemplo 2.3.5.1)	107
52	Predicción de crecimiento en altura (ejemplo 2.3.5.2)	108
53	Predicción de crecimiento en altura (ejemplo 2.3.5.3)	109
54	Predicción de crecimiento en altura (ejemplo 2.3.5.4)	110
55	Predicción de crecimiento en altura (ejemplo 2.3.5.5)	111
56	Predicción de crecimiento en altura (ejemplo 2.3.5.6)	112
57	Predicción de crecimiento en altura (ejemplo 2.3.5.7)	113
58	Predicción de crecimiento en altura (ejemplo 2.3.5.8)	114

ÍNDICE DE TABLAS

Tabla		Página
1	Muestra de validación de la red neuronal backpropagation entrenada para predecir crecimiento en diámetro	35
2	Muestra de validación de la red neuronal backpropagation entrenada para predecir crecimiento en altura	36
3	Comparación entre las técnicas de RNA y Lógica Difusa	49

INTRODUCCIÓN

El estado de Durango cuenta con vastos recursos naturales que son fuente de vida en su más amplio concepto, es la primera reserva nacional forestal.

Localizándose aproximadamente en sus zonas boscosas 250 mil habitantes que sobreviven del recurso forestal.

Donde de acuerdo con la Asociación de Industriales Forestales de Durango (AIFDAC), la mayoría de los predios forestales son de propiedad particular, un total de 662, mientras que 173 son ejidales y 64 comunales.

Esta gente dedicada al aprovechamiento de los recursos forestales ha mostrando una gran interés por el manejo y aprovechamiento racional del bosque involucrándose, en el proceso de certificación del buen manejo forestal.

El estado de Durango es el que tiene el proceso de mayor dinámica en la certificación y para 2002, 18 comunidades habían iniciado los primeros pasos para entrar al proceso y la mayor parte de ellas tiene una relación estrecha con industrias madereras que están ya preocupadas por las señales del mercado internacional en el sentido de que los productos deben salir con el certificado de buen manejo forestal.

Sin embargo la industria en el sector forestal sufre una aguda crisis traducida en cierre de plantas, pérdida de empleos, de mercados, ante la creciente invasión de los productos maderables de importación y el rezago tecnológico del sector.

Las importaciones de productos forestales provenientes de Sudamérica, Norteamérica y el sudeste asiático han invadido el mercado nacional y han puesto en situación muy delicada a las empresas nacionales. Con esta situación de crisis, resulta impensable pagar un sobreprecio a las comunidades certificadas proveedoras de materias primas.

La importación de productos forestales no se limita sólo a pulpa y papel sino que también se está incrementando la importación de madera sólida.

Algunas empresas que se abastecían de madera de ejidos y comunidades, hoy en día han dejado de hacerlo ya que encuentran mucho mejores precios en la madera proveniente de plantaciones chilenas.

Frente a una oferta externa a precios baratos, difícilmente la industria procesadora está en la posibilidad de pagar un sobreprecio para la madera local.

Las comunidades y ejidos en Durango son principalmente productores de materias primas forestales, son productores de madera en rollo y algunos de ellos producen madera aserrada. En su mayoría estas empresas no están en condiciones de entrar a los grandes mercados por su actual nivel tecnológico y nivel de integración vertical.

Así en la búsqueda de acceder a los mercados con productos forestales de alta calidad y a precios internacionales se hace necesario entre las diversas medidas para mejorar la competitividad. La implementación de nuevas tecnologías en el manejo y aprovechamiento del bosque. Una de las necesidades a satisfacer es el uso de nuevos

métodos más eficientes de predicción, con los cuales se logre observar el comportamiento de las variables que describen el crecimiento del bosque.

Siendo sumamente importante determinar la velocidad de crecimiento del diámetro y altura del árbol pues son las dimensiones que determinan la cantidad de madera que se va a obtener del árbol.

El estado de Durango es el único lugar del mundo donde crece la especie de pino *Pinus Cooperi*, Blanco (figura 1) árbol estudiado muy poco, prácticamente no existe información sobre esta especie. Sin embargo es de una gran importancia en su explotación por su aprovechamiento en la producción de madera y uso en la producción de hojas de triplay.

Pinus cooperi



Figura 1. La región sombreada muestra parte del Estado de Durango, México poblada por la especie endémica *Pinus Cooperi* Blanco.

A causa de su gran importancia como recurso maderable y por el casi absoluto desconocimiento como especie forestal, es que se hace necesario desarrollar un estudio en busca de hacer más eficiente su aprovechamiento.

Así al observar un bosque natural se presenta una amplia variedad de estructuras entre los individuos que lo forman, las predicciones en el crecimiento se dificultan, en estos

casos los modelos de crecimiento de árboles individuales son quizás la mejor opción para describir el crecimiento bajo estas condiciones tan complejas (Hilt, 1987). En estos rodales irregulares, con frecuencia se emplea modelos de crecimiento de árboles individuales en función de la distancia (Bruce, 1987). Sin embargo, en México la estimación de la competencia es un tema nuevo y muchas veces no se dispone de la información suficiente para su estimación; en este sentido, resulta de particular importancia para la silvicultura considerar índices que no incorporan la distancia, sino más bien las dimensiones de los árboles (diámetro del tronco, amplitud de copa, altura de copa, etc.) para estudiar la competencia independiente de la distancia (Glover, and Hool 1979).

La estimación de la competencia es de suma importancia para modelar el crecimiento y rendimiento en bosques naturales para aplicar un manejo más racional y predecir con precisión la dinámica de crecimiento.

En la toma de decisiones del manejo forestal, es indispensable conocer la predicción del crecimiento y rendimiento del bosque para poder definir en base a ello su comportamiento a través del tiempo (Hagem, 1962, Boosley, 1998, A. Law, 1990).

Sin embargo los modelos que consideran el efecto de competencia han sido poco utilizados en México, para el manejo de bosques de clima templado frío, (ver figura 2). En este trabajo de investigación se recolectaron datos dimensionales de los individuos de todo un bosque experimental ubicado en el estado de Durango, con los cuales se buscó determinar la capacidad de una red neuronal retroalimentada (backpropagation), entrenada con índices de competencia independientes de la distancia para estimar la velocidad de crecimiento en diámetro y altura de los individuos de un bosque. Dichos índices tratan de establecer la relación entre el individuo y su colectividad, intentan reflejar la lucha que enfrenta un individuo con sus vecinos competidores por los recursos para su subsistencia y considerando estos factores de competencia como fundamentales en la estimación de las velocidades de crecimiento de las dimensiones de un individuo como diámetro y altura. Para llevar a cabo este estudio se hace uso de una red neuronal con retro-propagación del error (Red Neuronal Feedforward FANN) y de una red neuro difusa ANFIS buscando sacar provecho de sus cualidades como son la de recoger información de las experiencias y utilizarlas para actuar ante situaciones futuras así como la capacidad para abstraer la información útil, más allá de los casos particulares y determinando cual de las dos metodologías resulta la más adecuada en la predicción del crecimiento de un bosque de *Pinus Cooperi* Blanco.

Las partes que integran este trabajo son:

CAPITULO I

1. 1. ANTECEDENTES, JUSTIFICACIÓN, MOTIVACIÓN, HIPÓTESIS, Y OBJETIVO.

1.1 ANTECEDENTES.

1.1.1 Definición del problema. El crecimiento de los árboles es el resultado de su tendencia al crecimiento ilimitado (dependiente de su potencial biótico, actividad fotosintética...) y de las restricciones impuestas por el medio (competencia con otros individuos, proceso de envejecimiento...). Si bien se trata de factores cuyo análisis es

complejo, se acepta que la interacción espacial con los árboles cercanos (denominada competencia) tiene el mayor peso en el proceso.

1.1.2. Estado del Arte. En años recientes, ha crecido el interés por el empleo de las redes neuronales artificiales (ANN) en varios campos de la ciencia (Maren 1990, Swingler 1996). Las redes neuronales constituyen una clase de modelos estadísticos, los cuales pueden ser utilizados en una amplia gama de aplicaciones, desde análisis exploratorio de datos y visualización hasta problemas de clasificación y regresión. En el modelado con redes neuronales, la meta es construir un modelo basado en datos de observaciones de un sistema, los cuales representen las propiedades esenciales del sistema de nuestro interés.

Las ANN también emergen como una aproximación alternativa en el modelado de fenómenos complejos no lineales en el área forestal (McRoberts et al. 1991; Gimblett and Ball 1995; Lek et al. 1996; Atkinson and Tatnall 1997). Por el potencial en su capacidad predictiva las ANN pueden proporcionar soluciones óptimas en el manejo de problemas de recursos forestales.

1.2. JUSTIFICACIÓN.

México es un país fuertemente importador de productos forestales y la balanza comercial del sector forestal es deficitaria en 1,850 millones de dólares (esta cantidad representa el 46 por ciento del total del déficit de la balanza comercial del país).

Los principales estados productores de madera en México son Durango en primer lugar con una aportación del 26 por ciento del total nacional.

El estado de Durango posee el 25 por ciento de los recursos forestales nacionales en una superficie de 9.1 millones de hectáreas; siendo la entidad más importante como productora forestal en el país y ocupa el primer lugar en la explotación de la especie maderable de pino.

Por lo que se hace necesario desarrollar un modelo que logre describir de forma precisa el crecimiento de un bosque, para así poder desarrollar nuevas técnicas silvícolas para un mejor manejo del bosque y así poder:

- Reducir las importaciones forestales del país
- Impulsar el desarrollo de la Industria forestal del estado de Durango.
- Lograr un mayor nivel de vida de la población dedicada al aprovechamiento del sector forestal.
- Hacer un uso adecuado del bosque en su explotación.
- Optimizar la producción y calidad de la madera.

1.3. MOTIVACIÓN.

Beneficios que se esperan alcanzar al contar con una herramienta que describa de forma detallada la evolución de un bosque de la especie de *Pinus Cooperi* Blanco.

1.4. HIPÓTESIS.

Fundamentación del tipo de herramienta a utilizar en el desarrollo del modelo que logre predecir el crecimiento de un bosque de la especie de *Pinus Cooperi* Blanco.

1.5. OBJETIVO.

Predecir la velocidad de crecimiento en diámetro y altura de la especie de pino (*Pinus Cooperi* Blanco) de un bosque haciendo uso de una red neuronal entrenada con de índices de competencia independientes de la distancia.

CAPITULO II

2. METODOLOGÍA, MATERIALES Y RESULTADOS.

2.1. MATERIALES.

2.1.1. Descripción del sitio. El estudio se efectuó en el área experimental, Cielito Azul, ubicada en el predio las “Veredas”, lote 4, de la zona asignada a la Unidad de Conservación y Desarrollo Forestal 4, en el municipio de San Dimas, Estado de Durango,

El bosque es de tipo coetáneo e in-coetáneo presentando una distribución en su dimensión de diámetro entre 5 y 80 centímetros principalmente.

El área basal promedio fluctúa entre 12 y 23 m² /ha. En ella se establecieron entre 1966 y 1968, 36 sitios permanentes de investigación silvícola en los que se han hecho mediciones dandométricas de las que se utilizaron las relativas a los años de 1982, 1986. Para el cálculo de los parámetros para estimar el crecimiento del bosque, de cada árbol se tomó medición de las siguientes variables: especie, tipo, diámetro normal, diámetro del tocón, grosor de la corteza, condición de daño, altura total, altura de fuste limpio, clase o dominancia, piso, vitalidad, tendencia dinámica y proyección de copa.

2.1.2. Selección de la muestra de validación. La selección de la muestra que se escogió buscó reflejar comportamiento del sistema, para lo cual en este trabajo la selección se realizó utilizando conocimiento a priori del sitio experimental de estudio y observando la frecuencia de aparición del crecimiento con respecto a las dimensiones del individuo. Así se consideró la influencia de la riqueza de nutrientes y agua del sitio su competencia, grupo de edad y la frecuencia de velocidad de crecimiento.

2.2. METODOLOGÍA.

2.2.1. Red Neuronal Backpropagation. La red neuronal backpropagation (Red Neuronal Feedforward FANN) corresponde a la clase de ANN más estudiada por el ámbito científico y la más utilizada en los diversos campos de aplicación.

Compuesta por tres elementos esenciales:

1. Las neuronas o nodos: son los elementos básicos de procesamiento
2. La arquitectura de la red: descrita por las conexiones ponderadas entre los nodos.
3. El algoritmo de entrenamiento: usado para encontrar los parámetros de la red.

Donde la red neuronal backpropagation es un sistema de procesamiento con la capacidad de examinar todos los patrones en paralelo. Con la capacidad de adaptarse a sí mismo para aprender la relación entre un conjunto de patrones dado como ejemplo y ser capaz de aplicar la misma relación a nuevos patrones de entrada. Poseyendo la capacidad de concentrarse en las características de una entrada arbitraria que se asemeje a otros patrones vistos previamente, pudiendo tratar con señales de ruido.

La principal ventaja de estas redes neuronales es que pueden comportarse como aproximadores universales, lo cual se demuestra por el teorema de Kolmogorov (1957).

2.2.2. Sistemas neuro-difusos (ANFIS). El avance de los sistemas de inteligencia computacional y la búsqueda de soluciones cada vez más completas para los problemas que pueden ser resueltos por estos métodos, han llevado a la unión de técnicas de soft computing (Li, 1996) haciendo que las debilidades de unos sistemas se compensen con las bondades de otros. Esta última, resulta ser una fuerte motivación para realizar la unión de las técnicas de Redes Neuronales Artificiales y Lógica Difusa.

ANFIS. En esta aproximación, los sistemas difusos y de RNA, trabajan juntos en una arquitectura homogénea que puede ser interpretada como una RNA con parámetros difusos o como un sistema difuso con parámetros o funcionamiento distribuidos.

Una red neuro difusa presenta como ventajas, su estructura permite el uso de métodos cualitativos y cuantitativos en la construcción de modelos, en la etapa de aprendizaje, en la de funcionamiento y en la realimentación de conocimiento al usuario, permitiendo así, todo el conocimiento disponible. Los métodos cuantitativos se observan en el aprendizaje que los sistemas neuro-difusos pueden realizar a partir de datos del ambiente y en su forma matemática de funcionamiento, mientras los métodos cualitativos se reflejan en el sentido lingüístico que tiene la base de conocimiento de los sistemas difusos y que genera una interacción bilateral sistema-hombre, permitiendo la introducción de conocimiento previo al sistema y la realimentación del estado final del modelo en forma de reglas expresadas lingüísticamente desde el sistema al hombre. Las otras ventajas son las ya conocidas ampliamente y que vienen heredadas de los sistemas difusos y de las RNA, tales como interpretación lingüística, introducción de conocimiento previo, autoaprendizaje, generalización (Haykin, 1994), interpolación, etc. Su principal desventaja El curso de la dimensionalidad: Los sistemas difusos presentan el problema del crecimiento exponencial del tamaño del modelo respecto al número de variables que poseen.

2.2.3. Parámetros para entrenamiento de la red neuronal. En este trabajo se estudia el desempeño de las redes neuronales entrenadas con índices de competencia independientes de la distancia en la predicción del crecimiento en diámetro y altura de un bosque irregular de pino de la especie *Pinus Cooperi* Blanco.

El crecimiento de un bosque de coníferas integrado por individuos de diferente edad y tamaño, suele ser caracterizado por índices de competencia independientes de la distancia, los cuales son construidos en base a las dimensiones del individuo tales como altura, diámetro cuadrático, área basal, amplitud de copa, longitud copa y son relacionadas a las respectivas dimensiones de la comunidad donde se encuentran.

En este trabajo son utilizados para entrenar la red neuronal una serie de índices independientes de la distancia, propuestos por Valles (2005) y un índice propuesto por Glover y Hool (1979) para predecir el crecimiento en diámetro, y altura.

2.2.4. Procedimiento de solución. Para determinar las características de la red utilizada en este trabajo, se muestra como se fueron seleccionando el pre-procesamiento de los patrones de entrenamiento, la arquitectura de la red, el tipo de función de activación de las capas ocultas, el algoritmo de optimización de los pesos de la red.

2.3. RESULTADOS.

Se muestran los resultados obtenidos al entrenar diferentes arquitecturas de red backpropagation en su capacidad de generalizar el crecimiento de un bosque de *Pinus Cooperi* Blanco y se muestran las redes que logran generalizar el crecimiento en altura y diámetro. Se presentan el funcionamiento de la red neuro-difusa ANFIS en la predicción generalizar el crecimiento en altura y diámetro de un bosque de *Pinus Cooperi* Blanco utilizando diversos índices de competencia independientes de la distancia.

CAPITULO III.

3. ANÁLISIS DE RESULTADOS, CONCLUSIONES, RECOMENDACIONES Y TRABAJO A FUTURO.

3.1. ANALISIS DE RESULTADOS.

Se analiza el comportamiento de la red neuronal backpropagation en lo que respecta a su manejo de la información respecto a su capacidad de generalización y en el proceso de búsqueda de la de la red neuronal con la mejor capacidad de generalización como las otras redes de mayor o menor tamaño se comportaron al tratar de predecir el crecimiento de un bosque de *Pinus Cooperi* Blanco.

Se comentan los resultados obtenidos de la utilización de una red neuro difusa ANFIS y el porque de sus resultados negativos al tratar de predecir el crecimiento de un bosque de *Pinus Cooperi* Blanco.

3.2. CONCLUSIONES.

Se presentan las características de la red neuronal backpropagation capaz de generalizar de forma correcta el crecimiento de un bosque de *Pinus Cooperi* Blanco.

Se concluye en base a los resultados obtenidos ratificando la capacidad una red neuronal backpropagation en el manejo de sistemas biológicos complejos.

Se muestra la ventaja de utilizar una red neuronal backpropagation en lugar de una neuro difusa ANFIS cuando se hace uso de una gran cantidad de parámetros.

3.3. RECOMENDACIONES.

Se comenta la posibilidad de poder modelar otros sistemas biológicos en base a los resultados obtenidos.

Se comenta sobre lo adecuado del procedimiento de selección de la información para predecir la evolución de un sistema que presenta periodos de tiempo de desarrollo muy grandes.

3.4. TRABAJO A FUTURO.

Como punto más importante se presenta como continuación del presente trabajo el utilizar la red neuro-difusa ANFIS haciendo uso de una mayor capacidad de memoria que permita observarla con mayor posibilidad de manejo de la información.

BIBLIOGRAFÍA.

ANEXO.

ESTIMACIÓN DEL CRECIMIENTO DE UN BOSQUE DE PINO (*PINUS COOPERÍ* BLANCO) POR MEDIO DE UNA RED NEURONAL BACKPROPAGATION ENTRENADA CON ÍNDICES DE COMPETENCIA INDEPENDIENTES DE LA DISTANCIA.

RESUMEN

El desarrollo de las técnicas de manejo de los bosques es similar en muchos lugares del mundo. Sin embargo; pocas teorías silvícolas pueden ser generalizadas para una región forestal específica, que permitan describir los procesos en el desarrollo futuro de los ecosistemas forestales; debido a que las condiciones climáticas, suelos y especies son muy diferentes. En estos casos se necesita conocer la dinámica de crecimiento en cada localidad, con base en las interacciones suelo-clima-vegetación y así conocer sus efectos en el desarrollo y crecimiento de los bosques naturales.

El tipo de bosques que caracterizan a la mayoría de los bosques en México, así como a la región boscosa del Estado de Durango es del tipo llamado irregular los cuales presentan una diversidad en edad y en estructura y/o tamaño, o bien en su distribución espacial en el sitio.

En estos bosques complejos, el conocer la producción presente y predecir el rendimiento futuro es difícil debido a la heterogeneidad, por lo cual se requiere más información silvícola y modelos basados en metodologías que puedan abordar de forma más adecuada la complejidad de un sistema.

En este trabajo se estudia el desempeño de las redes neuronales entrenadas con índices de competencia independientes de la distancia en la predicción del crecimiento en diámetro y altura de un bosque irregular de pino de la especie *Pinus Cooperi* Blanco.

El crecimiento de un bosque de coníferas integrado por individuos de diferente edad y tamaño, suele ser caracterizado por índices de competencia independientes de la distancia, los cuales son construidos en base a las dimensiones del individuo tales como altura, diámetro cuadrático, área basal, amplitud de copa, longitud copa y son relacionadas a las respectivas dimensiones de la comunidad donde se encuentran.

Los métodos de análisis estadístico pueden ser utilizados para entender un sistema biológico de nuestro interés el cual es demasiado complejo para ser modelado analíticamente.

La información del comportamiento del sistema es recolectada de la medición de variables seleccionadas. Esas observaciones son entonces analizadas aplicando métodos estadísticos o infiriendo conocimiento del sistema. Los resultados del análisis dependerán del conocimiento a priori acerca del sistema y de los métodos empleados durante el análisis. Esencialmente, todas las opciones consideradas durante el análisis, empezando desde la recolección de la información, y terminando con el análisis de resultados obtenidos de un modelo estadístico tienen efecto en la inferencia de conocimiento.

Las redes neuronales constituyen una clase de modelos estadísticos, los cuales pueden ser utilizados en una amplia gama de aplicaciones, desde análisis exploratorio

de datos y visualización hasta problemas de clasificación y regresión. En el modelado con redes neuronales, la meta es construir un modelo basado en datos de observaciones de un sistema, los cuales representen las propiedades esenciales del sistema de nuestro interés.

La red neuronal backpropagation entrenada con índices competencia independientes de la distancia pronosticó la velocidad de crecimiento del diámetro y altura del *Pinus Cooperi* Blanco con 90% y 94,5% de aproximación en un intervalo de confianza del

GROWTH EVALUATION OF A CONIFER FOREST (*PINUS COOPERI* BLANCO) USING A NEURAL NET BACKPROPAGATION TRAINED WITH DISTANCE INDEPENDENT COMPETITION MEASURES.

ABSTRACT

Development of handling practices to pine forests is similar around the world. Therefore, a few forest theories could be generalized to the same tree location and allow describing a future develop of the forest ecosystems; because meteorology conditions, soils quality and forest weed species could be very different. In these cases are necessary to recognize dynamics of growth in every location and their effects product of interrelations soil-climate-vegetation in forest location growing.

The characteristic forest in Mexico country and mountain region at Durango State is irregular forest. This kind of forest is composed of individual with very different size and aged or/and very different spatial distribution of the location.

In this very complex forest, to know production wood at the present and future time is very difficult to cause of heterogeneity of the individuals and location. Therefore is necessary more information about forest locations and approach to modeling the dynamics of complex systems.

In this work research neural nets performance training with distance independent competition measures forecasting diameter, height growth at a conifer specie *Pinus Cooperi* Blanco irregular forest location.

We model growth conifer trees with distance independent competition measures. These measures treat to explain variation in the height and diameter growth of individual conifer trees. This competition is based on the ratios of competing individual dimensions: diameter, height with the neighborhood dimensions.

Statistical data analysis methods can be applied to gain understanding to a process under interest, which is too complex to be modeled analytically. Data from the process is first collected by measuring selected variables. These observations are then analyzed by applying statistical methods to infer knowledge from the process.

The results of the analysis depend on the prior knowledge about the process and the methods employed during the analysis. Essentially, all the choices made during the analysis, starting from the data collection, and ending to analyzing the results gained from a statistical model have an effect on the inferred knowledge.

Neural networks constitute a class of statistical models, which can be applied in a broad range of applications, from exploratory data analysis and visualizations to classification and regression problems. In neural network modeling, the goal is to build a model based on the observed data, which can represent the essential properties of the process under interest.

A neural net backpropagation was training with distance independent competition measures and forecasting a diameter and height growth with an approximation of 90% and 94.5% with 90% confidence interval.

CAPITULO I

1. ANTECEDENTES, JUSTIFICACIÓN, MOTIVACIÓN, HIPÓTESIS, Y OBJETIVO.

1.1 ANTECEDENTES.

1.1.1. Definición del Problema. Un *bosque* puede definirse como un *ecosistema* (Bormann, 1979, Whittaker, 1975,) caracterizado por una extensa cubierta arbórea de mayor o menor densidad, donde los árboles son sus principales componentes y se interrelacionan con los demás organismos (hongos, insectos, fauna, etc) y el medio físico (Ovington, 1957, 1958, 1961).

Entre los diferentes tipos de bosques se pueden considerar:

a) **Bosques nativos:** Son aquellos que se desarrollan naturalmente en un área, y están integrados por especies adaptadas a la zona, desde hace muchas generaciones.

b) **Bosques implantados:** Son los que instaló el hombre y en general se trata de plantaciones de especies exóticas. En ellas los individuos se plantan en hileras paralelas y guardando entre sí distancias fijas, a esto se le llama plantación en macizo; también se plantan en forma de hileras individuales, a lo que se le llama cortina cortaviento. Suelen instalarse en áreas que carecían de bosque, pero a veces se llevan a cabo en bosques talados, propiciando el reemplazo de la especie nativa.

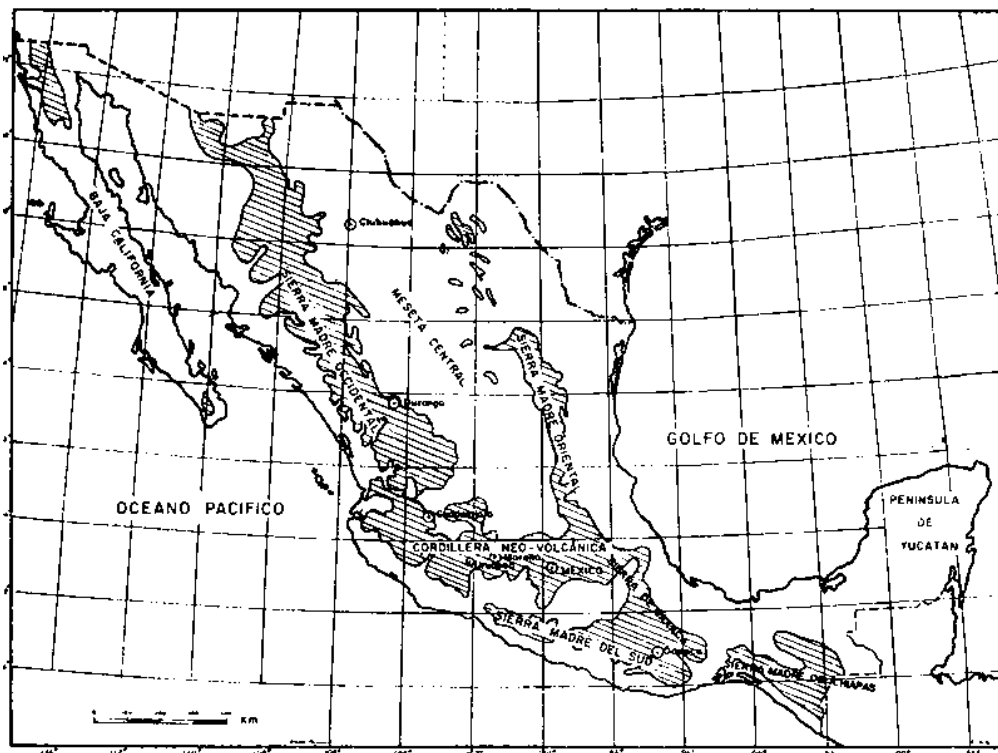


Figura 2 Distribución de la región boscosa en México

1.1.1.1 Crecimiento de los árboles: El crecimiento en altura está determinado por yemas apicales (brote largo), o sea que cuanto más rápido crecen éstas, más rápido crece el árbol en altura. En las zonas de clima frío, el crecimiento está caracterizado por

períodos de inactividad en invierno y de actividad en el resto del año. El reposo ocurre por una combinación de temperaturas bajas y escasa cantidad de hora de luz (fotoperíodo). La estacionalidad determina la formación de los anillos de crecimiento del tronco claramente observables cuando se realiza un corte transversal del mismo. El término leño temprano se atribuye al menos denso (el correspondiente al período de mayor crecimiento, como en primavera) y el leño tardío es el leño más denso (generalmente más oscuro, correspondiente a otoño).

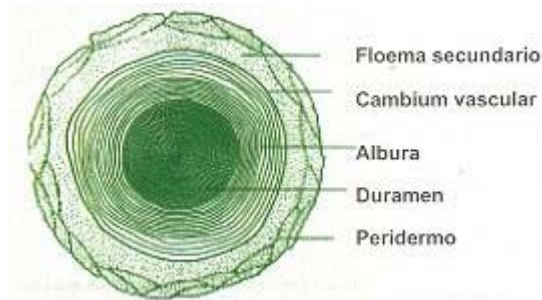


Figura 3. Partes que forman el tronco de un árbol

Los anillos de la madera se derivan del crecimiento en ancho del tronco debido a la actividad del *cambium* (*meristema* que genera tejidos). El *cambium* da lugar al crecimiento leñoso del xilema (hacia el interior) y del floema (hacia afuera). El floema activo es una delgada capa de células, aquella que más recientemente se ha diferenciado del *cambium*, y cumple funciones de transporte de asimilados orgánicos provenientes de la actividad fotosintética (savia elaborada). El *xilema* (leño) se divide en ALBURA y el leño central llamando DURAMEN. La albura es fisiológicamente activa en la acumulación de reservas y en el movimiento de agua y minerales desde las raíces hasta la copa (savia bruta). El duramen está formado por células muertas y sus principales componentes químicos de la madera son la celulosa y la lignina. (Figura 3)

El crecimiento biológico, es el resultado de un gran número de procesos complejos (Ovington, 1960, Watson, 1958). Cuando el organismo o población analizada es un árbol o masa forestal, podemos observar que los árboles poseen diferentes tasas o velocidad de crecimiento en las distintas etapas de su vida. En el crecimiento acumulado, tanto en altura como en diámetro, se distinguen tres etapas: un período inicial de crecimiento juvenil, que suele ser lento; un período en que el crecimiento se acelera hasta alcanzar un óptimo; y una etapa de madurez donde el crecimiento decrece hasta que alcanza la máxima dimensión.

En el caso de los árboles, su crecimiento es el resultado de la interacción entre dos factores: por una parte está la tendencia intrínseca hacia un crecimiento ilimitado, que depende del potencial biótico del individuo, de su actividad fotosintética, de la absorción de nutrientes, de los procesos catabólicos y anabólicos, etc.; en el lado opuesto se encuentran las restricciones al crecimiento, impuestas por el entorno en el que se desenvuelve dicho individuo (competencia con otros organismos, limitación de

recursos, procesos respiratorios y estrés) y por su propia condición de ser vivo (mecanismos de autorregulación del crecimiento y envejecimiento).

Si se representa la evolución en el tamaño (diámetro, altura, área basimétrica volumen, etc.) de un árbol o de una masa forestal a lo largo de su vida se obtiene una curva sigmoidea, cuyo patrón se mantiene sorprendentemente estable, con independencia de la especie de que se trate o de los cambios y fluctuaciones del entorno en que se desarrolle. La curva sigmoidea presenta dos características fundamentales, que derivan de la interacción entre los dos factores opuestos que condicionan el crecimiento, comentados anteriormente: tiene un punto de inflexión y una asíntota horizontal.

En la Figura 4 se ha representado una curva sigmoidea, que describe la evolución con el tiempo del tamaño de cualquier variable de un árbol o una masa forestal (diámetro, altura, área basimétrica, volumen, etc.), junto con la curva de su primera derivada, que corresponde al crecimiento corriente de dicha variable (representada a una escala mayor para facilitar su comparación).

En dicha figura se puede apreciar cómo inicialmente, en las primeras edades, el peso del factor positivo es mucho mayor que el de las restricciones o factor negativo, por lo que el crecimiento se dispara, dando lugar a una curva de evolución cóncava hacia arriba. Así en el caso de plantaciones en las primeras etapas de crecimiento habrá poca o ninguna competencia entre las plantas por la luz, nutrientes del suelo o agua. Bajo estas condiciones, las curvas de crecimiento tienen, generalmente, una forma exponencial, indicando que la ganancia anual de peso en seco aumenta continuamente con el aumento de la superficie foliar, y por tanto, de la capacidad de fotosíntesis. Así pues, la plantación crece a un ritmo siempre en aumento durante esta fase «exponencial». Siguiendo el curso normal, sin embargo, el vuelo se cierra y la competencia por la luz se hace severa y, según va formándose nuevo follaje en las partes superiores del árbol, las ramas inferiores muy asombradas se mueren. En las primeras etapas de la plantación el índice de superficie foliar será menor que 1, pero al proseguir el crecimiento, aumentará constantemente hasta que, al cerrar el vuelo, alcanza un valor constante, cuando la producción de nuevas hojas se compensa con la muerte del follaje de la región inferior del árbol.

A medida que se desarrolla el individuo las restricciones van aumentando, de tal manera que se llega a producir un cambio en la curvatura de la gráfica de evolución (se produce un punto de inflexión), que se corresponde con el valor máximo de la curva de crecimiento corriente. A partir de ese momento, y puesto que siguen aumentando las restricciones, el crecimiento comienza a ralentizarse, lo que significa que la curva de crecimiento disminuye paulatinamente y la de evolución aumenta cada vez más despacio. Esta tendencia continua hasta que, finalmente, los factores que favorecen y se oponen al crecimiento se igualan, por lo que cesa el crecimiento y se estabiliza el tamaño del individuo, hecho que corresponde matemáticamente con la asíntota horizontal.

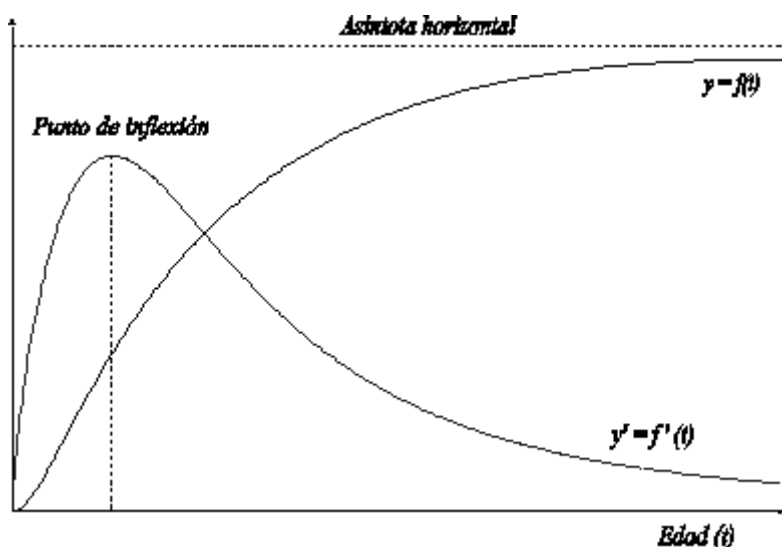


Figura 4 Curva de crecimiento arbóreo.

La altura de los árboles está generalmente relacionada con el tamaño y vigor de sus copas, lo que genera diferentes ritmos de crecimiento y condiciones de competencia.

Los crecimientos en diámetro dependen de las fases de desarrollo, la densidad, la clase de copa y la calidad del sitio.

El crecimiento en altura de los árboles depende fundamentalmente de la calidad de sitio (suelo, exposición, pendiente, temperatura, etc.).

El incremento alcanzado en un año se denomina crecimiento anual corriente.

En los bosques nativos se utiliza una clasificación de cuatro clases de copa: árboles dominantes que son aquellos cuyas copas alcanzan la altura máxima sobrepasando ligeramente el dosel general del bosque, y con un buen desarrollo lo que les permite una buena iluminación hasta cierta profundidad; los codominantes que se encuentran un poco más bajo y con menos volumen de copa que los dominantes, conforman partes del dosel superior y sus copas reciben luz sólo en la parte superior; árboles intermedios que se encuentran oprimidos por individuos dominantes y codominantes y alturas menores; y suprimidos ubicados totalmente bajo el dosel, de modo que no reciben luz solar directa, sino que sólo algo de luz difusa y por esta razón son débiles, de bajo crecimiento y tienden a morir.

Entre los factores que influyen en el crecimiento de la masa forestal se encuentran.

a) Calidad de sitio. A través de la altura de los árboles dominantes (árboles adultos) se puede determinar la calidad del sitio forestal, siendo éste la combinación de factores de suelo, pendiente, exposición, microclima, etc. La calidad de sitio indica la productividad potencial de una zona. Cuando mayor es la calidad de sitio de un lugar para una especie, más valioso desde el punto de vista económico es para el hombre ese terreno.

Si se observan árboles de 24 metros de altura, éste corresponde a un buen sitio forestal, es decir; suelos profundos con acumulación de materia orgánica, bien drenados, y un lugar protegido de los vientos. En una lomada expuesta a los vientos,

con suelos poco desarrollados, la altura de los árboles dominantes, sólo alcanza 3 metros de altura, esto nos indica un sitio muy poco productivo.

Como conclusión, un árbol con la misma información genética, en un sitio de buena calidad podría alcanzar 26 metros de altura y en un mal sitio forestal quedar con forma de arbusto de 1 metro de altura

El tipo de suelo figura 5 es un factor importante en la determinación de las clases de plantas que crecen en un área determinada. Las plantas usan elementos inorgánicos del suelo, tales como nitrógeno, potasio y fósforo, pero la comunidad de hongos, bacterias y otros organismos microscópicos que viven en el suelo también son vitales. Estos organismos vivos ayudan en la descomposición de plantas y animales muertos, incorporándolos así al suelo.

El suelo es afectado por el clima y las lluvias, la geología y la vegetación. La combinación de arena, limo, grava y arcilla le da texturas diferentes a suelos diferentes. Los suelos saludables, ricos en nutrientes, están constituidos por una mezcla de arena, limo y arcilla y son llamados suelos "francos". Los colores de los suelos se deben a diferentes minerales.



Figura 5 Suelo de un bosque

Las hojas caídas, y de hecho los árboles caídos completos, son el suelo de los bosques futuros. Esta materia orgánica es crucial debido a que contiene los nutrientes que eventualmente serán reincorporados en el suelo. También es importante en un estado parcialmente descompuesto ya que las hojas y maderas que se están pudriendo son capaces de almacenar humedad, como esponjas, y ayudan a que el suelo forestal retenga el agua de lluvia. Sin la materia orgánica de árboles y otras plantas del bosque, el suelo sería apenas rocas y arena.

Comparado a otros tipos de suelo, los arenosos son relativamente "jóvenes". El agua y el aire se mueven fácilmente entre los granos grandes de arena, por lo que este tipo de suelo retiene muy poca agua y es susceptible de secarse. Lo contrario sucede en los suelos arcillosos los cuales frecuentemente se presentan inundados debido a que

las pequeñísimas partículas de arcilla están fuertemente empaquetadas, lo que hace difícil el movimiento del agua y del aire. Los suelos arcillosos pueden ser lo suficientemente densas para dificultar el crecimiento en ellos de las raíces.

Por medio de sus raíces, los árboles extraen agua desde el subsuelo. Un árbol grande puede liberar hacia el aire hasta más de 100 galones diarios de agua. Este vapor de agua se convierte en parte del microclima del área, el cual sería mucho más seco sin la humedad proveniente de los árboles.

Demasiada precipitación puede producir el "lavado" del suelo. Esto sucede cuando el agua arrastra la mayor parte de los minerales y otros nutrientes. El agua también puede contribuir a la acidez del suelo debido a que, cuando se mueve por el suelo, retira minerales y deja átomos de hidrógeno. Mientras más átomos de hidrógeno haya en el suelo, más ácido es el suelo. Sin embargo, esto solamente llega a ser un problema en lugares donde ocurren frecuentes lluvias fuertes y, en estos casos, las plantas se han adaptado a crecer en suelos pobres. El lavado de los suelos en estos lugares es natural. En la mayoría de todos los demás casos, el movimiento del agua por el suelo es muy beneficioso para los bosques.

b) Importancia del agua en el desarrollo de un bosque. El agua dulce (no salada) es esencial para toda la vida sobre la Tierra, y los bosques son esenciales para el agua dulce. Los bosques filtran y limpian el agua, amortiguan las lluvias fuertes que de otra manera erosionarían los suelos, y mantienen en su lugar los bancos de los ríos. A su vez, el agua transporta nutrientes disueltos y los distribuye por todo el suelo del bosque. Los bosques actúan como "esponjas", capaces de recoger y almacenar grandes cantidades del agua de lluvia. Los suelos forestales absorben cuatro veces más agua de lluvia que los suelos cubiertos por pastos, y 18 veces más que el suelo desnudo.

Con sus profundos sistemas de raíces, los árboles son capaces de extraer agua de zonas profundas del suelo. El agua se mueve por el árbol y se usa en la fotosíntesis, en el enfriamiento, y en otros procesos de crecimiento. Se evapora, como vapor de agua, desde las hojas. En este ciclo, los árboles son "fuentes de agua" vivientes que redistribuyen el líquido: la humedad, que se quedaría atrapada en forma subterránea si no fuera por los árboles, es liberada a través de sus hojas hacia el aire, donde luego se condensa formando nubes y cae de nuevo en forma de lluvia. Sin árboles que distribuyan esta agua, el clima en muchas regiones sería mucho más seco. Esta reserva subterránea y constante de agua es liberada lenta y gradualmente por los árboles, ayudando a evitar las inundaciones y sequías estacionales.

El dosel del bosque intercepta una gran parte de la precipitación que cae sobre él. De no ser así, esta lluvia y nieve caería directamente sobre el suelo, erosionándolo gradualmente. El bosque evita que suceda esta erosión. El agua que cae en el dosel del bosque luego cae suavemente al suelo con mucha menor fuerza que la lluvia directa.

c) Exposición a la luz solar. La fuente de energía para el crecimiento de las plantas la brinda el sol, junto al agua y el dióxido de carbono, la energía solar es convertida por la fotosíntesis en las fibras de madera de los árboles en desarrollo. Su radiación

afecta y determina los patrones de clima en la tierra y su energía lumínica es convertida por las plantas verdes en energía química mediante el proceso de la fotosíntesis.

El ritmo de crecimiento del árbol dependerá de la velocidad con que aumente la superficie foliar, ya que la última determina su capacidad fotosintética final. De ahí que una especie que desarrolla rápidamente una gran superficie foliar tendrá un ritmo de crecimiento grande.

Las plantaciones más eficientes son aquéllas que interceptan completamente la luz incidente, y la emplean eficazmente en la fotosíntesis.

Ciertos árboles de copa frondosa muestran en las primeras etapas de crecimiento incrementos mayores que las coníferas, pero esta situación se invierte después cuando cierra el vuelo, pues se admite que la productividad del bosque de coníferas es considerablemente mayor que el de frondosas (Ovington, 1957, 1960). Esta productividad mayor de las coníferas en masas de espesura normal, a pasar del mayor potencial en ritmo de crecimiento de ciertas frondosas como individuos, es probablemente debido en parte a que la cubierta de coníferas intercepta en mayor proporción la luz incidente que la cubierta de muchas frondosas (Ovington 1958, 1961) ha señalado también que la forma cónica de la copa de cada árbol en las coníferas asegura una mayor superficie de hoja bien iluminada. Así pues, se ha atribuido la mayor productividad de los bosques de coníferas a poseer hojas persistentes, a la densidad de la cubierta y a la forma cónica de la copa.

Otro factor importante que afecta a la producción anual de materia seca por el árbol es la duración de la superficie foliar, que es la integral del índice de superficie foliar a lo largo del período de crecimiento, la cual toma en consideración la magnitud y la persistencia del índice de superficie foliar (Watson, 1956, 1958). Claro está que las coníferas de hoja persistente son capaces de realizar fotosíntesis siempre que las condiciones de luz, temperatura y agua sean favorables. Aunque existen diferencias de opinión sobre si la fotosíntesis de las coníferas en invierno tiene alguna importancia, parece no existir duda de que en regiones con inviernos suaves se da un incremento significativo de la materia seca en este período (Hagem, 1962).

Probablemente es más importante, sin embargo, la disposición de las coníferas de hoja persistente (figura 6) para hacer la fotosíntesis en la primavera temprana o en otoño tardío, cuando las condiciones de luz y temperatura sean favorables.

Otro factor importante que afecta a la superficie de hoja y al ritmo de crecimiento de los árboles es el tipo estacional del crecimiento longitudinal y radial. Como es bien sabido, algunas especies arbóreas continúan con un crecimiento activo longitudinal y producción de hojas durante toda la estación de crecimiento, mientras que en otras especies el crecimiento longitudinal se limita a la expansión de primordios foliares encerrados en la yema durmiente y, una vez terminada esta expansión, ya no desarrollan más hojas. Intermedias entre estos dos tipos extremos, muchas especies presentan una forma continua de actividad apical y expansión de nuevas hojas después de la expansión de la yema, pero el crecimiento longitudinal cesa en pleno verano, es decir, al final de junio o en julio. Ahora bien, es evidente que el tipo de crecimiento longitudinal de los brotes en un árbol afectará marcadamente a su

producción de hoja y ritmo de crecimiento. En las especies en que el crecimiento longitudinal es mantenido todo el año, la superficie de hoja, y por tanto, la capacidad fotosintética del árbol aumenta constantemente durante toda la estación. Por el contrario, en los árboles en que el crecimiento longitudinal se limita a una breve expansión en primavera, la superficie foliar permanece constante la mayor parte de la estación y el nuevo material formado por fotosíntesis no se invierte en el desarrollo de nuevas hojas sino en el desarrollo de yemas, en el crecimiento radial y en la formación de reservas. Los asimilados empleados en el crecimiento radial no son desperdiciados, por supuesto, pero es obvio que para la máxima eficacia de producción de madera es preciso establecer un equilibrio entre el crecimiento radial, que no aumenta la capacidad fotosintética del árbol, y la inversión de asimilados en material foliar que posteriormente aumentará su capacidad productiva. El género *Pinus* ofrece un caso interesante al darse en sus árboles el crecimiento longitudinal en expansiones muy rápidas.



Figura 6. Agujas (Hojas) de pino

1.1.1.2 Un bosque como un sistema. Un sistema biológico presenta un alto grado de complejidad en la descripción de su comportamiento. Donde la complejidad en los sistemas naturales típicamente surge del esfuerzo colectivo de un gran número de componentes. Siendo a menudo prácticamente imposible predecir el comportamiento de un componente en particular de forma detallada, o el preciso comportamiento del sistema completo (Weisbuch, 1991, Law, 1990).

Un sistema como un todo puede sin embargo puede mostrar un comportamiento total definido, y este comportamiento usualmente posee características importantes.

Quizás la más importante, es su robustez, donde típicamente no es afectado por fallas o perturbaciones de sus componentes individuales.

En nuestro problema buscamos predecir el crecimiento en diámetro y altura de un pino donde este árbol en sí mismo es un sistema complejo formado de millones de células que interactúan entre ellas y son afectadas por factores externos; el árbol mismo se integra como un individuo formando parte de una colectividad llamada bosque. Los integrantes de este sistema interactúan entre ellos con otros organismos y con el medio ambiente.

Y así de esta serie de interrelaciones entre el árbol y su entorno y las propias interrelaciones entre las células que lo forman, se debe extraer un comportamiento característico que logre abstraer el crecimiento de los anillos de la madera que derivan en el crecimiento en ancho del tronco debido a la actividad del *cambium*.

Un sistema que esta sujeto a complejas influencias externas posee una dinámica que puede ser modelada estadísticamente (Lev Davidovich, 1969). El tratamiento estadístico simplifica la compleja dinámica estocástica impredecible a un sistema sencillo (N. G. van Kampen, 1981). En el tiempo, las influencias aleatorias causan en el sistema una distribución en el universo de comportamientos de forma gaussiana.

El principio de la aplicación de este tratamiento estadístico es la de simplificar la complejidad de la interacción entre los elementos de un sistema haciendo un tratamiento de los estados microscópicos que describen el comportamiento de equilibrio macroscópico del sistema (Kerson, 1987).

Así el principio general de simplicidad descansa en un promedio de las dinámicas microscópicas rápidas en la escala del tiempo de las observaciones macroscópicas (Teorema de ergodicidad) y un promedio de las variaciones microscópicas espaciales.

Así se puede considerar el crecimiento como un cambio gradual en el tiempo de un ser vivo que es apreciado a través del proceso natural fisiológico y de igual manera es representado de forma simplificada por medio de una curva de crecimiento siendo esta una representación gráfica de las etapas de desarrollo en el transcurso de su vida.

Al estudiar el crecimiento de un ser vivo, puede ser visualizado como un sistema del cual se logra su representación a través de un modelo (Bruce, 1987, Tabachnick, 1989).

1.1.2. Estado del Arte. La toma de decisiones adecuadas para el manejo de un bosque natural con una distribución espacial del terreno heterogénea integrado de una población compuesta de una mezcla de diferentes especies se incrementa en un alto grado de dificultad. El acceso a la nueva tecnología de inteligencia artificial (AI), permite un procesamiento del conocimiento que puede ser catalogado como herramientas de apoyo en toma de decisiones. La utilización de redes neuronales artificiales (ANN), conocidas por su capacidad de procesamiento paralelo, distribuido de la información, para predecir comportamientos de sistemas no lineales se presenta como una alternativa muy atractiva para resolver este tipo de problemas.

El cerebro es el órgano en el que se alojan las habilidades cognoscitivas de los seres humanos. Puede decirse que el cerebro es un elemento de procesamiento de la

información extremadamente complejo, cuyo modo de funcionamiento es eminentemente paralelo y cuyo comportamiento no puede describirse por medio de modelos sencillos como son los lineales.

El modelo conexionista fue desarrollado tratando de imitar el sistema de computación más complejo de los que se conocen hasta ahora, que es el cerebro. El cerebro esta formado por millones de células llamadas neuronas.

Desde el punto de vista computacional, una ANN puede ser descrita como un conjunto de autómatas celulares (*neuronas*), por el cual se establece un flujo de información mediante una topología de interconexiones (*sinapsis*).

De esta forma, las ANN imitan en cierto modo la estructura física y el modo de operación de un cerebro. Teniendo en cuenta que el cerebro presenta las cualidades de procesamiento paralelo, procesamiento distribuido y adaptabilidad, un sistema ANN posee también estas características.

El sistema resulta ser intrínsecamente paralelo porque esta formado por unidades elementales de procesamiento llamadas neuronas. Cada neurona realiza un tipo de procesamiento muy simple.

El sistema es distribuido. Esto quiere decir que la información no se almacena localmente en ciertas zonas concretas de la RNA sino que se halla presente por toda ella, en concreto, se almacena en la sinapsis entre las neuronas. De igual forma, la computación es también distribuida. Al calcular la respuesta de la red neuronal, intervienen todos y cada uno de los procesadores elementales, los cuales se hallan distribuidos por toda la arquitectura de la red. Además, éste carácter distribuido hace que la red presente tolerancia a fallos (si se pierde una parte de las neuronas no se pierde toda la información)

Una red neuronal presenta además un grado de adaptabilidad que se concreta en las capacidades de aprendizaje y generalización. Por aprendizaje entendemos la capacidad para recoger información de las experiencias y utilizarlas para actuar ante situaciones futuras. Íntimamente relacionada con el aprendizaje esta la generalización, que podría definirse como la capacidad para abstraer la información útil, más allá de los casos particulares. De esta manera, la ANN es capaz de responder ante casos desconocidos.

Así una posible definición (Haykin, 1994) de Red Neuronal Artificial es:

Una red neuronal es un procesador masivamente paralelo distribuido que es propenso por naturaleza a almacenar conocimiento experimental y hacerlo disponible para su uso.

Este mecanismo se parece al cerebro en dos aspectos:

1. El conocimiento es adquirido por la red a través de un proceso que se denomina **aprendizaje**.
2. El conocimiento se almacena mediante la modificación de la fuerza o **peso sináptico** de las distintas uniones entre neuronas.

En el estado actual de la fisiología del sistema nervioso los conocimientos acerca del funcionamiento del cerebro todavía son muy limitados. Incluso atendiendo al nivel celular los modelos disponibles son muy simples y no justifican completamente su funcionamiento. En relación con las redes neuronales artificiales, ciertos tipos de redes hoy en día parecen poco adecuadas como candidatos a modelar el funcionamiento del sistema nervioso de un ser vivo, aunque compartan a grandes rasgos las propiedades que se han enumerado antes.

Una red neuronal está constituida por nodos, o unidades, que se encuentran unidas mediante conexiones. A cada conexión se le asigna un peso numérico. Los pesos constituyen el principal recurso de memoria de largo plazo en las redes neuronales, y el aprendizaje usualmente se realiza con la actualización de tales pesos. Algunas de las unidades están conectadas con el ambiente externo, y son designadas ya sea como unidades de entrada o unidades de salida. Los pesos son modificados de manera tal que la conducta de entrada / salida de la red esté más acorde con la del ambiente que produce las entradas.

Las unidades constan de un conjunto de conexiones de entrada provenientes de otras unidades, un conjunto de vínculos de salida que van hacia otras unidades, un nivel de activación del momento y recursos para calcular cuál será el nivel de activación del siguiente paso, con base en sus entradas y pesos respectivos. Lo importante es que en cada una de las unidades se efectúa un cálculo local con base en las entradas que le proporcionan sus vecinas, pero sin que sea necesario un control global en todo el conjunto de unidades.

La información no se almacena localmente en ciertas zonas concretas de la red neuronal sino que se halla presente por toda ella, en concreto, se almacena en la sinapsis entre las neuronas. De igual forma, la computación es también distribuida. Al calcular la respuesta de la red neuronal, intervienen todos y cada uno de los nodos. Además, este carácter distribuido hace que la red presente tolerancia a fallos (si se pierde una parte de las neuronas no se pierde toda la información)

Una red neuronal presenta además un grado de adaptabilidad que se concreta en las capacidades de aprendizaje y generalización. Por aprendizaje entendemos la capacidad para recoger información de las experiencias y utilizarlas para actuar ante situaciones futuras. Íntimamente relacionada con el aprendizaje esta la generalización, que podría definirse como la capacidad para abstraer la información útil, más allá de los casos particulares. De esta manera, la red neuronal es capaz de responder ante casos desconocidos.

Una red neuronal se puede visualizar como un grafo dirigido y ponderado donde cada uno de los nodos son neuronas artificiales y los arcos que unen los nodos son las conexiones sinápticas. Al ser dirigido, los arcos son unidireccionales. Esto significa que la información se propaga en un único sentido, desde una neurona pre-sináptica (neurona origen) a una neurona post-sináptica (neurona destino). Por otra parte es ponderado, lo que significa que las conexiones tienen asociado un número real, un peso, que indica la importancia de esa conexión con respecto al resto de las conexiones. Si dicho peso es positivo la conexión se dice que es excitadora, mientras que si es negativa se dice que es inhibidora.

La unidad básica de la red neuronal artificial es la neurona. Esta se puede representar de acuerdo a la figura 7.

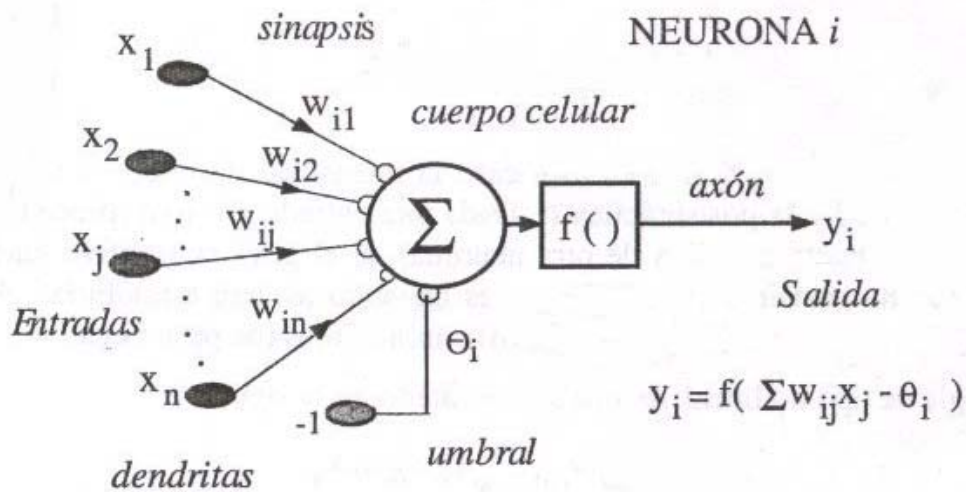


Figura 7: Neurona Artificial tipo McCulloch-Pitts

Una neurona artificial es un procesador elemental, en el sentido de que procesa un vector $x (x_1, x_2, \dots, x_N)$ de entradas y produce una respuesta o salida única. Los elementos clave de una neurona artificial son los siguientes:

- a) Las entradas que reciben los datos de otras neuronas. En una neurona biológica corresponderían a las dendritas.
- b) Los pesos sinápticos W_{ij} . Al igual que en una neurona biológica se establecen sinapsis entre las dendritas de una neurona y el axón de otra, en una neurona artificial a las entradas que vienen de otras neuronas se les asigna un peso, un factor de importancia. Este peso, que es un número, se modifica durante el entrenamiento de la red neuronal, y es aquí por tanto donde se almacena la información que hará que la red sirva para un propósito u otro.
- c) Una regla de propagación. Con esas entradas y los pesos sinápticos, se suele hacer algún tipo de operación para obtener el valor del potencial post sináptico (valor que es función de las entradas y los pesos y que es el que se utiliza en último término para realizar el procesamiento). Una de las operaciones más comunes es sumar las entradas, pero teniendo en cuenta la importancia de cada una (el peso sináptico asociado a cada entrada). Es lo que se llama suma ponderada, aunque otras operaciones también son posibles.
- d) Una función de activación. El valor obtenido con la regla de propagación, se filtra a través de una función conocida como función de activación y es la que nos da la salida de la neurona. Según para lo que se desee entrenar la red neuronal, se suele escoger una función de activación u otra en ciertas neuronas de la red.

Regularmente las neuronas se agrupan en capas de manera que una red neuronal está formada por varias capas. Aunque todas las capas son conjuntos de neuronas, según la función que desempeñan, suelen recibir un nombre específico.

Las más comunes son las siguientes:

- *Capa de entrada:* estas neuronas reciben los datos que se proporcionan a la Red Neuronal para que los procese.

- *Capas ocultas*: estas capas introducen grados de libertad adicionales en la Red Neuronal. El número de ellas puede depender del tipo de red que estemos considerando. Este tipo de capas realiza gran parte del procesamiento.
- *Capa de salida*: Esta capa proporciona la respuesta de la red neuronal. Normalmente también realiza parte del procesamiento.
-

En la clasificación de una ANN hay dos criterios que son usados regularmente, la arquitectura y el tipo de aprendizaje.

En la arquitectura se encuentran dos posibilidades distintas. Si la arquitectura de la red no presenta ciclos, es decir, no se puede trazar un camino de una neurona a sí misma, la red se llama **unidireccional** (prealimentadas). Por el contrario, si podemos trazar un camino de una neurona a sí misma la arquitectura presenta ciclos, este tipo de redes se denominan **recurrentes** o realimentados figura 8.

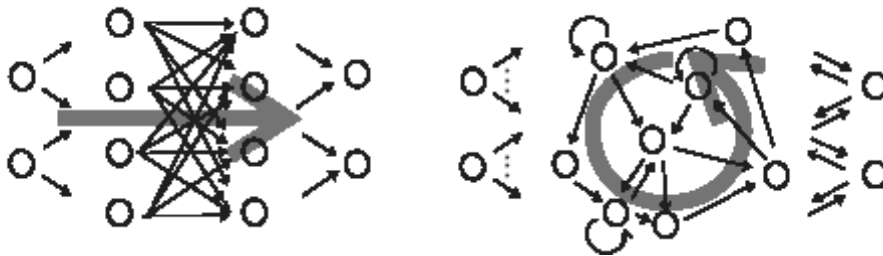


Figura 8. Clasificación de Redes.

En la clasificación de acuerdo al tipo de aprendizaje hay cuatro clases distintas:

1. *Aprendizaje supervisado*: En este tipo de aprendizaje se le proporciona a la Red Neuronal una serie de ejemplos consistentes en unos patrones de entrada, junto con la salida que debería dar la red. El proceso de entrenamiento consiste en el ajuste de los pesos para que la salida de la red sea lo más parecida posible a la salida deseada. Es por ello que en cada iteración se use alguna función que nos de cuenta del error o el grado de acierto que esta cometiendo la red.
2. *Aprendizaje no supervisado o autoorganizado*: En este tipo de aprendizaje se presenta a la red una serie de ejemplos pero no se presenta la respuesta deseada. Lo que hace la red es reconocer regularidades en el conjunto de entradas, es decir, estimar una función densidad de probabilidad $p(x)$ que describa la distribución de patrones x en el espacio de entrada R_n .
3. *Aprendizaje Híbrido*: Es una mezcla de los anteriores. Unas capas de la red tienen un aprendizaje supervisado y otras capas de la red tienen un aprendizaje de tipo no supervisado. Este tipo de entrenamiento es el que tienen redes como las RBF.
4. *Aprendizaje reforzado (reinforcement learning)*: Es un aprendizaje con características del supervisado y con características del auto-organizado. No se proporciona una salida deseada, pero si que se le indica a la red en cierta medida el error que comete, aunque es un error global.

El aprendizaje se convierte sencillamente en el procedimiento de ajuste de los parámetros para que correspondan a los datos del conjunto de capacitación, proceso conocido por los estadísticos como regresión no lineal.

Entre los diferentes métodos de aprendizaje se tienen:

a) *Corrección de error*: A la red inicial se le asignan pesos aleatoriamente. La red se actualiza en un esfuerzo porque sea congruente con los ejemplos.

Para ello se efectúan pequeños ajustes en los pesos y así disminuir la diferencia entre los valores observados y los predichos. La principal diferencia con los algoritmos lógicos es la necesidad de repetir la fase de actualización varias veces por cada uno de los ejemplos con el fin de lograr la convergencia.

Es común dividir el proceso de actualización en épocas. Cada época implica la actualización de todos los pesos para todos los ejemplos. En el caso de los perceptrones la regla de actualización de peso resulta muy sencilla. Si la salida predicha de la unidad de salida única es O, y la salida correcta debería ser T, el error se calcula de la siguiente manera: $Err=T-O$. Si el error es positivo, entonces necesitamos aumentar O; si es negativo, hay que disminuir O.

b) *Aprendizaje por propagación posterior*: El aprendizaje en este tipo de red se efectúa de la misma manera que en los perceptrones: se le presentan a la red las entradas de los ejemplos, y si ésta calcula un vector de salida que coincida con la meta, no hay nada que hacer. Pero si existe un error, los pesos se ajustan para disminuir el error.

El secreto consiste en evaluar las consecuencias de un error y dividirlos entre todos los pesos contribuyentes. En los perceptrones es fácil hacer lo anterior, puesto que sólo hay un peso entre cada entrada u salida. Pero en las redes de varios niveles hay muchos pesos que conectan cada entrada con una salida, y cada uno de estos pesos contribuye a más de una salida.

El algoritmo de propagación posterior es un método razonable para distribuir la contribución de cada uno de los pesos. Como en el caso del algoritmo de aprendizaje del perceptrón, se trata de reducir al mínimo el error entre cada salida meta y la salida realmente calculada por la red. El algoritmo de propagación posterior (back propagation) es el utilizado para entrenar una red multicapa feedforward. En este algoritmo se introduce un vector de entrenamiento a partir del cual el algoritmo ajusta los pesos de las interconexiones, al relacionar las entradas con las salidas de la red proporcionadas en el vector de entrenamiento. En el proceso de entrenamiento se produce un error entre el vector de salida de la red y el valor de vector de salida deseado, este error produce una alteración de los pesos de las interconexiones de la red. El algoritmo presenta dos tipos de información: una fase de flujo hacia delante donde la activación de cada elemento de procesamiento es pasado a la siguiente capa hasta alcanzar la salida y una fase hacia atrás donde se propaga la diferencia entre el valor actual y el deseado de la respuesta de la capa de salida, de tal forma que se van alterando los valores de los pesos de los nodos de las capas previas en dirección de la capa de salida a la capa de entrada.

El algoritmo pertenece a la categoría del tipo de entrenamiento supervisado, donde se genera un error de la diferencia entre el valor del patrón de entrenamiento y el valor producido por la red, esta diferencia es estimada vía una función de error, la cual proporciona el error global de la red. La función de error más comúnmente usada es la suma del cuadrado del error, donde las diferencias individuales de cada unidad de procesamiento son elevadas al cuadrado y sumadas. El proceso de entrenamiento se enfoca en minimizar esta función. Esta función es representada por una superficie que presenta pendientes, mesetas, mínimos locales y globales. El objetivo es evitar los quedar atrapado en los mínimos locales, y encontrar el mínimo global deteniendo la búsqueda.

El algoritmo toma en cuenta el gradiente del vector de la superficie del error, moviéndose a lo largo de dirección indicada por el gradiente, con pasos grandes o

pequeños hacia el mínimo global. Si y_j representa la respuesta de salida de la neurona j , donde

$$y_j = [1 + \exp(-u_j)]^{-1} \quad (1)$$

El algoritmo back propagation realiza una corrección Δw_{ij} al peso sináptico w_{ij} acorde a la regla delta (Haykin 1999). Para una iteración dada:

$$\Delta w_{ij} = \alpha \delta_j y_j \quad (2)$$

Donde α es el parámetro de velocidad de aprendizaje, y δ_j es el gradiente local, determinado por:

$$\delta_j = e_j \varphi_j'(u_j) \quad (3)$$

Así, la señal de error de cada neurona se vuelve muy importante en la estimación del ajuste de peso Δw_{ij} . Su cálculo depende de la topología de la neurona j . Consecuentemente:

- Si j está en la capa de salida entonces $e_j = (d_j - y_j)$ y δ_j es estimado utilizando (3).
- Si j está en una capa oculta, entonces:

$$\delta_j = \varphi_j'(u_j) \sum_k \delta_k w_{kj} \quad (4)$$

De aquí la estimación del ajuste del peso sináptico es producto de la estimación del gradiente local de la siguiente capa, de aquí el término back propagation.

Se puede considerar otro tipo de clasificación: por su modo de operación.

Modo de operación. Al hablar de modo de operación se está haciendo referencia a la manera en que la red neuronal procesa los estímulos externos y crea la respuesta de salida. Puede considerarse a una red neuronal como perteneciente a una de dos grandes categorías:

- *Redes estáticas.* En este tipo de red una vez establecido el valor de las entradas las salidas alcanzan un valor estacionario independientemente de las entradas en el instante anterior, y en un tiempo siempre por debajo de una determinada cota. Estas redes se pueden caracterizar estructuralmente por la *inexistencia de bucles de realimentación y de elementos de retardo entre los distintos elementos de proceso que las forman*. Debido a su modo de funcionamiento, estas redes tienen una capacidad limitada para sintetizar funciones dependientes del tiempo.
- *Redes dinámicas.* Este tipo de red responde de manera distinta ante diferentes secuencias de entradas, haciendo uso de manera implícita o explícita de la variable *tiempo*. Esta característica las hace en principios más idóneas que las redes estáticas para la síntesis de funciones en las que aparezca de alguna manera el parámetro tiempo

1.1.2.1. ANN en Predicción y Clasificación. A partir de 1986 (Rumelhart y McClelland, 1986; McClelland y Rumelhart, 1986) las llamadas redes neurales o modelos conexionistas han ido progresivamente utilizándose como herramientas de predicción y clasificación.

De forma breve una red neural es un sistema informático reticular (de inspiración neuronal) que aprende de la experiencia mediante la auto-modificación de sus conexiones (Hecht-Nielsen, 1990; Hertz, Krogh y Palmer, 1991; Wasserman, 1989, Hiler y Martínez, 1995; Martín y Sanz, 1997).

Como herramientas de predicción/clasificación, muchos autores están estudiando las relaciones entre las técnicas estadísticas convencionales y los modelos conexionistas (ver p.e. Cherkassky, Friedman y Wechler, 1994; Flexer, 1995; Michie, Spiegelhalter y Taylor, 1994; Ripley, 1996; Sarle, 1994, 1998). De hecho las redes neuronales han sido conceptuadas como técnicas estadísticas no paramétricas al estar libres del cumplimiento de los supuestos teóricos de la estadística paramétrica (ver p.e. Smith, 1993), o también como técnicas de regresión no lineal (ver p.e. Sarle, 1998). El problema surge cuando encontramos resultados contradictorios a la hora de determinar qué modelos son más eficientes en la solución de problemas concretos de predicción/clasificación. Así mientras algunos trabajos empíricos no encuentran diferencias entre los resultados hallados por unos y otros modelos (Croall y Mason, 1992; Michie et al, 1994; Ripley, 1993; Thrun, Mitchell y Cheng, 1991), otros resultados tienden a apoyar una ligera superioridad de las redes neuronales sobre las técnicas estadísticas (Garson, 1991; Huang y Lippman, 1987; White, 1994).

1.1.2.2. ANN en el modelado de la dinámica de crecimiento de un bosque. En años recientes, ha crecido el interés por el empleo de las redes neuronales artificiales (ANN) en varios campos de la ciencia (Maren 1990, Swingler 1996). Las redes neuronales constituyen una clase de modelos estadísticos (Hwang, 1994, Ripley, 1993, 1996, 1997), los cuales pueden ser utilizados en una amplia gama de aplicaciones, desde análisis exploratorio de datos y visualización hasta problemas de clasificación y regresión. En el modelado con redes neuronales, la meta es construir un modelo basado en datos de observaciones de un sistema, los cuales representen las propiedades esenciales del sistema de nuestro interés (Devijver, 1982).

Los modelos de crecimiento forestal que describen la dinámica de un bosque (por ejemplo: regeneración, crecimiento, mortalidad y supervivencia) han sido ampliamente utilizados en el manejo del bosque para actualizar el inventario, para predecir la producción forestal en el futuro y asegurar la composición de las especies y la estructura del ecosistema bajo la influencia de las cambiantes condiciones ambientales (Bormann, 1979). Mucho progreso ha sido hecho en el área desde la utilización de ANN para modelar la mortalidad de árboles individualmente en 1991 (Guan and Gertner, 1991a). En el mismo año, Guan y Gertner (1991b) desarrollaron un modelo basado en ANN exitosamente, que predecía la supervivencia de un árbol de pino rojo (*Pinus resinosa*). Ellos encontraron que el modelo de supervivencia del pino rojo basado en una ANN no solo ajustaba los datos mejor que un modelo estadístico, sino que tenía un mejor funcionamiento en datos a futuro. El modelo fue lo suficientemente flexible para modelar el crecimiento de pinos rojos pequeños y grandes. Su aproximación fue mejorada utilizando un algoritmo de entrenamiento adecuado y la plataforma para modelar la probabilidad de supervivencia del árbol por Guan y Gertner (1995). Por otra

parte, Hasenauer y Merkl (1997) demostraron haciendo uso de una red neuronal con aprendizaje no supervisado para predecir la mortalidad individual en modelos de crecimiento y producción en Austria. Ellos encontraron que redes neuronales trabajan ligeramente mejor que los modelos estadísticos convencionales basados en la aproximación LOGIT (Aldrich, 1984). Guan (1997) propuso un marco de referencia para asegurar localidad de la predicción del proceso basado en la mecánica de los modelos de crecimiento del bosque. El método consiste en cuatro pasos: (1) Asumir distribuciones para los valores de los parámetros, (2) examinar los parámetros, (3) esbozar el comportamiento del modelo por medio de muestreo y (4) aproximar el comportamiento del modelo en base a los puntos muestreados. Este método propuesto fue aplicado a un modelo de crecimiento basado en un balance de carbono desarrollado por Valentine (1988), y ha demostrado ser muy efectivo en el análisis de modelos grandes y complejos.

1.2. JUSTIFICACIÓN.

México es un país fuertemente importador de productos forestales y la balanza comercial del sector forestal es deficitaria en 1,850 millones de dólares (esta cantidad representa el 46 por ciento del total del déficit de la balanza comercial del país).

Los principales estados productores de madera en México son Durango en primer lugar con una aportación del 26 por ciento del total nacional para el año 2000 con 2 millones 371 metros cúbicos de madera en rollo, seguido por Chihuahua con 2 millones 091 mil metros cúbicos y Michoacán con un millón 394 mil metros cúbicos.

El estado de Durango posee el 25 por ciento de los recursos forestales nacionales en una superficie de 9.1 millones de hectáreas; siendo la entidad más importante como productora forestal en el país y ocupa el primer lugar en la explotación de la especie maderable de pino, aportando el 26 por ciento de la producción nacional con 1.9 millones de metros cúbicos de madera en rollo.

De la superficie forestal de Durango 4.9 millones de hectáreas están conformadas por bosques de acuerdo con el Inventario Forestal de la Secretaría del Medio Ambiente y Recursos Forestales (SEMARNAT).

Las principales especies explotadas son pino, con una extracción de 2 millones 033 metros cúbicos en el año 2000. Los principales municipios productores forestales son San Dimas, Pueblo Nuevo, Tepehuanes, Guanaceví, y Durango.

La necesidad de implementación de nuevas técnicas de predicción en la búsqueda de hacer un uso racional del aprovechamiento de un recurso renovable como lo es un bosque. Y para así en base a un conocimiento más preciso del crecimiento de un bosque lograr desarrollar técnicas silvícolas sustentables para ser aplicadas a una variedad de objetivos de manejo y utilización de los bosques. Logrando encaminar los bosques hacia una diversidad de regímenes para el manejo de la densidad, dependiendo del criterio, el cual puede ser productivo, económico o ecológico. Y satisfaciendo la necesidad de conocer las densidades para mejorar el crecimiento de los árboles, la calidad de la madera. También poder estandarizar un nivel de densidad que minimice el impacto sobre los recursos agua/suelo; importantes para lograr el equilibrio del ambiente forestal.

Por lo que se hace necesario el comprender el crecimiento de un bosque de forma detallada en base al uso de modelos desarrollados haciendo uso de metodologías que logren describir las interrelaciones que surgen de la complejidad de un sistema biológico como es un bosque. Esto con el fin de desarrollar técnicas de manejo del bosque que logren una mayor producción y calidad de la madera y conservación del recurso forestal.

1.3. MOTIVACIÓN.

Así al contar con una herramienta que describa de forma detallada la evolución de un bosque productor de madera se podrán alcanzar algunos objetivos como:



Figura 9. Bosque antiguo. Sierra Madre Occidental del estado de Durango

- Buscar reducir las importaciones forestales del país logrando un mejor aprovechamiento de los recursos forestales (figura 9) en base al uso de mejores herramientas de pronóstico que permitan desarrollo de técnicas silvícolas más eficientes.
- Impulsar el desarrollo de la Industria forestal del estado de Durango al proveer materias primas de alta calidad y menor precio.
- Lograr un mayor nivel de vida de la población dedicada al aprovechamiento del sector forestal al lograr acceder al mercado nacional como internacional en base a la calidad y precio de su producto.
- Satisfacer la necesidad de hacer un uso adecuado del bosque en su explotación.
- Evaluar la capacidad de producción a futuro para poder satisfacer la demanda de madera por el mercado.
- Modelar las dimensiones de crecimiento del pino en el bosque para optimizar la producción y calidad de la madera.

1.4. HIPÓTESIS.

El desarrollo de un modelo basado en la aplicación de una red neuronal artificial resultará adecuado para representar el comportamiento de un sistema complejo si se cuenta con la suficiente información que describa la dinámica de dicho sistema.

Así para seleccionar la herramienta para desarrollar el modelo que estime el crecimiento de un bosque, la decisión se ha hecho en base a los siguientes puntos.

1.- Características de la información que describe el sistema.

Se trata de desarrollar un modelo de un sistema biológico altamente complejo por la gran cantidad de interrelaciones que presenta con el medio ambiente y otros sistemas biológicos.

Por su alto grado de complejidad se presenta el problema de contar con escasez de datos, errores de medición en los datos producidos en y redundancia de datos a causa de la dificultad en su obtención. Es por lo que una red neuronal se presenta como una buena alternativa de herramienta de solución (Bishop, 1995, Ripley, 1996,1997), por su capacidad para lidiar hasta cierto grado con mayor ventaja con este tipo de información que otro tipo de modelo alterno.

2.-. Se requiere la utilización de una red estática.

El tipo de información que se tiene para el entrenamiento de la red; consiste en índices de competencia que están evaluando dimensiones de un individuo en relación a sus vecinos en instantes fijos de tiempo por lo cual se requiere de un tipo de red estático, que logre extraer de esta información la dinámica de este sistema biológico altamente complejo.

3.-Contar con una herramienta en la solución de problemas de predicción y clasificación.

De hecho las redes neuronales han sido conceptualizadas como técnicas estadísticas no paramétricas (Smith, 1993), al estar libres del cumplimiento de los supuestos teóricos de la estadística paramétrica, o también como técnicas de regresión no lineal (Sarle, 1998).

1.5. OBJETIVO.

Predecir la velocidad de crecimiento en diámetro y altura de la especie de pino (*Pinus Cooperi* Blanco ver figura 10) de un bosque haciendo uso de una red neuronal entrenada con de índices de competencia independientes de la distancia.



Figura 10 *Pinus Cooperi* Blanco.

CAPITULO II

2. METODOLOGÍA, MATERIALES Y RESULTADOS.

2.1. MATERIALES.

2.1.1. Descripción del sitio. El estudio se efectuó en el área experimental, Cielito Azul, ubicada en el predio las “Veredas”, lote 4, de la zona asignada a la Unidad de Conservación y Desarrollo Forestal 4, en el municipio de San Dimas, Estado de Durango, geográficamente localizado entre 24 22’ y 24 23’ norte y 105 53’ y 105 54’ oeste, con una altitud promedio de 2700 metros. La topografía es de ondulada a plana, con una pendiente promedio de 15%.

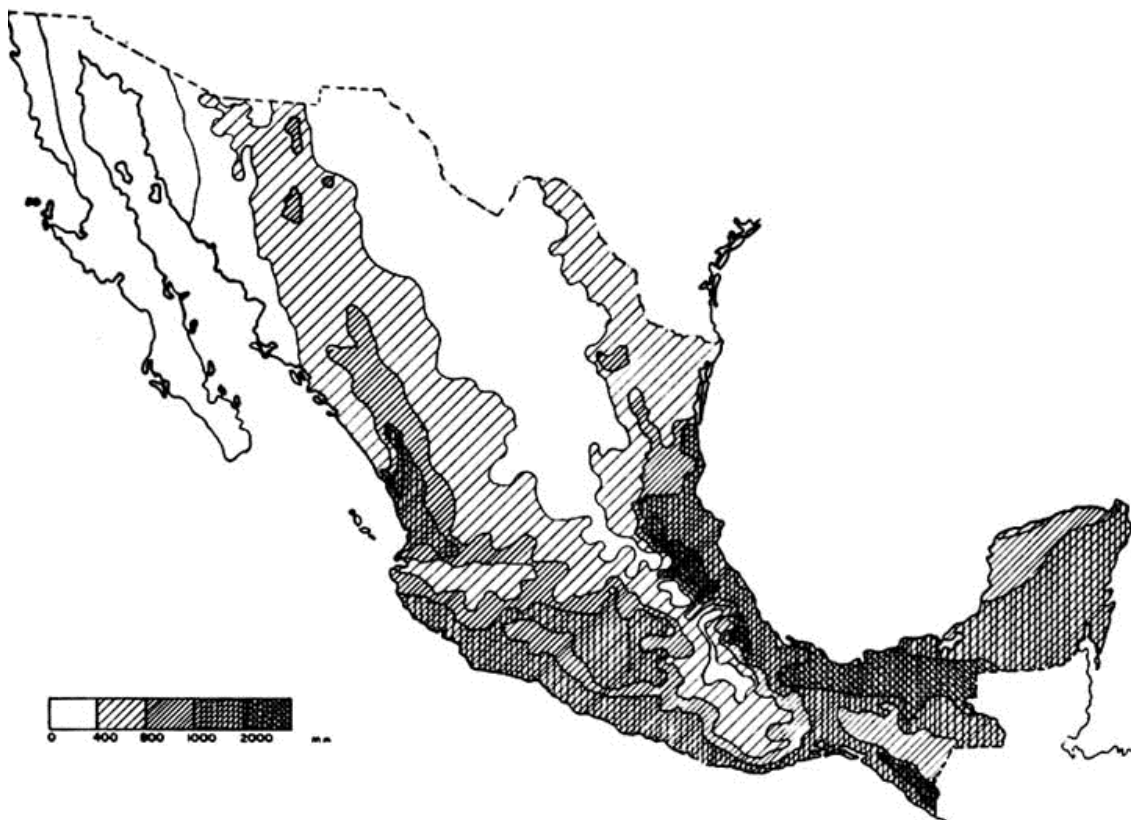


Figura 11 Precipitación media anual de lluvia

El sustrato se caracteriza por presentar una asociación de cambisol éutrico, predominan los suelos limo-arcillosos y limo-arenosos con un PH de 5.0 y un espesor de la capa de materia orgánica de 5 centímetros. El clima presenta características de templado semifrío subhúmedo, con lluvias en verano y precipitación durante el invierno mayor a 10.2 mm.

La precipitación media anual en el área es de 800 mm (figura 11). La vegetación arbórea esta constituida por estratos con especies de los géneros Pinus y Quercus, con dominio de Pinus Cooperi Blanco y Pinus Durangensis Martinez. Se encuentran, también diversos géneros de vegetación arbustiva y herbácea.

El bosque es de tipo coetáneo e in-coetáneo presentando una distribución en su dimensión de diámetro entre 5 y 80 centímetros principalmente.

El área basal promedio fluctúa entre 12 y 23 m² /ha. En ella se establecieron entre 1966 y 1968, 36 sitios permanentes de investigación silvícola (SPIS), aplicando cinco tratamientos de aclareo y un testigo, en los que se han hecho mediciones dendrométricas en el año de 1972, 1979, 1982, 1986, 1993 y 2004. Cada sitio está dividido en cuatro cuadrantes de 25 X 25 metros Manzanilla (1993). En cada cuadrante se estableció un punto de origen en sus vértices, a partir del cual se midieron las distancias en dos ejes (X,Y) para cada uno de los árboles y se ubicaron espacialmente figura 12.



Figura 12 Sitio experimental “cielito azul” de estudio del *Pinus Cooperi* Blanco

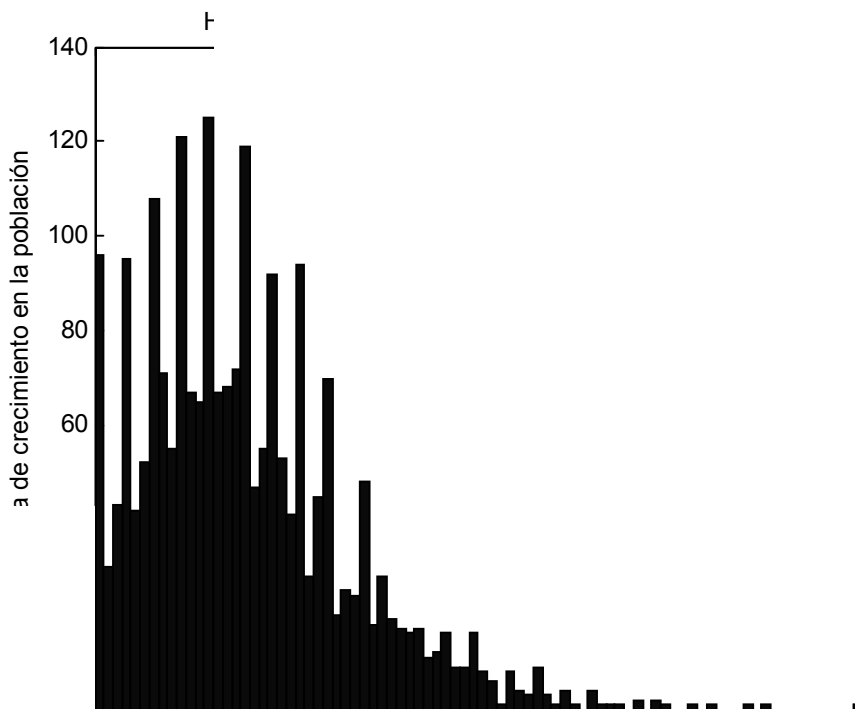
Además, de cada árbol se tomó medición de las siguientes variables: especie, tipo, diámetro normal, diámetro del tocón, grosor de la corteza, condición de daño, altura total, altura de fuste limpio, clase o dominancia, piso, vitalidad, tendencia dinámica y proyección de copa. Para este trabajo se utilizaron las mediciones dentro de un periodo de 5 años, acuerdo a como lo propone Manzanilla (1993), de 1982 y 1986 de con un total de 2195 pinos los cuales se utilizaron para el entrenamiento de las redes y la validación de los resultados.

2.1.2. Selección de la muestra de validación. La selección de la muestra de validación del aprendizaje de la red neuronal que se escogió buscó reflejar el comportamiento del sistema, para lo cual la selección se realizó utilizando conocimiento a priori del sitio experimental de estudio y observando la frecuencia de aparición del crecimiento con respecto a las dimensiones del individuo. Así se consideró la riqueza de nutrientes y agua del sitio su competencia, grupo de edad y la frecuencia de velocidad de crecimiento. Se seleccionó una muestra de validación 21 individuos para la predicción de crecimiento en diámetro y de 30 para el crecimiento en altura.

Se seleccionó una muestra de validación de 21 individuos para la predicción de crecimiento en diámetro. La muestra de validación de crecimiento en diámetro se encuentra dentro de un rango de velocidad de crecimiento de 0.6 a 3.8 cm / 5 años (tabla 1). Como se puede observar en el histograma de frecuencias de velocidad de crecimiento en diámetro (gráfica 1) Estos valores se encuentran dentro de la región de mayor frecuencia de aparición dentro de la población. El rango de crecimiento de toda la población esta entre un rango de crecimiento de 0 a 11 centímetros en cinco años como se puede ver en el histograma de frecuencias (gráfica 1).

La frecuencia del número de individuos con la misma velocidad de crecimiento de la muestra fluctúa en la población de un número de 21 para el crecimiento de 3.8 cm hasta 75 para un crecimiento de 2.2 cm.

En lo que respecta al crecimiento en altura se seleccionó una muestra de validación 30 individuos para la predicción de crecimiento en altura (gráfica 2). Se escogió entre los datos que aparecen con mayor frecuencia en las mediciones y algunos datos que se presentan con poca frecuencia pero cercanos a datos con gran frecuencia de aparición en la población.



Gráfica 1 Histograma de frecuencias de la velocidad de crecimiento en diámetro de la población de un bosque de *Pinus Cooperi* Blanco.

Como se puede observar 30 datos de la muestra de validación en crecimiento en altura (tabla 2), corresponden a 10 rangos de crecimiento en el histograma de frecuencias de la población. La mayoría de los 30 datos aparece con una frecuencia de 8 valores dentro de

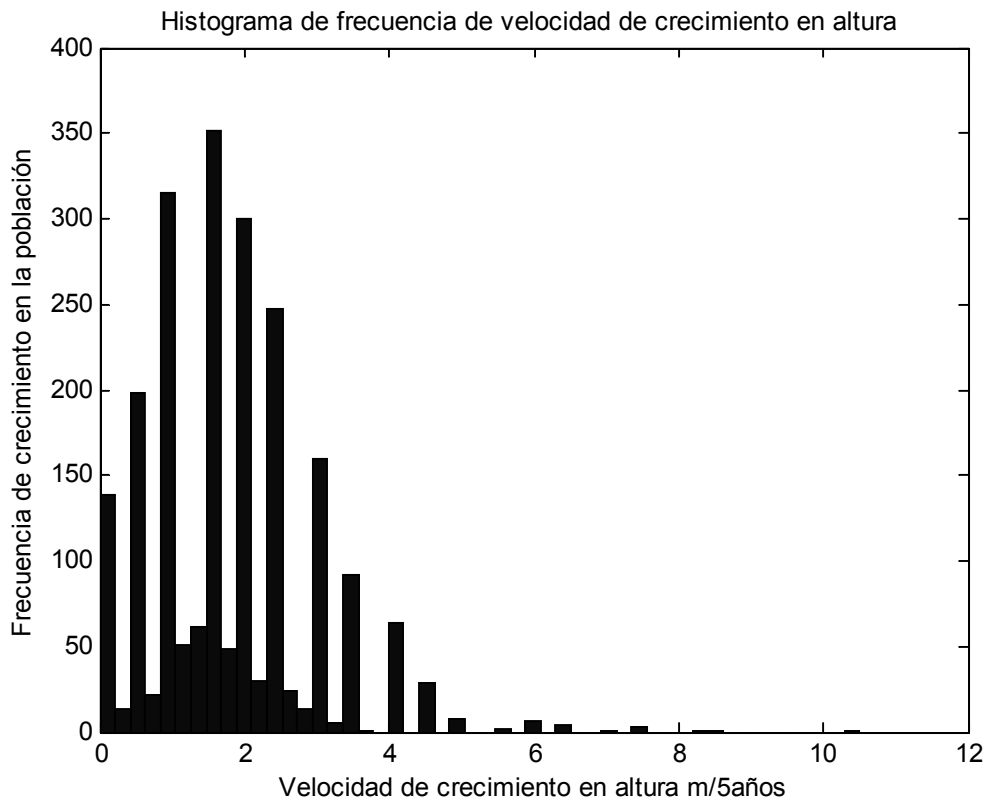
la muestra y corresponden a un crecimiento de 1.0 metro / 5 años, mientras en el histograma de frecuencias (gráfica 2) aparece con una frecuencia de 302 observaciones. En segundo lugar en numero de aparición dentro de la muestra se tiene el crecimiento de 1.5 metros/ 5 años con 6 datos mientras en el histograma de frecuencias de toda la población aparece con una frecuencia de 324 datos. Así la mayoría de las observaciones incluidas en la muestra aparecen con una frecuencia entre 157 a 324 en el histograma de frecuencias de toda la población.

Tres datos en la muestra aparecen con una frecuencia igual a 1, y estos corresponden a los crecimientos en altura de 2.6, 1.8, y 1.6 metros/ 5 años y su frecuencia en la población total del bosque es 13, 20 y 25 datos. Sin embargo una característica de rangos de crecimiento en altura que aparecen con poca frecuencia es que están entre rangos de crecimiento con una alta frecuencia de aparición.

En la selección de la muestra de crecimiento en diámetro y altura además de tomar en cuenta la frecuencia de aparición dentro de la población se consideraron otros factores como el relacionar la velocidad de crecimiento en diámetro y altura con respecto a parámetros como altura, diámetro, calidad del sitio, y los valores de sus índices de competencia independientes de la distancia.

Crecimiento en diámetro		
Frecuencia en la muestra	Velocidad de crecimiento m/5años	Frecuencia en la población
2	0.6	41
1	0.7	51
1	0.9	51
2	1.0	70
1	1.1	54
1	1.2	65
1	1.3	54
1	1.5	64
1	1.6	61
1	1.7	62
1	1.9	67
1	2.1	42
1	2.2	75
1	2.5	46
1	2.7	52
1	3.1	27
1	3.2	44
1	3.4	31
1	3.8	21

Tabla 1. Muestra de validación de la red neuronal back propagation entrenada para predecir crecimiento en diámetro.



Gráfica 2 Histograma de frecuencias de la velocidad de crecimiento en altura de la población de un bosque de *Pinus Cooperi* Blanco.

Crecimiento en altura		
Frecuencia en muestra	Velocidad de crecimiento m/5años	Frecuencia en la población
1	0.5	190
8	1.0	302
6	1.5	324
1	1.6	25
1	1.8	20
2	2.0	272
5	2.5	235
1	2.6	13
3	3.0	157
2	3.5	88

Tabla 2. Muestra de validación de la red neuronal backpropagation entrenada para predecir crecimiento en altura.

2.1.3. Determinación del tamaño de muestra.

La capacidad de generalización de cada red utilizada tendrá que ser comprobada contra una muestra de crecimiento en diámetro y otra en altura. Las muestras utilizadas para determinar una buena generalización de la red requieren que el tamaño de la muestra este dentro de un intervalo de confianza que sea aceptable. Esto se determinó por medio del procedimiento (Montgomery, 1991) siguiente:

$$\sigma = \frac{\sqrt{\sum (E - e)^2}}{(n - 1)} \quad (5)$$

$$E = \sum e \quad (6)$$

$$n = \frac{N * Z^2 * \sigma^2}{\Delta^2 * (N - 1) + Z^2 * \xi} \quad (7)$$

N = Tamaño de la población.

n = Tamaño de la muestra.

e = diferencia entre el valor pronosticado y el valor de validación.

Z = 0.95, 1.96, 3.0 para un intervalo de confianza de 90, 95, respectivamente.

Δ = error tipo I, 0.1, 0.05 0 para un intervalo de confianza de 90, 95, respectivamente.

ξ = Parámetro con un valor igual a 1.0

2.2. METODOLOGÍA.

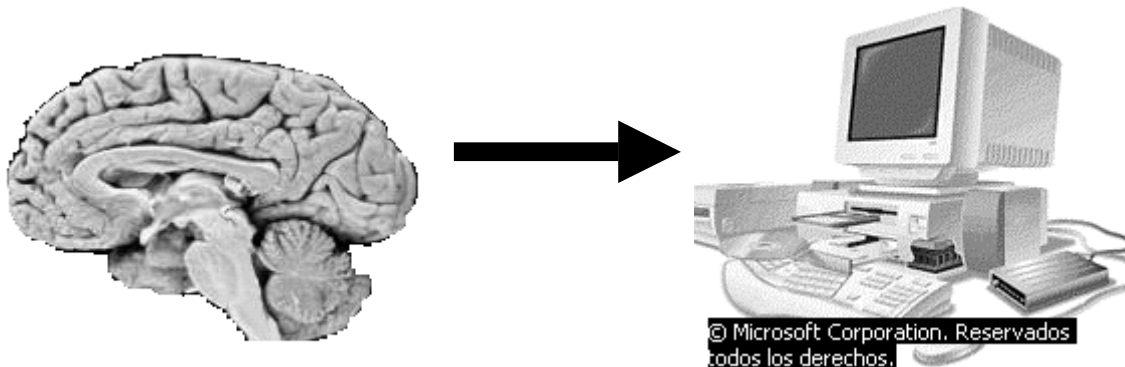


Figura 13. Se ha tratado de modelar el funcionamiento del cerebro por medio de las redes neuronales artificiales.

2.2.1 Red neuronal backpropagation (Red Neuronal Feedforward FANN). Esta red neuronal corresponde a la clase de ANN más estudiada por el ámbito científico y la más utilizada en los diversos campos de aplicación (Blum, 1991).

Los tres elementos esenciales en una FANN son:

1. Las neuronas o nodos: son los elementos básicos de procesamiento (Freeman, 1991, Searle, 1998).
2. La arquitectura de la red: descrita por las conexiones ponderadas entre los nodos.
3. El algoritmo de entrenamiento: usado para encontrar los parámetros de la red.

Matemáticamente, las FANN consisten en un conjunto de neuronas organizadas en capas como se muestra en la figura 14. Las capas se componen en capa de entrada, oculta y de salida (Rojas, 1995, Hassoun, 1995).

El número de unidades de la capa de entrada m está determinado por la aplicación. La arquitectura o topología de la red $A\lambda$ se refiere a la distribución topológica de las conexiones, donde λ representa el número de neuronas en la capa oculta.

Para las arquitecturas con una sola capa oculta, el número de unidades ocultas indexa diferentes clases de modelos ANN ($S\lambda$) ya que es un descriptor no-ambiguo de dimensionalidad d del vector paramétrico.

Dado un conjunto de observaciones, la tarea del aprendizaje neuronal es construir un estimador $g\lambda(x, w)$ de la función desconocida $h(x)$,

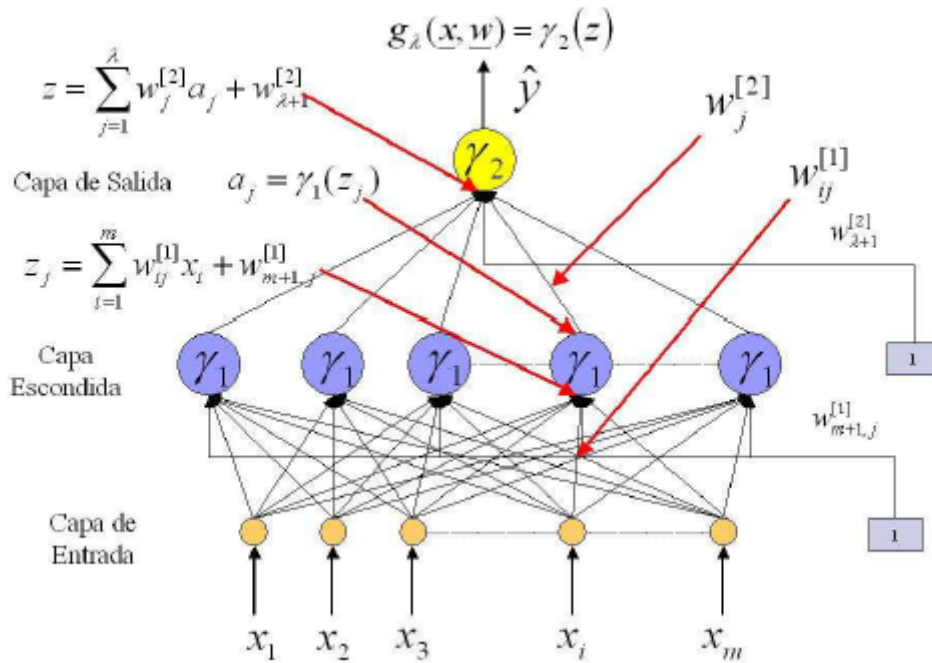


Figura 14: Una ANN feedforward de para aproximar $h(\cdot)$

$$\hat{y}_k = g_\lambda(\underline{x}, \underline{w}) = \gamma_2\left(\sum_{j=1}^{\lambda} w_{jk}^{[2]} \gamma_1\left(\sum_{i=1}^m w_{ij}^{[1]} x_i + w_{m+1,j}^{[1]}\right) + w_{\lambda+1,k}^{[2]}\right) \quad (8)$$

donde

- $w = (w_1, w_2, \dots, w_d)^T$ es el vector paramétrico a ser estimado compuesto por los pesos de las conexiones entre las neuronas de la red.
- γ_1 es una función no-lineal
- γ_2 es una función que puede ser lineal o no lineal
- λ es el parámetro de control que indica el número de neuronas escondidas.

La función γ_1 debe ser continua, monótona, acotada y diferenciable. Típicamente esta función es la sigmoideal dada por la ecuación 9,

$$\gamma_1(z) = \frac{1}{1 + e^{-z}} \quad (9)$$

Si la función γ_2 se elige no-lineal, debe ser continua, estrictamente monótona, acotada y diferenciable. La función sigmoideal dada por la ecuación 9, satisface estos requisitos. El parámetro estimado $\hat{\underline{w}}_n$ (ecuación 10) es obtenido minimizando iterativamente una función de costo $L_n(\underline{w})$, es decir:

$$\hat{\underline{w}} = \arg \min \{L_n(\underline{w}) : \underline{w} \in W\} \quad W \subseteq \mathbb{R}^d \quad (10)$$

$L_n(w)$ representa normalmente la función del promedio de los mínimos cuadrados, es decir (ecuación 11):

$$L_n(w) = \frac{1}{n} \sum_{i=1}^n (y_i - g(x_i, w))^2 \quad (11)$$

2.2.1.1 Algoritmo de Aprendizaje

El proceso de aprendizaje (Hastie, 1994, Karampiperis, 2002, White, 1994, Parisi, 1996) de una red neuronal feedforward es supervisado pues los parámetros de la red, conocidos como pesos, son estimados a partir de un conjunto de patrones de entrenamiento compuesto por patrones de entrada y salida $\{(x, y)\}_{i=1 \dots n}$. Para estimar los parámetros de la red, en 1986 Rumelhart, Hinton y Williams redescubrieron el algoritmo de *backpropagation* (Hinton, 1986, Rumelhart, 1986, Maclelland, 1986). El aprendizaje de redes neuronales feedforward es una generalización de la regla delta propuesta por Widrow-Hoff (1960).

Sea el error generado por el patrón (x^t, y^t) , $t = 1 \dots n$ dado por $E^t = \frac{1}{2} (y^t - g\lambda(x^t, wn))^T (y^t - g\lambda(x^t, wn))$, y sea E la medida de error total dado por $E = \sum_t E^t$, y considerando una red neuronal feedforward de tres capas con K neuronas de salidas descrita por la ecuación 7. Sea:

$z_j^t = \sum_{i=1}^m w_{ij}^{(1)} x_i^t + w_{m+1, j}^{(1)}$ la entrada a la neurona escondida j para el patrón t ;
 $a_j^t = \gamma_1(z_j^t)$ la salida de la neurona escondida; $z_k^t = \sum_{j=1}^{\lambda} w_{jk}^{(2)} a_j^t + w_{\lambda+1, k}^{(2)}$ la entrada a la neurona de salida k , y $a^t k = \gamma_2(z_k^t)$ la salida de la neurona de salida k .

Finalmente sean γ_1 y γ_2 diferenciables. Si se considera el parámetro α como la velocidad de aprendizaje, entonces la actualización de los pesos de la neurona de salida considerando el patrón t , viene dada por la ecuación 12,

$$\Delta_t w_{jk}^{[2]} = \alpha \delta_k^t z_j^t \quad (12)$$

como $\frac{\partial E^t}{\partial a_k^t} = -(y_k^t - a_k^t)$ entonces

$$\delta_k^t = -\frac{\partial E^t}{\partial z_k^t} = -\frac{\partial E^t}{\partial a_k^t} \frac{\partial a_k^t}{\partial z_k^t} = (y_k^t - a_k^t) \gamma_2'(z_k^t) \quad (13)$$

La actualización de los pesos de la neurona oculta j viene dado por la ecuación 14,

$$\Delta_t w_{ij}^{[1]} = \alpha \delta_j^t x_i^t \quad (14)$$

donde

$$\delta_j^t = -\frac{\partial E^t}{\partial z_j^t} = -\frac{\partial E^t}{\partial a_j^t} \frac{\partial a_j^t}{\partial z_j^t} = \gamma_1'(z_j^t) \sum_{k=1}^K \delta_k^t w_{jk}^{[2]} \quad (15)$$

donde la última igualdad se obtuvo a partir de la ecuación 16

$$\frac{\partial E^t}{\partial a_j^t} = \sum_{k=i}^K \frac{\partial E^t}{\partial z_k^t} \frac{\partial z_k^t}{\partial a_j^t} = \sum_{k=i}^K \frac{\partial E^t}{\partial z_k^t} \frac{\partial (\sum_{j=1}^{\lambda} w_{jk}^{[2]} a_j^t + w_{\lambda+1,k}^{[2]})}{\partial a_j^t} = \sum_{k=i}^K \frac{\partial E^t}{\partial z_k^t} w_{jk}^{[2]} = - \sum_{k=1}^K \delta_k^t w_{jk}^{[2]}$$

(16)

Un caso particular del algoritmo backpropagation descrito anteriormente es si se considera la función de activación la función sigmoideal descrita por la ecuación (9), cuya derivada es:

$f'(z) = f(z)(1 - f(z))$, donde $f(z) = \frac{1}{1+e^{-z}}$ es la función sigmoideal. En este caso al usar la función sigmoideal en γ_1 y γ_2 en las ecuaciones 13 y 15 la actualización de los pesos vendría dada por la ecuación 17:

$$(w_{jk}^{[2]})_{nuevo} = (w_{jk}^{[2]})_{anterior} + \alpha (y_k^t - z_k^t) z_k^t (1 - z_k^t) z_j^t \quad (17)$$

para las neuronas de salida, y para las neuronas ocultas se tendría la ecuación 18:

$$(w_{jk}^{[2]})_{nuevo} = (w_{jk}^{[2]})_{anterior} + \alpha (y_k^t - z_k^t) z_k^t (1 - z_k^t) z_j^t \quad (18)$$

2.2.1.2. Aplicaciones de las ANN

Las redes neuronales artificiales son modelos no-paramétricos flexibles utilizados para resolver un gran número de problemas (Ripley, 1996, Rumelhart and Hinton, 1986, Rojas, 1995). En particular, en el campo de la estadística computacional puede ser utilizado como modelos de regresión (Flexer, 1995, Hectht, 1990, Abhijit, 1995), en clasificación (Anderson, 1995, Singh, 1997).y en series temporales.

a) *Regresión* (Ripley, 1996, 1986). El problema consiste en modelar una función desconocida $h(x_1, \dots, x_n)$ teniendo como información solamente el conjunto de datos $\{x_1^i, y^i\}_{i=1}^n$. La salida es generada por el modelo donde se le puede agregar un ruido gaussiano i.i.d., es decir, $y_i = h(x_1^i, x_2^i) + a_i$, donde $a \sim N(0; \sigma_a^2)$.

b) *Clasificación*. El problema consiste en clasificar un determinado patrón que está caracterizado por un conjunto de características $\{x_1\}_{i=1}^n$ en la clase $k = 1 \dots K$.

c) *Series Temporales*. Un proceso estocástico $\{X_t\}_{t \in Z}$ se dice que es un proceso *Autoregresivo no lineal* de orden p ($NAR(p)$) si se cumple la ecuación 19,

$$x_t = h(x_{t-1}, x_{t-2}, \dots, x_{t-p}) + \varepsilon_t \quad (19)$$

donde $h(\cdot)$ es alguna función no-lineal continua y derivable; además se asume que

$$\mathbb{E}[\varepsilon_t | x_{t-1}, x_{t-2}, \dots] = 0 \quad \text{y que la varianza de } \varepsilon_t \text{ es } \sigma_\varepsilon^2.$$

El predictor óptimo del mínimo error cuadrático medio (MSE) de x_t dado $x_{t-1}, x_{t-2}, \dots, x_{t-p}$ viene dado por la esperanza condicional (para $t \geq p+1$)

$$\hat{x}_t = \mathbb{E}[x_t | x_{t-1}, \dots, x_{t-p}] = h(x_{t-1}, x_{t-2}, \dots, x_{t-p}) \quad (20)$$

Un problema central de los NAR es construir la función $h : \mathbb{R}^p \rightarrow \mathbb{R}$ para expresar x_t en la forma dada por la ecuación (13), donde h es una función no-lineal, derivable desconocida y ε_t denota ruido.

Se supone que $\mathbb{E}[\varepsilon_t | x_{t-1}, x_{t-2}, \dots] = 0$ y que ε_t tiene una varianza finita σ_ε^2 . Bajo estas condiciones, el predictor óptimo MSE de x_t dado $x_{t-1}, x_{t-2}, \dots, x_{t-p}$ está dado por la ecuación (20).

2.2.1.3. Utilización de modelos FANN.

Las tres aplicaciones presentadas anteriormente pueden ser representadas por una red neuronal feedforward (FANN), en la cual la FANN provee una aproximación no lineal al sistema a través de la ecuación (8). La red neuronal feedforward provee una estimación de la función no-lineal proceso que es llevado a cabo mediante la estimación de los parámetros $w = (w^{(1)T}, w^{(2)T})^T$, con $w^{(1)} = (w^{(1)}_{1,1}, \dots, w^{(1)}_{p+1,\lambda})^T$ y $w^{(2)} = (w^{(2)}_1, \dots, w^{(2)}_{\lambda+1})^T$, de la función $g\lambda(\cdot)$ mediante la minimización de la suma de los residuos al cuadrado de manera similar a la descrita en la ecuación (11). La estimación se efectúa a través del algoritmo de backpropagation (Bishop, 1995, Fogelman, 1994, Werbos, 1974).

2.2.1.4. Algoritmo de aprendizaje. Método de optimización.

Método del Gradiente Conjugado: Este algoritmo no involucra el cálculo de las segundas derivadas de las variables y converge al mínimo de la función cuadrática en un número finito de iteraciones. El algoritmo del gradiente conjugado (Moller, 1993, Caruana, 2000), sin aplicarlo aún al algoritmo de propagación inversa consiste en:

1. Seleccionar la dirección de p_0 , la condición inicial, en el sentido negativo del gradiente:

$$p_0 = -g_0 \quad (21)$$

Donde

$$g(k) \equiv \nabla e(x) |_{x=x_k} \quad (22)$$

2. Seleccionar la parámetro de velocidad de aprendizaje α_k para minimizar la función a lo largo de la dirección

$$x_{k+1} = x_k + \alpha_k p_k \quad (23)$$

3. Seleccionar la dirección siguiente de acuerdo a la ecuación

$$p_k = -g_k + \beta_k p_{k-1} \quad (24)$$

con

$$\beta_k = \frac{\Delta \mathbf{g}_{k-1}^T \mathbf{g}_k}{\Delta \mathbf{g}_{k-1}^T \mathbf{p}_{k-1}} \quad \text{o} \quad \beta_k = \frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}} \quad (25)$$

4. Si el algoritmo en este punto aún no ha convergido, regresamos al paso 2

Este algoritmo no puede ser aplicado directamente a una red neuronal porque el error no es una función cuadrática; lo que afecta al algoritmo en dos formas, primero no es hábil para minimizar la función a lo largo de una línea como es requerido en el paso 2; segundo, el error mínimo no será alcanzado normalmente en un número finito de pasos y por esto el algoritmo necesitará ser inicializado después de un número determinado de iteraciones (Wasserman, 1989).

A pesar de estas complicaciones, esta modificación del algoritmo Backpropagation converge en muy pocas iteraciones, y es incluso uno de los algoritmos más rápidos para redes multicapa (Pham, 2001, Hertz, 1991).

2.2.1.5. Ventajas del uso de redes neuronales.

La potencia computacional de una red neuronal deriva de su estructura de cálculo distribuido paralelo y de su capacidad de aprender y generalizar entendiendo por generalización el producir salidas razonables frente a entradas desconocidas (Chater, 1995, Haykin, 1994, Zurada, 1992). Estas propiedades harán posible la resolución de problemas que necesitarían gran cantidad de tiempo en ordenadores “clásicos” (Anderson, 1995, Ripley 1993, Searle, 1994).

El uso de redes neuronales ofrece las siguientes ventajas:

- a) No linealidad: Una neurona es un elemento no lineal por lo que una interconexión de ellas (red neuronal) también será un dispositivo no lineal (Michie, 1994, White, 1994). Esta propiedad permitirá la simulación de sistemas no lineales y caóticos.
- b) Capacidad de establecer relación entrada-salida: Una de las ramas principales de investigación en el campo de las redes neuronales es lo que se conoce como aprendizaje supervisado (Mehra, 51, Russell 1999, Thrun, 1991). Aquí se le presentan a la red una serie de entradas junto con las salidas deseadas de la red para esas entradas.

La red en su proceso de aprendizaje establece la relación entre las entradas y las salidas de manera que, ante entradas desconocidas, es capaz de dar una respuesta “aproximada” a la salida deseada.

- c) Adaptabilidad: Una red neuronal tiene la capacidad de modificar sus pesos de acuerdo con los cambios que se produzcan en su entorno de trabajo (cambios en las entradas, presencia de ruido, etc. (Soucek, 1991, Dayhoff, 1998, Reed, 1999).
- d) Tolerancia a fallos: Una red neuronal, al ser un sistema distribuido (Rumelhart, D.E., Durbin, 1995), permite el fallo de algunos elementos individuales (neuronas) sin alterar significativamente la respuesta del sistema total.

e) Posibilidad de implementación en VLSI: Esta posibilidad permite que estos sistemas puedan ser aplicados en sistemas de tiempo real, simulando sistemas biológicos mediante elementos de silicio.

f) Uniformidad de análisis y diseño: En todos los dominios de aplicación de las redes neuronales se usa la misma notación, además todas las redes neuronales tienen como elementos básicos las neuronas por lo que se hace posible teorías conjuntas para los diferentes algoritmos y aplicaciones de las redes neuronales (Cherkassky, 1994, Ripley, 1997, Smith, 1993).

g) Analogía biológica: El campo de las redes neuronales tiene como objetivo simular la capacidad de análisis y cómputo del cerebro humano (Chater, 1995, Hertz, 1990, Hiler, 1991). Este intento de simulación conlleva una colaboración entre biólogos e ingenieros.

2.2.1.6. Limitaciones del uso de redes neuronales.

En cuanto a los inconvenientes de las redes hay que destacar que son necesarios datos para su entrenamiento y, si la red es de un gran tamaño se necesita un número elevado de estos (McClelland, 1986, Rumelhart, 1986). En segundo lugar, la implementación de estos sistemas lleva asociado un conocimiento de programación.

2.2.2 Sistemas neuro-difuso (ANFIS).

Sistemas difusos

Cualquier sistema dinámico o estático que hace uso de la lógica difusa o de conjuntos difusos, se considera como un sistema difuso. Los sistemas difusos se pueden emplear en modelado y control (a nivel directo o supervisor) de sistemas, en reconocimiento de patrones, etc. Específicamente, en modelado, el sistema se puede describir mediante una relación difusa o un conjunto de reglas proposicionales difusas (cualitativas), que pueden tener distintas formas. Una proposición difusa se basa en la utilización de términos cualitativos A_i asociados a una variable x definida dentro de un universo acotado. Por ejemplo, la variable temperatura dentro del rango $[0,45]$ grados puede ser descrita mediante términos cualitativos: “*temperatura alta*”, “*temperatura baja*”, etc., con los que se pueden establecer proposiciones respecto a un objeto (proceso) particular, del tipo “*Si la potencia es alta entonces la temperatura aumenta*”, para el caso de modelado difuso, o “*Si la temperatura es alta, entonces disminuir la potencia*”, para el caso de control difuso. Principalmente, existen dos tipos de sistemas difusos, conocidos como “Mamdani” (26), y Takagi-Sugeno (27), de acuerdo con la forma del consecuente de cada regla:

$$r^l : \text{If } x \text{ is } A_i^l \text{ then } y \text{ is } B_j^l, 1 \leq i \leq n_x, 1 \leq j \leq n_y \quad (26)$$

$$r^l : \text{If } x \text{ is } A_i^l \text{ then } y = f_j(x), 1 \leq i \leq n_x, 1 \leq j \leq n_y \quad (27)$$

Un conjunto de reglas r^l ($1 \leq l \leq N$), forma una base de reglas. En (26) se representa un sistema difuso con antecedentes y consecuentes proposicionales, en cambio, en (27) se representa un sistema con entradas proposicionales y salidas precisas descritas mediante funciones estáticas o dinámicas. Es bien conocido que existen diferentes métodos de inferencia del resultado de cada regla, del tipo “Modus Ponens Generalizado”, en

conjunto con métodos de “conjunción” de proposiciones (para el caso de sistemas difusos con más de una entrada) y de “agregación” de los resultados individuales, para determinar el valor difuso o preciso de la salida del sistema completo, para un determinado valor de las entradas.

Existen también, diversos métodos para convertir valores precisos en difusos y viceversa (denominados métodos de “fusificación” y de “defusificación”, respectivamente).

Una de las propiedades más importantes de las bases de reglas difusas, es su capacidad de aproximación de funciones no lineales, con error de aproximación acotado. Esto permite asegurar que, con un conjunto “infinito” de reglas es posible aproximar cualquier función con error de aproximación tendiendo a “cero”, dentro del universo definido para las entradas de la base de reglas. Este compromiso entre la precisión y la complejidad del modelo difuso debe ser considerado, pues no resulta práctico operar con una base de reglas demasiado grande, además de la pérdida de comprensión de la misma por un experto.

La extensión al caso multivariable es directa y en ese caso, el número de reglas se relaciona con el número de conjuntos difusos definidos para cada una de las “n” variables de entrada:

$$N = \prod_{k=1}^n n_k$$

(28)

Esta “maldición” de la dimensionalidad, o crecimiento explosivo de reglas a mayor cantidad de términos difusos por variable de entrada, se puede reducir significativamente mediante esquemas jerarquizados de sistemas difusos, o bien, combinando algebraicamente algunas de las variables de entrada, a lo que se denomina “fusión sensorial” (Jamshidi, 1997).

Aunque el diseño del sistema difuso (modelo) se realice tomando en cuenta el conocimiento de un experto, o mediante algún mecanismo de optimización para el ajuste de sus parámetros, finalmente éste constituye una *aplicación* (mapeo) no lineal de las entradas en las salidas, en su versión más básica.

2.2.2.1. ANFIS, estructura y funcionamiento.

ANFIS. Significa Adaptive Neuro-Fuzzy Inference System y denomina al tipo de red desarrollada por Jyh-Shing Jang que combina conceptos y algoritmos de redes neuronales tipo Backpropagation junto con elementos de lógica borrosa o fuzzy logic (Braae, 1979).

La red en sí es básicamente una red Feed Forward con una estructura muy rígida que modela una serie de reglas de inferencia de Takagi-Sugeno.

Esta red combina nodos sintonizados por entrenamiento con nodos fijos de diverso tipo.

Estructura y funcionamiento. El ANFIS (Jan, 1993,1997) implementa el razonamiento borroso de tipo 3 (Kosko, 1994), que implica un proceso de defuzzificación algebraico en la cual la salida o parte consecuente de cada regla es una combinación lineal de las

variables de entrada y la salida final del sistema se obtiene como una suma ponderada de todas las reglas presentes.

La ponderación de cada regla es obtenida en base a la premisa de la regla.

La premisa consta de una serie de funciones de pertenencia asociadas a variables lingüísticas (una por cada variable de entrada) que son evaluadas y combinadas mediante la aplicación de una *norma-t*. Donde una *norma-t* puede pensarse como una especie de generalización del *AND lógico*.

La figura 15 muestra este tipo de razonamiento en forma esquemática.

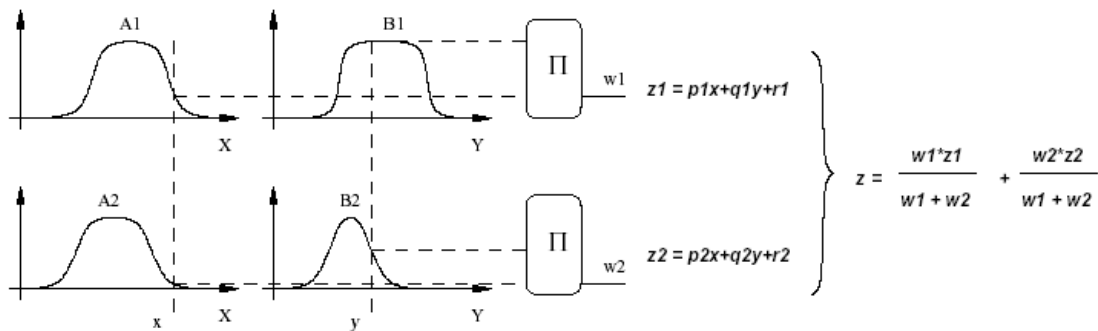


Figura 15: Razonamiento borroso tipo Takagi-Sugeno: El peso de cada parte consecuente es obtenido en base a las funciones de pertenencia de esa regla evaluadas para las variables de entrada. La parte consecuente es calculada como un producto lineal de las entradas. Por último la salida total es la suma ponderada de todas las partes consecuentes

En el ejemplo se muestra un sistema de este tipo que toma dos variables de entrada x e y definiéndose 2 reglas de inferencia (1 y 2), cada una con sus funciones de pertenencia:

A_1 y B_1 para x e y en la regla 1 y A_2 y B_2 para la regla 2.

La salida de cada regla i , z_i , es obtenida como combinación lineal más un término independiente de las entradas x e y en base a los parámetros p_i , q_i y r_i definidos para cada regla. Esta salida es el modelo de la parte consecuente de la regla para este tipo de razonamiento borroso.

Las funciones de pertenencia son evaluadas para cada regla y los pesos obtenidos para cada regla serán entonces

$$w_1 = A_1(x)B_1(y) \tag{29}$$

para la regla 1 y

$$w_2 = A_2(x)B_2(y) \tag{30}$$

Para la regla 2.

Los pesos obtenidos son entonces normalizados entre 0 y 1

$$\overline{w}_1 = \frac{w_1}{w_1 + w_2}$$

$$\overline{w}_2 = \frac{w_2}{w_1 + w_2}$$

(31)

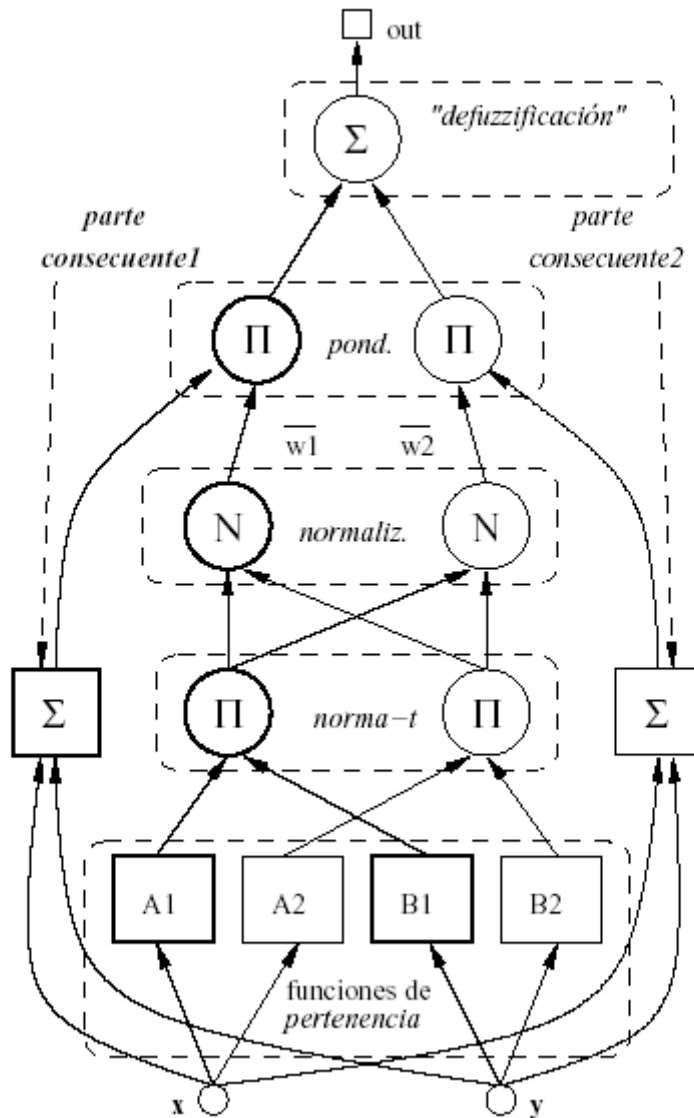


Figura 16. Estructura de una red ANFIS. Los bloques y las conexiones resaltados con trazo grueso corresponden a lo que se necesita agregar para modelar una regla.

Por último, la salida del sistema, Z , está dada por la suma ponderada de las salidas de las partes consecuentes de cada regla en base a los pesos obtenidos:

$$Z = \overline{w}_1 z_1 + \overline{w}_2 z_2$$

(32)

Para cada nueva regla incorporada a la red, se crean una serie de unidades y conexiones que la modelan, es decir, distintos tipos de funciones de red y salida (no se utiliza la activación en estas unidades).

La figura 16 muestra la estructura necesaria para definir el sistema de inferencia determinado por las reglas del ejemplo anterior. En la figura pueden verse los tipos de nodos utilizados para cada parte del sistema:

- Las funciones de pertenencia están definidas como unidades capaces de ser entrenadas cuya función de salida es de tipo campana “*bell*” (u otra función cuya forma sea similar). Cada una de estas recibe como única entrada a una de las variables de entrada.
- Las *normas-t* son implementadas como nodos no viables de entrenar y utilizando la función identidad como función de activación y de salida.

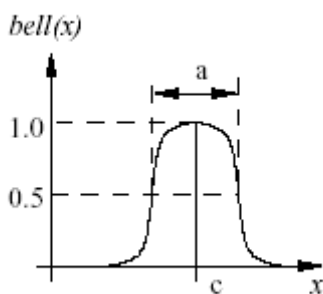


Figura 17. $bell(x)$, utilizada como función de pertenencia a las variables lingüísticas definidas en la premisa de las reglas

- La normalización es efectuada por unidades no viables de entrenar cuya función de red es especialmente diseñada con el fin de producir una salida que es la versión normalizada de la primera entrada con respecto al total de las entradas.
- La ponderación es efectuada por unidades análogas a las utilizadas para la *norma-t*.
- La parte consecuente es implementada como una unidad capaz de ser entrenada de tipo ADALINE, es decir, cuya función de red es la suma ponderada de pesos adaptables que representan los parámetros de la combinación lineal antes mencionada.
- La salida de la red es una unidad no viable de entrenar cuya función de red es de tipo sumatoria con los pesos de entrada fijos a 1.

En la figura 17 puede verse la forma de la función $bell(x)$ según los tres parámetros a , b y c .

2.2.1.2. Entrenamiento de ANFIS:

El algoritmo de entrenamiento es un híbrido de Backpropagation con el de mínimos cuadrados.

Esta técnica de entrenamiento efectúa en la primera época un paso de adaptación utilizando la técnica de los mínimos cuadrados. En este paso se adaptan los parámetros de la parte consecuente solamente siendo estos los coeficientes de la combinación lineal y el término independiente de las entradas que forma la salida de cada regla i ; en el ejemplo p_i , q_i y r_i respectivamente.

Este paso, según el autor, reduce drásticamente los tiempos de entrenamiento de la red.

Luego de esta pasada se fijan los valores obtenidos para la parte consecuente y se continúa la adaptación mediante el algoritmo estándar de backpropagation con momentum para adaptar a los otros parámetros variables de la red (en este caso a los parámetros de las funciones de pertenencia).

2.2.2.3. Sistemas híbridos neuro difusos.

El avance de los sistemas de inteligencia computacional y la búsqueda de soluciones cada vez más completas para los problemas que pueden ser resueltos por estos métodos, han llevado a la unión de técnicas de soft computing (Li, 1996) haciendo que las debilidades de unos sistemas se compensen con las bondades de otros. Esta última, resulta ser una fuerte motivación para realizar la unión de las técnicas de Redes Neuronales Artificiales y Lógica Difusa (LD) (Nebot, 1997, Boosley, 1998, Tsoukalas, 1997). En el tabla 3 se pueden observar las similitudes y diferencias entre estas dos técnicas.

Re-definiendo las motivaciones para crear los sistemas híbridos neuro-difusos, podemos esbozar las siguientes:

- Los problemas a resolver con estas técnicas inteligentes tienen generalmente una componente de datos empíricos y otra de conocimiento previo, que con los sistemas neuro-difusos pueden ser explotados a cabalidad mezclando los métodos cuantitativos y cualitativos de las RNA y la LD respectivamente (Kosko, 1991).
- La expresión de conocimiento como una base de reglas lingüísticas, enriquecida o surgida con o a través de datos cuantitativos, es una forma adecuada para hacer realimentación al usuario humano de la forma como el sistema realiza las tareas.
- Habiendo encontrado que la representación del conocimiento como un Sistema Difuso se acerca a la forma de razonamiento humano, es natural intentar adicionar a estos sistemas capacidades de aprendizaje, de forma que se emule mejor tal comportamiento o que se mejoren los resultados obtenidos con el sólo Sistema Difuso (Mehrotra, 1997, Fausset, 1994, Nauck 1995). (tabla 3).

	RNA	Lógica Difusa
Paralelismo	Bueno	Bueno
Dominio	Continuo	Continuo
Linealidad	No lineal	No lineal
Base de Conocimiento	Distribuida en pesos	Distribuida en Base de Reglas
Método de	Matemática mediante	Razonamiento sobre la

Solución	propagación de la información de entrada	Base de Reglas
Representación del conocimiento	Numérica	Lingüística
Conocimiento explotado	Datos empíricos	Conocimiento humano previo
Capacidad de aprendizaje	Alta	No tiene - Es una forma lógica de tratamiento matemático.

Tabla 3

Comparación entre las técnicas de RNA y Lógica Difusa.

2.2.2.4. Taxonomía de los sistemas neuro-difusos (NF). Muchas de las aproximaciones de los sistemas NF implementadas hasta hoy, pueden ser enmarcadas dentro de un grupo (Nauk, 1993) que se basa en las diferentes combinaciones de las RNA y los Sistemas Difusos:

- Modelos Concurrentes: Cuando la RNA y el Sistema Difuso trabajan juntos pero sin interactuar el uno en el otro, es decir, ninguno determina las características del otro.
- Modelos Cooperativos: Cuando la RNA se usa para determinar los parámetros de funcionamiento del sistema difuso. En estos modelos, se distinguen dos fases: la de entrenamiento y la de funcionamiento. En la primera, la RNA interactúa con el sistema difuso determinando los parámetros del mismo, mientras en la segunda, la RNA desaparece dejando sólo el sistema difuso. Ejemplo de estos sistemas es la determinación de parámetros de control de un sistema difuso.
- Modelos Híbridos: En esta aproximación, los sistemas difusos y de RNA, trabajan juntos en una arquitectura homogénea que puede ser interpretada como una RNA con parámetros difusos o como un sistema difuso con parámetros o funcionamiento distribuidos. Las ANFIS (Jan, 1993,1997) y el NEFCON (Nauk, 1995).

2.2.2.5. Ventajas de los sistemas neuro-difusos

La estructura neuro-difusa permite el uso de métodos cualitativos y cuantitativos en la construcción de modelos, en la etapa de aprendizaje, en la de funcionamiento y en la realimentación de conocimiento al usuario, permitiendo así, todo el conocimiento

disponible (Rumelhart and Durbin, 1995, Skapura, 1995). Los métodos cuantitativos se observan en el aprendizaje que los sistemas neuro-difusos pueden realizar a partir de datos del ambiente y en su forma matemática de funcionamiento, mientras los métodos cualitativos se reflejan en el sentido lingüístico que tiene la base de conocimiento de los sistemas difusos y que genera una interacción bilateral sistema-hombre, permitiendo la introducción de conocimiento previo al sistema y la realimentación del estado final del modelo en forma de reglas expresadas lingüísticamente desde el sistema al hombre. Las otras ventajas (Martín, 1997, Werbos, 1993,), son las ya conocidas ampliamente y que vienen heredadas de los sistemas difusos y de las RNA (Tarssenko 1998,), tales como interpretación lingüística, introducción de conocimiento previo, autoaprendizaje, generalización, interpolación, etc.

2.2.2.6. Limitaciones de los sistemas neuro-difusos.

1. La dificultad de la dimensionalidad: Los sistemas difusos presentan el problema del crecimiento exponencial del tamaño del modelo respecto al número de variables que poseen, haciendo impracticable su uso para sistemas de más de cuatro dimensiones. Este crecimiento se puede representar en dos formas diferentes:

a) Tamaño del modelo: El número de conjuntos multivariados de la entrada crece exponencialmente con el número de entradas:

$$p = \prod_{i=1}^n p_i \quad (33)$$

de la ecuación (33) p es el número de conjuntos multivariados, n es el número de entradas y p_i es el número de conjuntos difusos para la entrada i . Como resultado de este crecimiento exponencial, el costo de cálculo y de implementación crece en igual forma y la capacidad de interpretación se limita cada vez más (Jamshidi, 1997, Hagan, 1994).

b) Tamaño de las muestras: Si N_1 es el tamaño de muestra apropiado para un sistema univariado, entonces N , representado en la ecuación 34,

$$N=N_1 \quad (34)$$

es el tamaño de muestra requerido para un sistema de n variables. En este caso, se presentan los mismos resultados que en el caso anterior, y además el espacio muestral puede verse tan poblado que puede representar pobremente al sistema.

2. Calidad de los datos: Existen problemas asociados con la cantidad, distribución y ruido en y de los datos. Si los datos están mal distribuidos, el aprendizaje puede no ser muy bueno, presentando especialmente problemas de generalización de interpolación si faltan muchos datos en cierta región. Este inconveniente es herencia de los sistemas neuronales y puede ser superado interpolando los datos, haciendo un tratamiento previo o buscando una representación adecuada de los mismos.

3. Conocimiento previo: En los sistemas neuro-difusos es deseable tener el conocimiento previo y confiable de un experto en el sistema, pues de otro modo, se debe realizar un modelamiento por "Fuerza Bruta" exclusivamente a través de los datos numéricos empíricos y existirán problemas al no haber descripción cualitativa del sistema, a menos que se apliquen técnicas de minería de datos (descubrimiento de conocimiento) a la base de conocimiento del sistema neuro-difuso.

2.2.3. Parámetros para entrenamiento de la red neuronal. Los índices de competencia independientes de la distancia utilizados para entrenar la red neuronal son parámetros construidos de dimensiones sencillas de medir de los árboles como el diámetro, altura, amplitud de copa, etc. En este trabajo son utilizados para entrenar la red neuronal una serie de índices independientes de la distancia, propuestos por Valles (2005) y un índice propuesto por Glover y Hool (1979) para predecir el crecimiento en diámetro (Hilt, 1987) y altura los cuales a continuación son descritos:

Índice de competencia de Glover y Hool (IAB). El índice representa la relación del área basal de un árbol con respecto al área basal media sobre la parcela de muestreo. Donde el índice se calcula como (ecuación 35):

$$IAB = \frac{[(\sum (\pi(D_j / 2)^2)) / n]}{(\pi(D_i / 2)^2)} \quad (35)$$

Índice de Glover y Hool.

Donde IAB_i es el índice de área basal para el árbol i , D_i , es el diámetro normal del árbol sujeto, D_j , es el diámetro del competidor, π (3.1416) es la constante para calcular el área basal, n es el número de árboles dentro de la parcela.

Índice de competencia de área basal por densidad (IABD). Una variación del índice de Glover y Hool se define como el índice de área basal número de árboles para el árbol i sujeto a la competencia D_i , D_j , π y n como en la ecuación 36. D es representado como la densidad dentro del sitio o parcela.

$$IABD = \frac{[(\sum (\pi(D_j / 2)^2)) / n]}{(\pi(D_i / 2)^2)} (na / ha) \quad (36)$$

Índice de competencia de área basal por densidad.

na/ha = número de árboles por hectárea.

Índice de competencia de altura media (ecuación 37):. Índice independiente de la distancia para medir la competencia por luz al considerar la altura de los árboles sujetos a competencia con sus vecinos dentro de la parcela. IAM_i , es el índice de altura media del árbol i sujeto a competencia, AT_j es la altura de los competidores j en la parcela de

muestreo, n es el número de árboles en la parcela, AT_i es la altura del árbol sujeto a competencia.

$$IAM = \frac{[\sum(AT_j)/n]}{(AT_i)} \quad (37)$$

Índice de competencia de altura media.

Índice de competencia de altura media por densidad. Una variante del índice anterior es adicionar la densidad establecida dentro de la parcela y de esta manera conocer el efecto sobre este índice para estimar la competencia, en donde AT_j , AT_i , y n representan los parámetros de la ecuación 38, D representa el mismo parámetro de la ecuación 8.

$$IAMD = \frac{[\sum(AT_j)/n]}{(AT_i)}(na/ha) \quad (38)$$

Índice de competencia de altura media por densidad.

na/ha = número de árboles por hectárea.

Índice de competencia de diámetro cuadrático. Es la relación del diámetro cuadrático obtenido en la parcela al diámetro normal del árbol i sujeto a competencia (ecuación 39):

$$IDC = \frac{\left[\sqrt{\left[\frac{\sum D_j^2}{n} \right]} \right]}{(D_i/100)} \quad (39)$$

Índice de competencia de diámetro cuadrático.

Índice de competencia de diámetro cuadrático por densidad (IDCD). Este índice además de medir la competencia en base al tamaño medio de los árboles dentro de la parcela con relación al tamaño de cada uno de los árboles i sujetos, a la competencia también depende de la población de árboles en la parcela. Por tal razón se agrega la densidad para estudiar su contribución en la estimación de la competencia individual (ecuación 40).

$$IDCD = \frac{\left[\sqrt{\left[\frac{\sum D_j^2}{n} \right]} \right]}{(D_i/100)}(na/ha) \quad (40)$$

Índice de competencia de diámetro cuadrático por densidad

na/ha = número de árboles por hectárea.

Índice de competencia de longitud de copa. La longitud de copa viva del árbol es representada (ILC) como una medida cuantitativa que muestra el grado de presión lateral ejercida sobre el árbol sujeto a competencia por los árboles vecinos. La

influencia de la longitud de copa viva tiende a disminuir a medida que se ejerce mayor competencia como resultado de una mayor poda natural (ecuación 41):

$$ILC = \frac{[(\sum(AT_j - AF_j))/n]}{(AT_i - AF_i)} \quad (41)$$

Índice de competencia de longitud de copa.

Donde ILC_i es el índice de competencia de longitud de copa de árbol sujeto a competencia; LC_i es la longitud de copa viva del árbol sujeto i ; AF_j es la altura del fuste limpio del competidor j ; AT_j es la altura total del competidor.

Índice de competencia de longitud de copa por densidad. Este índice mide los efectos de las longitudes de copa promedio del sitio con relación a la longitud de cada árbol i sujeto a competencia y además ponderado por la densidad o población total de árboles dentro de la misma (ecuación 42):

$$ILCD = \frac{[(\sum(AT_j - AF_j))/n]}{(AT_i - AF_i)}(na/ha) \quad (42)$$

Índice de competencia de longitud de copa por densidad.

na/ha = número de árboles por hectárea.

Índice de competencia de amplitud de copa (IAC). En este índice se considera la amplitud de copa promedio de los árboles competidores, la cual es relacionada con las áreas de las copas de los árboles i sujetos a competencia (ecuación 43):

$$IAC = \frac{[(\sum Rcn_j + Rcs_j + Rce_j + Rco_j)/4]/n}{(Rcn_i + Rcs_i + Rce_i + Rco_i)/4} \quad (43)$$

Índice de competencia de amplitud de copa.

Los parámetros $Rcn_j, Rcs_j, Rce_j, Rco_j$, son el radio de copa norte, sur, este, oeste para el competidor j ; $Rcn_i, Rcs_i, Rce_i, Rco_i$, son el radio de copa norte, sur, este, oeste para el árbol i sujeto a competencia; n es el número de árboles dentro de la parcela.

Índice de competencia de amplitud de copa por densidad (ecuación 44). Al agregar el factor de ponderación de la densidad es con el objetivo de observar el efecto de dicho factor y estudiar su contribución en predicción del crecimiento arbóreo.

$$IACD = \frac{[(\sum Rcn_j + Rcs_j + Rce_j + Rco_j)/4]/n}{(Rcn_i + Rcs_i + Rce_i + Rco_i)/4}(na/ha) \quad (44)$$

Ecuación 15. Índice de competencia de amplitud de copa por densidad.

na/ha = número de árboles por hectárea.

Índice de competencia de área basal total. Este índice relaciona el área basal total con el área basimétrica de los árboles i sujetos a competencia dentro de un sitio circular de muestreo (ecuación 45):

$$IABT = \frac{\left[\sum (\pi((D_j/2)^2)) \right]}{(\pi(D_i/2)^2)} \quad (45)$$

Ecuación 16. Índice de competencia de área basal total.

Índice de sitio (Polimorfico2). A partir de la función de Chapman-Richards (Valles, 2000) se desarrolló el índice de sitio polimorfico 2 el cual tiene la forma (ecuación 46):

$$ISP\ 2 = \beta_1 \left[1 - \exp(-\beta_2 * eb) \right]^{\beta_3} \quad (46)$$

$$\beta_3 = \ln(hd / \beta_1) / \ln(1 - \exp(\beta_2 * edad)) \quad (47)$$

Ecuación 17 Índice polimorfico2 (de sitio)

Donde:

Isp2: Índice con el modelo polimorfico2

eb: es la edad base o de referencia para estimar el índice de sitio IS.

hd: Altura dominante en metros.

β_1 β_2 : Parámetros del modelo

2.2.4. Software Utilizado

- las redes neuronales backpropagation se construyeron haciendo uso de la toolbox neural nets del software matlab versión 6.5 (Demuth, 2003).

2.2.5. Procedimiento de Solución.

En la predicción del crecimiento del diámetro de pino (*Pinus Cooperi* Blanco), se procedió en la solución de la siguiente forma:

2.2.5.1 Selección de Variables.

Como primer paso el cual es básico, se hizo una selección de las variables que describen el sistema, pues si las variables que se seleccionan no son las adecuadas no habrá una red que logre describir el comportamiento del sistema en estudio, en este trabajo los índices de competencia independientes de la distancia y dimensiones del pino estudiado fueron seleccionadas como las variables significativas para describir el crecimiento de la especie *Pinus Cooperi* Blanco.

2.2.5.2 Pre- tratamiento de datos de entrenamiento.

Para lograr un funcionamiento adecuado de la red fue necesario dar a los datos de entrenamiento de la red un pre-tratamiento para la inicialización de los pesos y bias.

Entre las alternativas que se tuvieron para normalizar los patrones de entrenamiento, para entrenar la red neuronal backpropagation se selecciono entre:

Normalizar los datos de entrenamiento con media cero y desviación unitaria.

$$pn = (p - \text{meant}) / \text{stdt} \quad (48)$$

P = Matriz RXQ de vectores de entrada

PN= matriz RXQ de vectores de entrada normalizados con media cero desviación estándar uno.

MEANT= Vector RX1 conteniendo la media para cada vector de la matriz P.

STDT= Vector RX1 conteniendo la desviación estándar para cada vector de P.

$$tn = (t - \text{meant}) / \text{stdt} \quad (49)$$

TN - Matriz SxQ SxQ de vectores objetivo normalizados.

MEANT – Vector Sx1 conteniendo la media para cada T.

STDT - Vector Sx1 conteniendo la desviación estándar para cada T.

O hacer un escalamiento dentro del rango [-1, +1] de acuerdo al intervalo de la función de activación tangencial.

$$pn = 2*(p - \text{minp}) / (\text{maxp} - \text{minp}) - 1; \quad (50)$$

PN - Matriz RxQ de vectores de entrada normalizados.

MINP- Vector Rx1 conteniendo los valores mínimos para cada P.

MAXP- Vector Rx1 conteniendo los valores máximos para cada P.

$$tn = 2*(t - mint)/(maxt - mint) - 1; \quad (51)$$

TN - Matriz SxQ de vectores objetivo normalizados.

MINT- Vector Sx1 conteniendo los valores mínimos para cada T.

MAXT- Vector Sx1 conteniendo los valores máximos para cada T.

Selección del tipo de pre-procesamiento de los patrones de entrenamiento.

Para observar el efecto del preprocesamiento de los patrones de entrenamiento en el tiempo de entrenamiento de una red. Se hizo un escalamiento en los patrones de entrenamiento en un rango de [-1, +1]. Posteriormente se entrenó una red con una arquitectura de una capa oculta de 90 neuronas con función de activación tangencial y una capa de salida de una neurona con función de activación lineal, utilizando el algoritmo de gradiente conjugado escalado. Se consumió un tiempo de entrenamiento de 1082.416 segundos para alcanzar un error cuadrático medio de 1×10^{-2} en 6150 épocas.

Los parámetros que se utilizaron para el entrenamiento fueron:

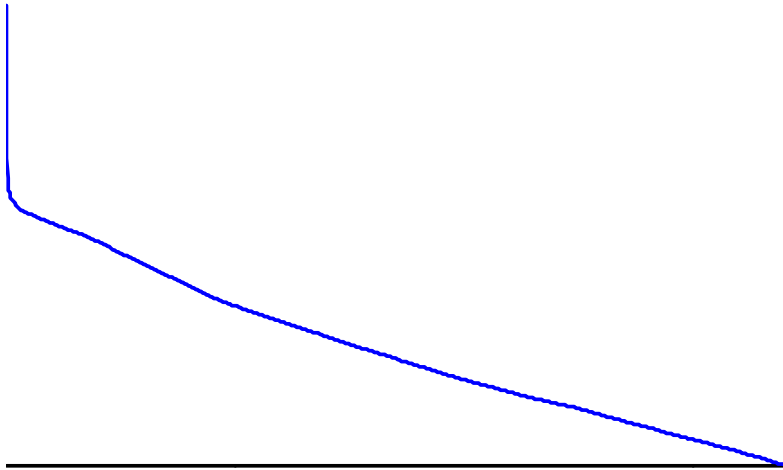
Las dimensiones: área basal por cuadrante, incremento de área basal por cuadrante, diámetro cuadrático por cuadrante, edad, diámetro a 0.3 cm., diámetro a 1.3 cm., relación entre diámetros 0.3 y 1.3, altura.

Y los índices: índice de sitio, índice de Glover y Hool, índice de área basal número de árboles, índice de altura media basal número de árboles, , índice de diámetro cuadrático número de árboles, índice de longitud de copa número de árboles, y índice de amplitud de copa número de árboles.



Gráfica 3 Escalamiento [-1 +1] de patrones de entrenamiento

Entrenando la misma red pero ahora haciendo un pre-procesamiento a los patrones de entrenamiento a media cero y varianza uno, el resultado que se obtuvo fue con un tiempo de 229.59 segundos en 1306 épocas hasta alcanzar un error cuadrático medio de 1×10^{-2} .



Gráfica 4 Escalamiento Media 0 varianza 1 de patrones de entrenamiento

Por lo que se consideró más conveniente utilizar un reprocesamiento de los patrones de entrenamiento a valores con media cero y varianza uno.

2.2.5.3. Selección de Arquitectura de la ANN.

El siguiente paso fue entrenar la red por lo cual se debió escoger la morfología más adecuada; así se selecciono la reportada como aproximador universal (Kolmogorov, 1957, Hornik, 1989, Cybenko, 1989, Blum, 1991).

Ya habiendo escogido la morfología con funciones de activación sigmoidales de las unidades de procesamiento en las capas ocultas y lineales en la capa de salida, se tuvo que escoger entre el tipo de funciones de activación tangencial o logística.

Selección de la función de activación.

Las alternativas que se tuvieron para escoger como función de activación en las capas ocultas fueron la función sigmooidal logística y la función sigmooidal tangencial. Para escoger entre ellas se muestra un ejemplo donde se les compara y donde se observo que la convergencia al error fue más rápida utilizando la función tangencial.

Número de entradas 21.

Número de salidas 1.

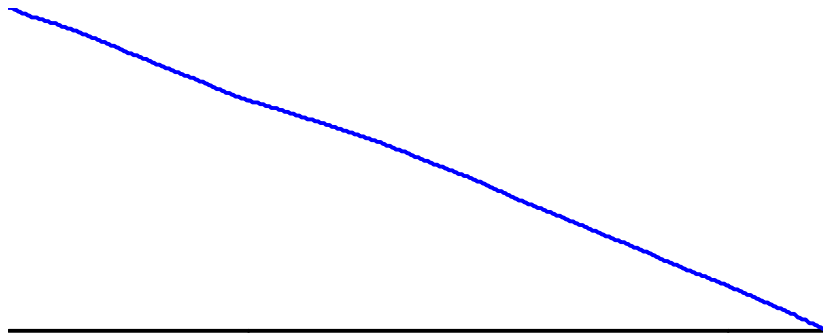
Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

Morfología de la red.

Se entrenó una red de 6 capas, con 5 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función sigmooidal logística capa de salida función lineal.

Algoritmo de aprendizaje. Gradiente conjugado.



Gráfica 5 Entrenamiento con función sigmooidal logística

El tiempo en que convergió del error 0.1 al error 0.01 fue 1909.125 segundos.
Utilizando una función tangencial en lugar de la logística.

Número de entradas 21.

Número de salidas 1.

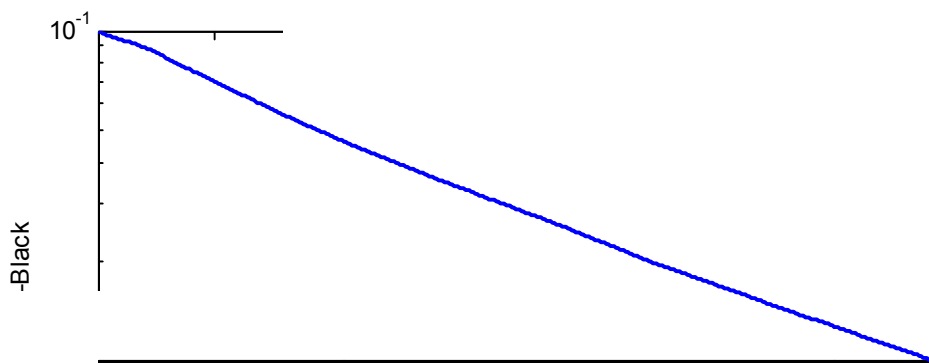
Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

Morfología de la red.

Se entrenó una red de 6 capas, con 5 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función sigmooidal tangencial capa de salida función lineal.

Algoritmo de aprendizaje. Gradiente conjugado.



Gráfica 6 Entrenamiento con función tangencial

El tiempo en que convergió del error 0.1 al error 0.01 fue 971.607 segundos.

Así como se muestra en el ejemplo anterior se observa que la convergencia al error fue más rápida utilizando la función tangencial como lo reporta Bishop (1995). Por lo cual se escogió la función tangencial

Para el entrenamiento de la red se debió escoger el algoritmo de optimización del ajuste de pesos y bias.

Selección del algoritmo optimización de entrenamiento.

Donde la red tiene que ser entrenada para aproximar una función (regresión no-lineal) se tuvo como alternativas de algoritmo de entrenamiento: Rprop (backpropagation resiliente) (Riedmiller, 1993), gradiente conjugado (Moller, 1993). Estas alternativas se seleccionaron en base a las características de las morfologías utilizadas para predecir el crecimiento del *Pinus Cooperi* Blanco. El problema demanda redes de un tamaño que satura la memoria de la computadora si se usaran métodos que requieren el calcular el Hessiano y no serían adecuadas.

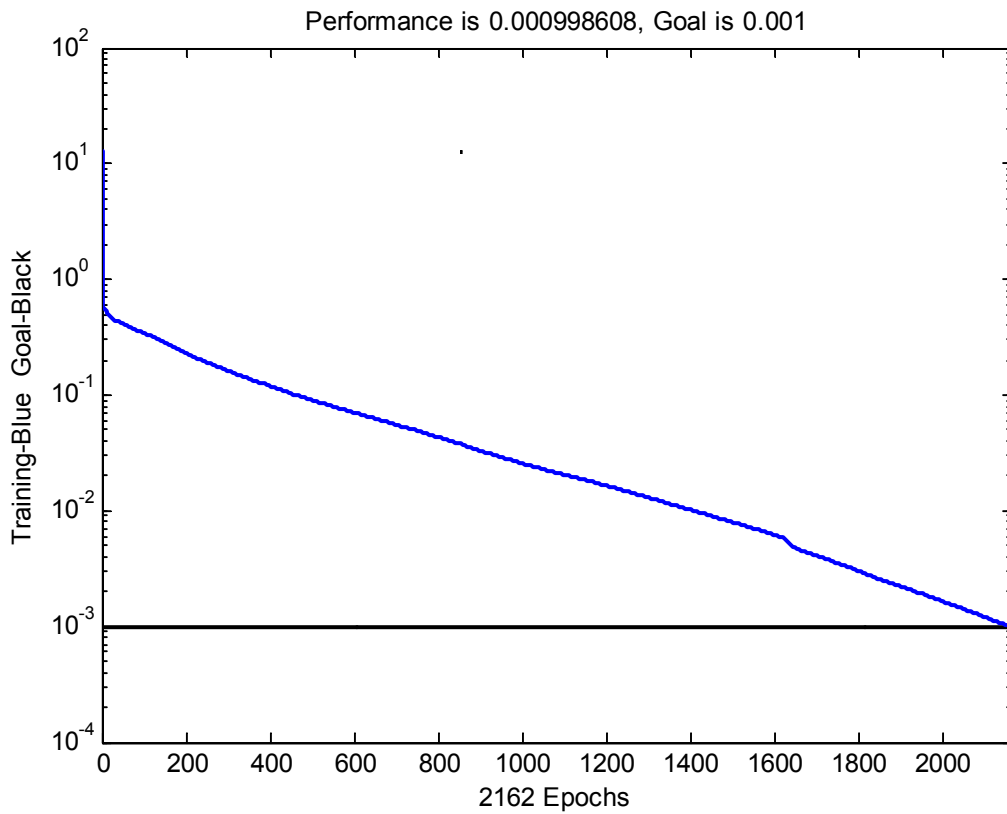
Así se comparó el desempeño de la red con cada una de las alternativas consideradas:

Morfología de la red.

Una capa oculta de 90 neuronas con función de activación tangencial y una capa de salida de una neurona con función de activación lineal.

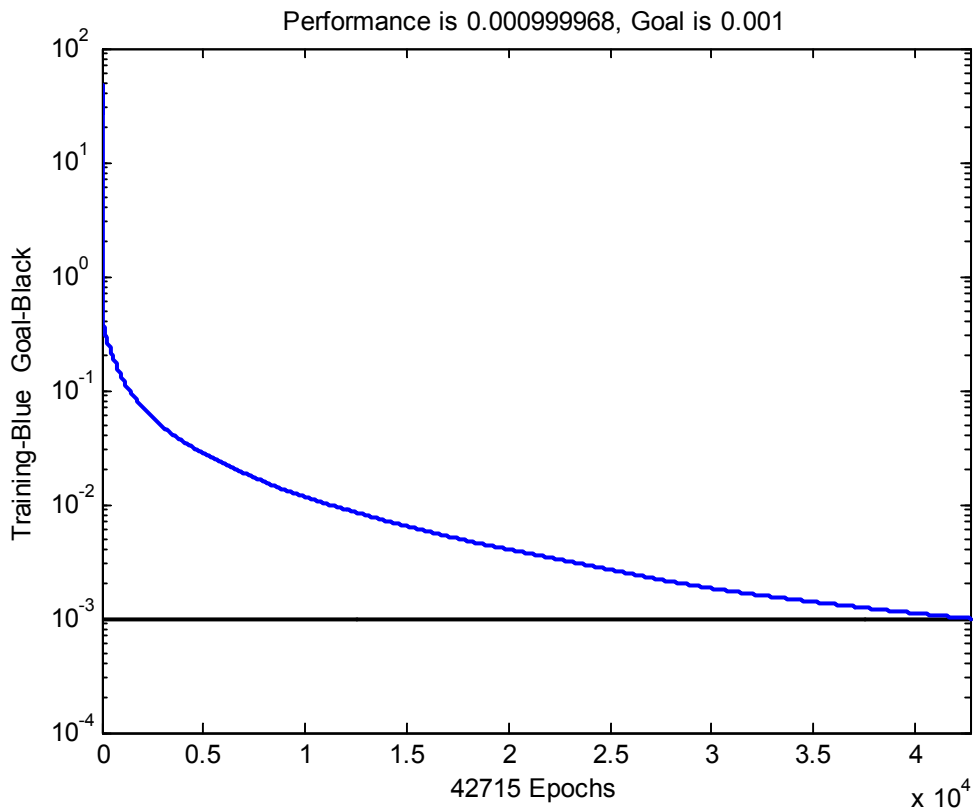
Algoritmo de entrenamiento: gradiente conjugado.

Resultado: consumió un tiempo de entrenamiento de 388.058 segundos para alcanzar un error cuadrático medio de 1×10^{-3} en 2162 épocas.



Gráfica 7 método gradiente conjugado.

Entrenando la misma red pero ahora con el algoritmo RPROP. El tiempo que se consumió para alcanzar un error medio cuadrático de 1×10^{-3} fue de 4326.782 segundos en 42 715 épocas.



Gráfica 8 Método RPROP

Por el menor tiempo en converger al error se optó por utilizar el algoritmo de gradiente conjugado escalado.

Ya normalizados los datos de entrenamiento se busco la mejor arquitectura de la red para predecir el crecimiento en diámetro y altura del *Pinus Cooperi* Blanco.

La inicialización de los pesos y bias se hizo de forma aleatoria.

Se decidió utilizar, para el entrenamiento de la red toda la información disponible pues se esta tratando con un sistema altamente complejo.

Para seleccionar la morfología que predijera de forma satisfactoria el crecimiento de la especie de pino el *Pinus Cooperi* Blanco, se optó por implementar diferentes morfologías a partir de una red con una capa oculta y una de salida e ir aumentando el tamaño de la red en el número de capas como el número de neuronas por capa.

La diferencia que se observo entre utilizar una red con una capa oculta con una cantidad alta de neuronas y distribuir esas mismas neuronas en varias capas se muestra en el siguiente ejemplo

Red neuronal dos capas; predicción de crecimiento en diámetro.

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento:

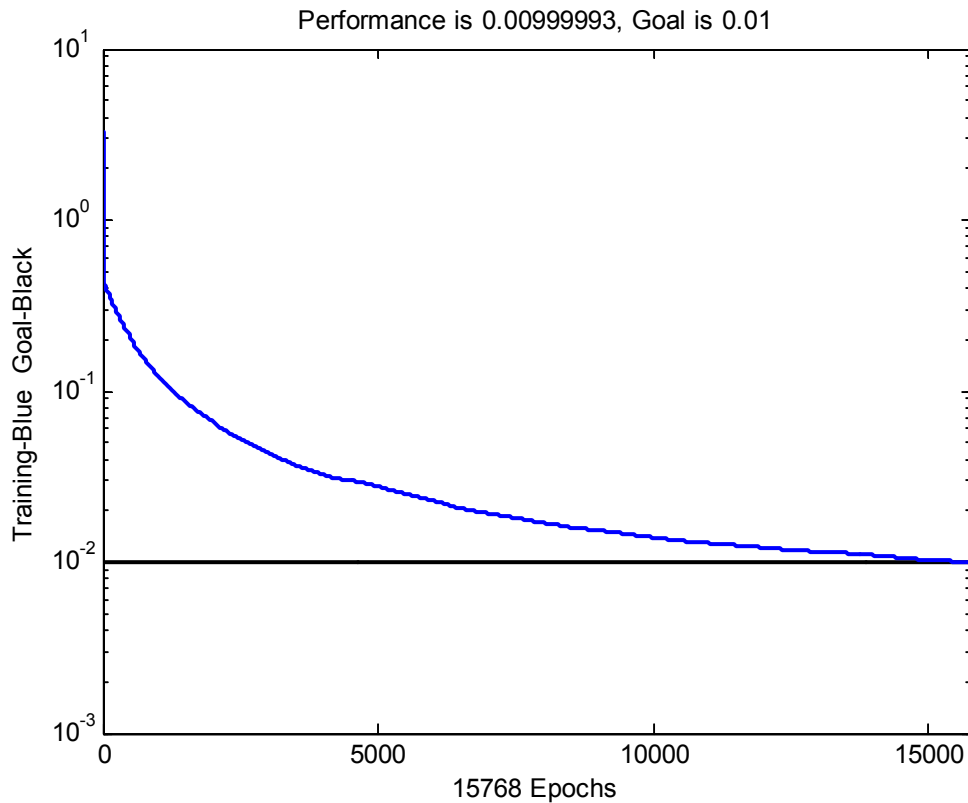
Media cero (0) desviación uno (1).

Morfología de la red.

Se entrenó una red de 2 capas, con 1 capa oculta con 90 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

Algoritmo de aprendizaje. Gradiente conjugado.



Gráfica 9 Red una capa oculta

El tiempo de entrenamiento para alcanzar el error de 0.01 igual a 6533.755 segundos.

Distribuyendo las 90 neuronas de la capa oculta en tres capas se tiene:

Red neuronal Cuatro capas; predicción de crecimiento en diámetro.

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento:

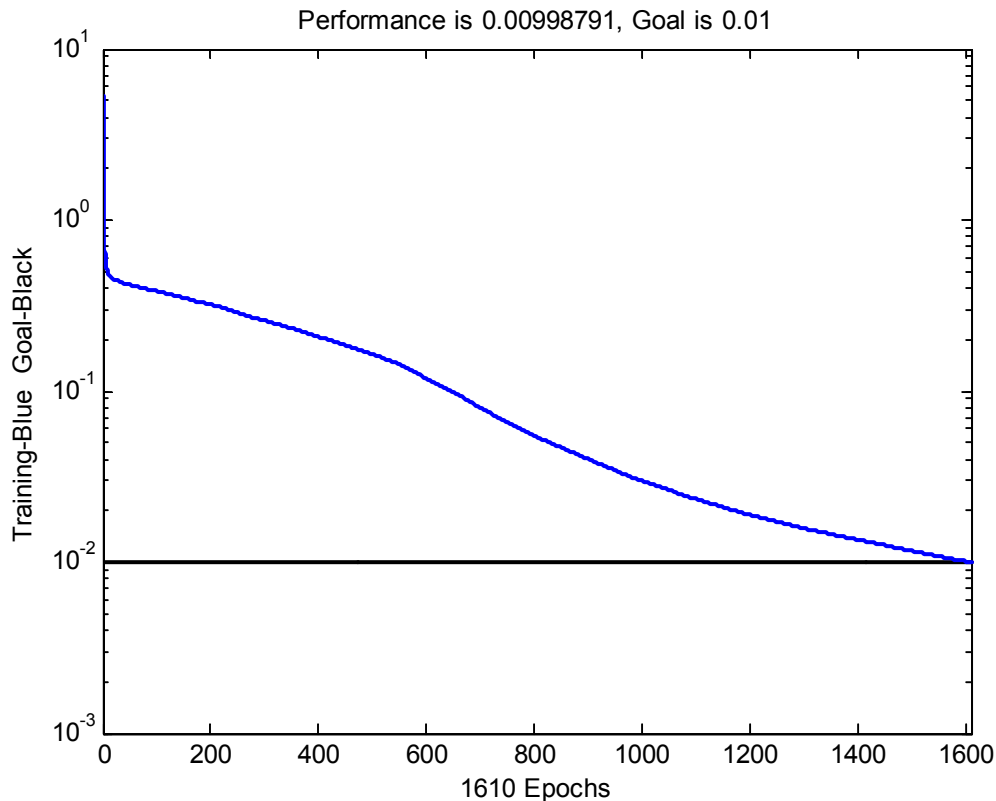
Media cero (0) desviación uno (1).

Morfología de la red.

Se entrenó una red de 4 capas, con 3 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

Algoritmo de aprendizaje. Gradiente conjugado.



Gráfica 10 Red tres capas ocultas

El tiempo de entrenamiento para alcanzar el error de 0.01 en 1610 épocas fue 662.853 segundos.

Por el ejemplo se puede concluir que una red neuronal con varias capas ocultas en lugar de solo una capa oculta aprende con mayor rapidez.

El entrenamiento de las redes en el desarrollo de este trabajo se detuvo a diferentes errores 0.1, 0.01, 0.001, 0.0001, 0.00001, como la capacidad de aprendizaje de la red lo permitió, con el fin de observar como la red fue teniendo la capacidad de generalizar.

No fue posible hacer uso de la estrategia detención temprana (early stopping) por las características de la información del crecimiento del bosque de *Pinus Cooperi* Blanco. Pues esta estrategia requiere de dos grupos de datos equivalentes uno para entrenar y otro para validar. Lo cual con la cantidad y forma en que esta distribuida la información no es posible, pues se tendrían que formar los dos grupos perdiendo una gran cantidad de información, siendo aún más difícil predecir el crecimiento del bosque.

2.2.5.4. Evaluación de la capacidad de generalización de la ANN.

Ya entrenada la red se verificó su capacidad de predicción con una muestra de validación

Selección de la muestra de validación. La selección de la muestra que se escogió buscó reflejar el comportamiento del sistema, para lo cual en este trabajo la selección se realizó utilizando el conocimiento de que se tendrá un mayor crecimiento del árbol a magnitudes pequeñas de índices de competencia y donde es de esperarse que a valores altos de competencia el árbol tenga un menor crecimiento.

Aunado a lo anterior se consideraron en la muestra datos que aparecían en los rangos de crecimiento de mayor frecuencia.

Otros factores también considerados fueron la riqueza de nutrientes y agua de la parcela, grupo de edad y la velocidad de crecimiento.

Así se seleccionó una muestra de validación 21 individuos para la predicción de crecimiento en diámetro y de 30 para el crecimiento en altura.

2.3. RESULTADOS.

2.3.1. Parámetros utilizados.

Se utilizó un total de 2195 árboles de los cuales se tomó una muestra de 21 y 30 individuos para validar los resultados de la red en el crecimiento en diámetro y altura.

Se utilizaron como parámetros:

- 1) Índice de sitio en metros por cuadrante.
- 2) Área basal por cuadrante en m^2 / ha.
- 3) Incremento en área basal por cuadrante m^2 /ha.
- 4) Diámetro cuadrático por cuadrante en centímetros.
- 5) Edad media por cuadrante en años.
- 6) Diámetro del árbol a 0.3 metros de la base del suelo en centímetros.
- 7) Diámetro normal del árbol a 1.3 metros sobre la base del suelo en cm.
- 8) Altura total del árbol en metros.
- 9) Índice de área basal de Glover y Hool del árbol “i”
- 10) Índice de competencia de área basal número de árboles.
- 11) Índice de altura media del árbol “i” valor absoluto.
- 12) Índice de altura media número de árboles.
- 13) Índice de diámetro cuadrático del árbol “i”
- 14) Índice de diámetro cuadrático número de árboles.
- 15) Índice de longitud de copa del árbol “i”
- 16) Índice de longitud de copa número de árboles.
- 17) Índice de amplitud de copa del árbol “i”
- 18) Índice de amplitud de copa número de árboles.
- 19) Índice de competencia de área basal total del árbol “i”
- 20) Índice de sitio polimórfico 2 en metros por sitio experimental
- 21) Relación de diámetros 0.3/1.3 m del árbol,

2.3.2. Estimación de la velocidad de crecimiento en diámetro con una Red Neuronal Backpropagation.

Se tomó una muestra de 21 individuos del total de la población de 2195 para validar la capacidad de generalización de la red neuronal buscada.

Se procedió a seleccionar la morfología de red que pudiera predecir la velocidad de crecimiento en diámetro dentro de un intervalo de confianza aceptable.

Se buscó la morfología. Para lo cual se comparó el funcionamiento de diversas arquitecturas seleccionándose la que se encontró más adecuada.

2.3.2.1. Predicción de crecimiento en Diámetro

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento:

Media cero (0) desviación uno (1).

Morfología de la red.

Se entrenó una red de dos capas, con una capa oculta con 90 neuronas y una capa de salida con una neurona.

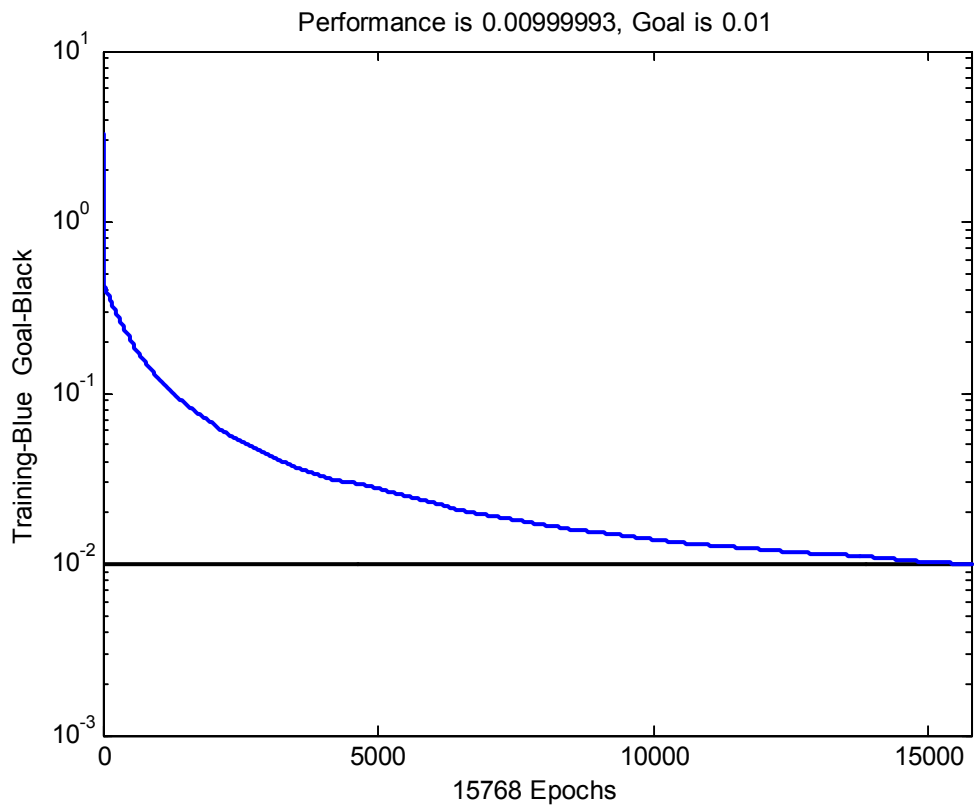
Función de activación. Capa oculta función tangencial capa de salida función lineal.

Algoritmo de aprendizaje. Gradiente conjugado.

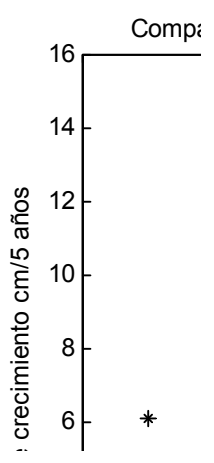
Tratando de analizar como la red va ajustando los valores de la muestra de validación, se fue entrenando la red en etapas con los siguientes resultados:

Se entrenó la red durante 15768 épocas hasta alcanzar un error cuadrático medio de 0.01

Al observar los resultados obtenidos (gráfica 12) con diferentes errores de entrenamiento se desechó esta morfología y se experimento con otra.



Gráfica 11 Entrenamiento de la red (ejemplo 2.3.2.1.)



Gráfica 12 predicción de crecimiento en diámetro (ejemplo 2.3.2.1.)

2.3.2.2. Predicción de crecimiento en Diámetro

Número de entradas 21.

Número de salidas 1.

Tratamiento a los patrones de entrenamiento. Media cero (0) desviación uno (1).

Morfología de la red.

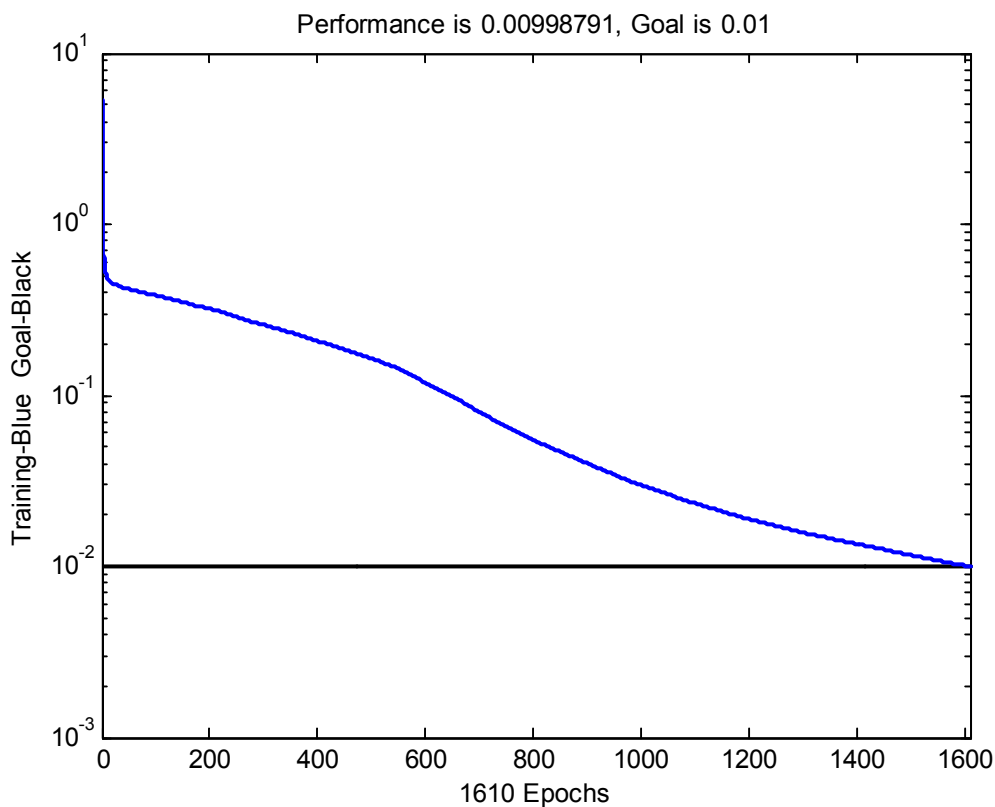
Se entrenó una red de 4 capas, con 3 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

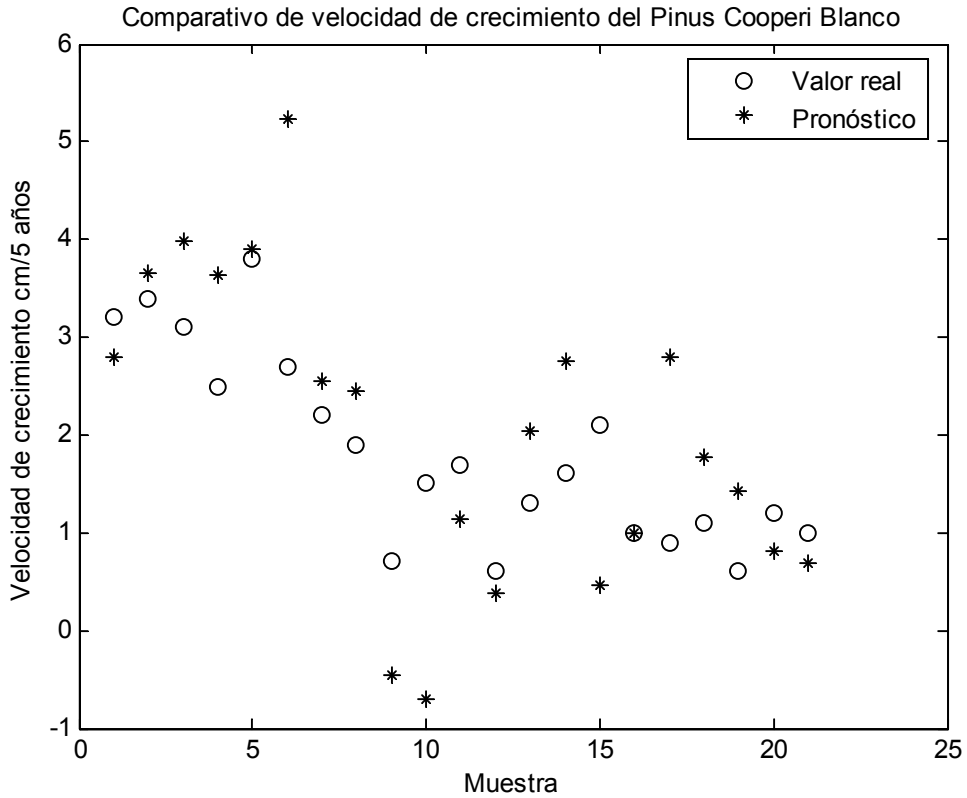
Algoritmo de aprendizaje. Gradiente conjugado.

Tratando de analizar como la red va ajustando los valores de la muestra de validación, se fue entrenando la red en etapas con los siguientes resultados:

Se entrenó la red durante 2408 épocas hasta alcanzar un error cuadrático medio de 0.01



Gráfica 13 Entrenamiento de la red (ejemplo 2.3.2.2.)



Gráfica 14 predicción de crecimiento en diámetro (ejemplo 2.3.2.2.)

La red tuvo la capacidad de aproximar algunos valores de la muestra sin embargo otros fueron totalmente equivocados. La opción es agregar mas capas a la red.

Sin embargo se puede observar la mayor capacidad de pronóstico entre la red de cuatro capas con respecto a la de dos capas, aunque claro la red de cuatro capas tiene un mayor número de pesos.

2.3.2.3. Predicción de crecimiento en Diámetro

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

Morfología de la red.

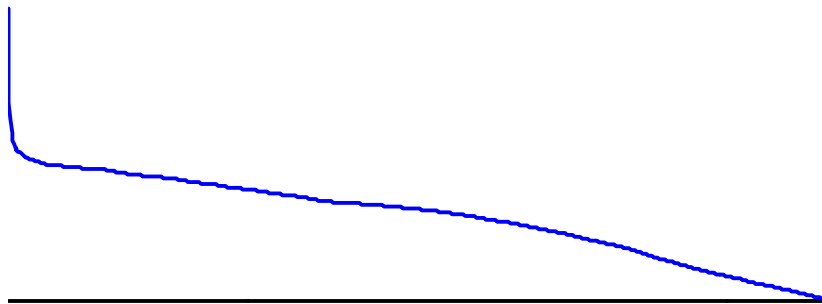
Se entrenó una red de 6 capas, con 5 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

Algoritmo de aprendizaje. Gradiente conjugado.

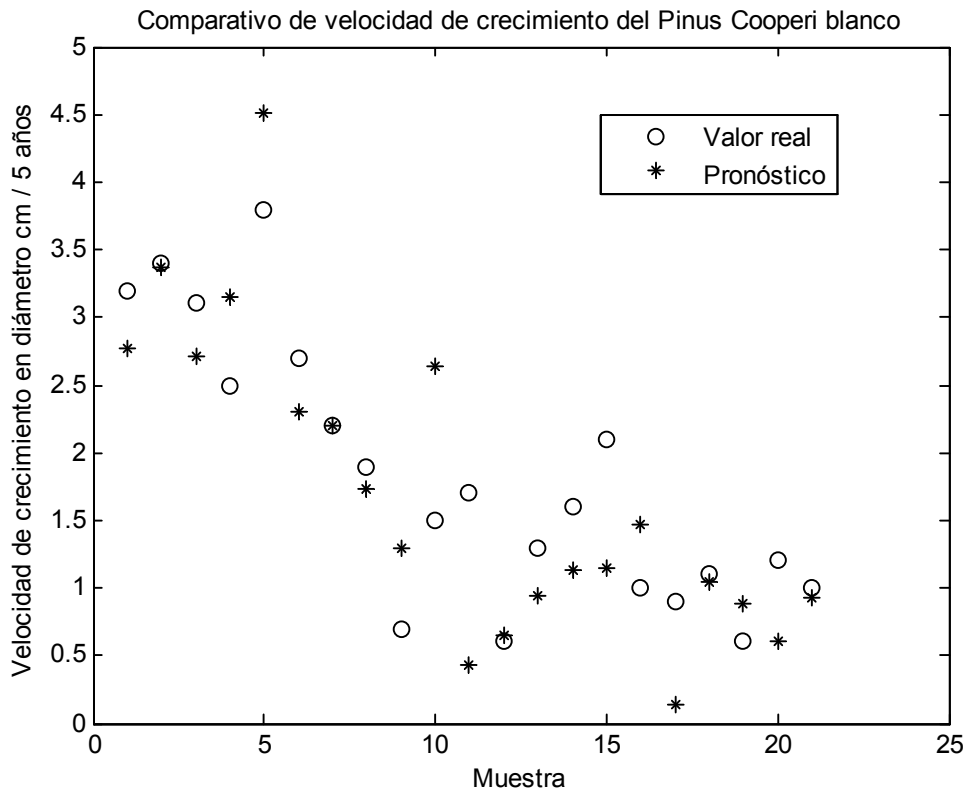
Tratando de analizar como la red va ajustando los valores de la muestra de validación, se fue entrenando la red en etapas con los siguientes resultados:

Se entrenó la red durante 957 épocas hasta alcanzar un error cuadrático medio de 0.1



Gráfica 15 Entrenamiento de la red (ejemplo 2.3.2.3)

:



Gráfica 16 predicción de crecimiento en (diámetro ejemplo 2.3.2.3)

Como se puede observar los resultados no fueron los deseados por lo que se prosiguió entrenando la red en una siguiente etapa.

2.3.2.4. Predicción de crecimiento en Diámetro.

Entrenamiento hasta un error de 0.01

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

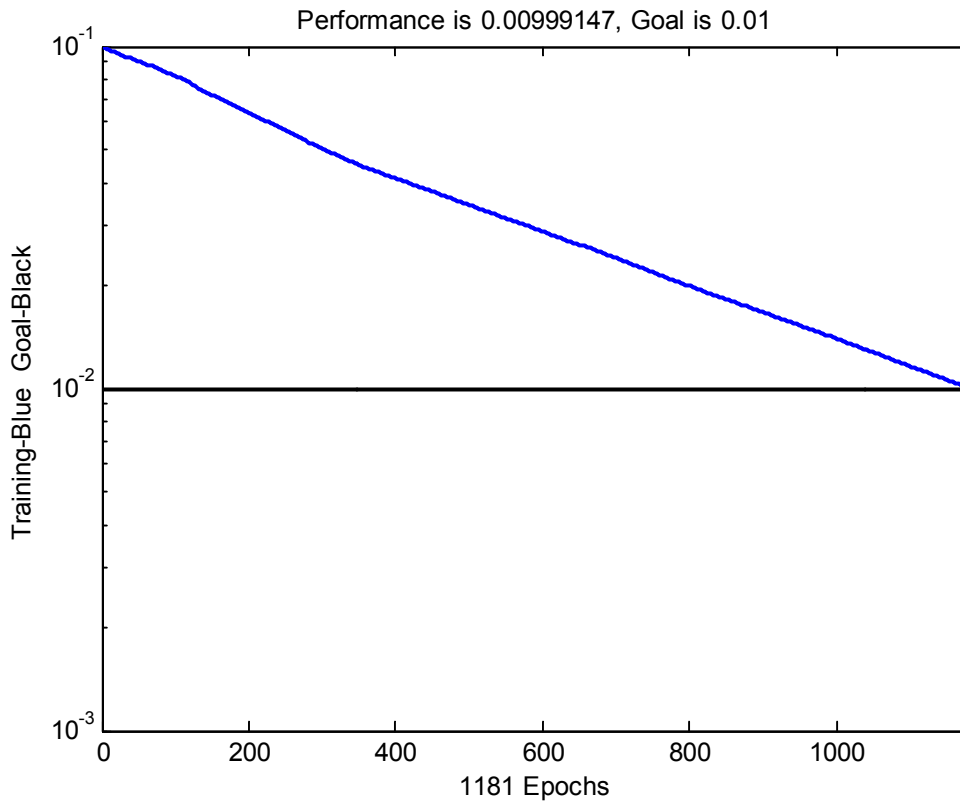
Morfología de la red.

Se entrenó una red de 6 capas, con 5 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

Algoritmo de aprendizaje. Gradiente conjugado.

Se entrenó la red hasta alcanzar un error medio cuadrático de 0.01 durante 1181 épocas.



Gráfica 17 Entrenamiento de la red (ejemplo 2.3.2.4.)

Gráfica 18 predicción de crecimiento en diámetro (ejemplo 2.3.2.4.)

Como se puede observar los resultados obtenidos son excelentes, la red **aproximó los valores** de la muestra de validación en un **90%** de aproximación al error promedio dentro de un intervalo de confianza del **90%**.

2.3.2.5. Predicción de crecimiento en Diámetro.

Entrenamiento hasta un error de 0.001

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

Morfología de la red:

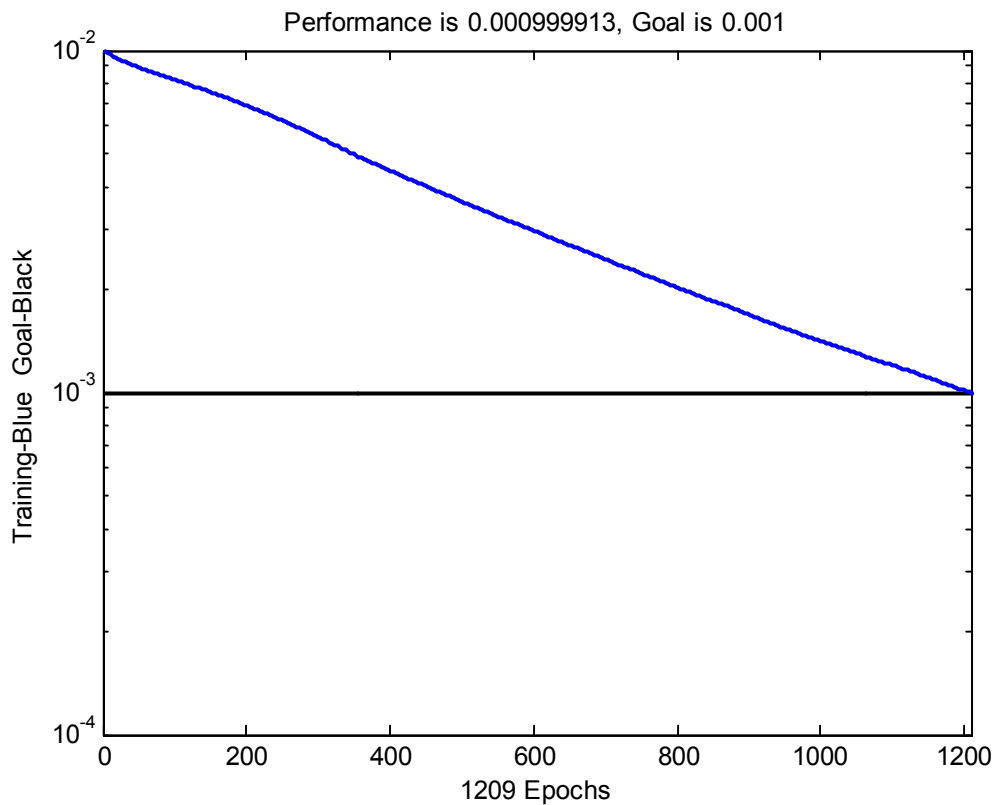
Se entrenó una red de 6 capas, con 5 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

Algoritmo de aprendizaje. Gradiente conjugado.

Al continuar entrenando la red para ver si se puede mejorar su aproximación a los valores de la muestra se obtienen los siguientes resultados:

Se entrena la red durante 1209 épocas hasta alcanzar un error cuadrático medio de 0.001



Gráfica 19 Entrenamiento de la red ejemplo (2.3.2.5.)

Gráfica 20 predicción de crecimiento en diámetro (ejemplo 2.3.2.5.)

Al seguir entrenando la red durante 1209 épocas hasta alcanzar un error medio cuadrático de 0.001, la aproximación del pronóstico se degrada como se puede observar en el resultado.

2.3.2.6. Predicción de crecimiento en Diámetro.

Se aumenta a 11 capas la arquitectura de la red, para observar el efecto en la calidad del pronóstico en una red más grande

Entrenamiento hasta un error de 0.01

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

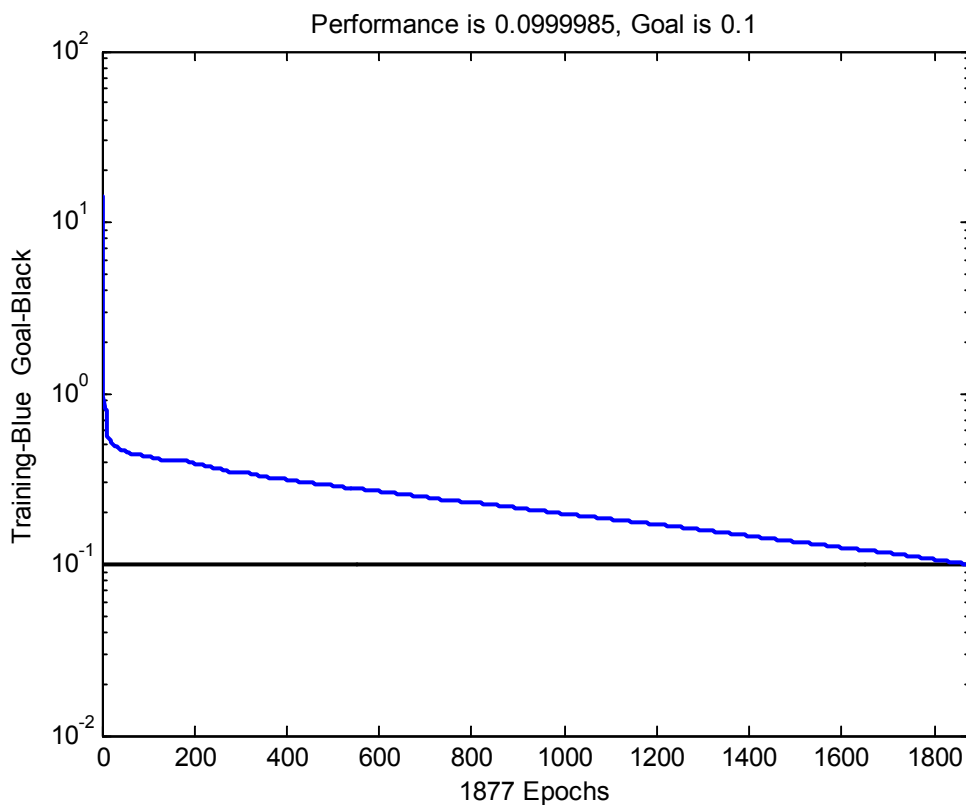
Morfología de la red:

Se entrenó una red de 11 capas, con 10 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

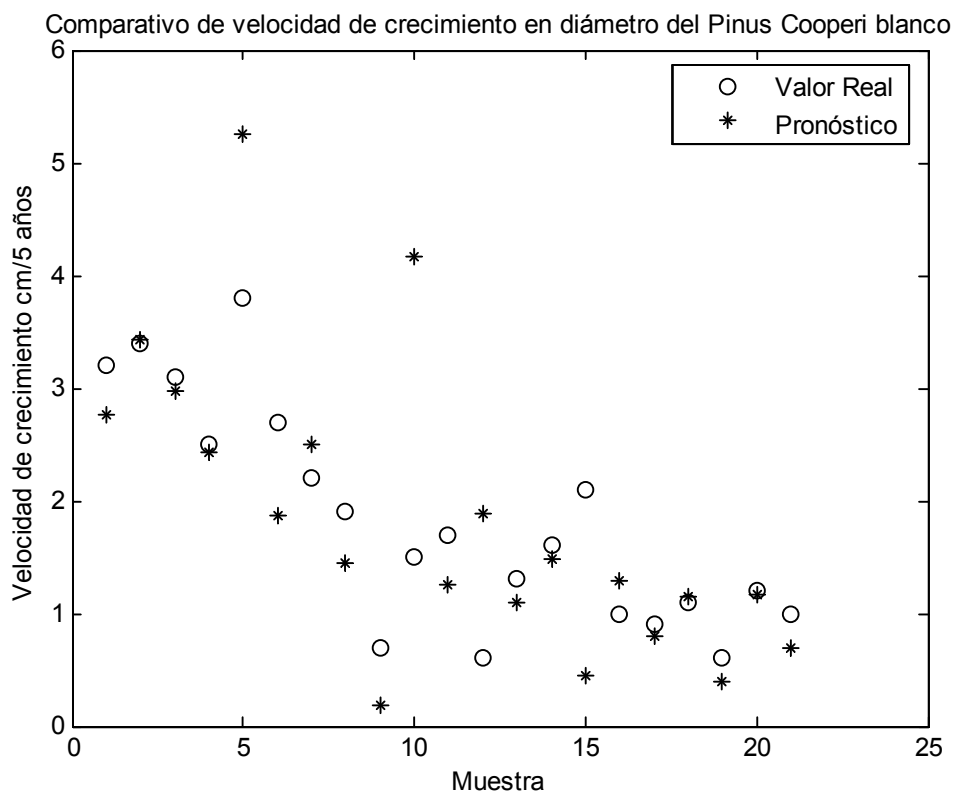
Algoritmo de aprendizaje. Gradiente conjugado.

Se entrenó la red hasta alcanzar un error cuadrático medio en el entrenamiento de 0.1, Durante 1877 épocas.



Gráfica 21 Entrenamiento de la red (ejemplo 2.3.2.6.)

Al validar el poder de predicción de la red con la muestra de 21 datos se producen los siguientes resultados:



Gráfica 22 predicción de crecimiento en diámetro (ejemplo 2.3.2.6.)

Al analizar los resultados podemos concluir que el entrenamiento no ha sido el suficiente para que la red sea capaz de predecir con una precisión significativa.

2.3.2.7. Predicción de crecimiento en Diámetro.

Así se continúa entrenando a la red hasta un error cuadrático medio de 0.01

Entrenamiento hasta un error de 0.01

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

Morfología de la red:

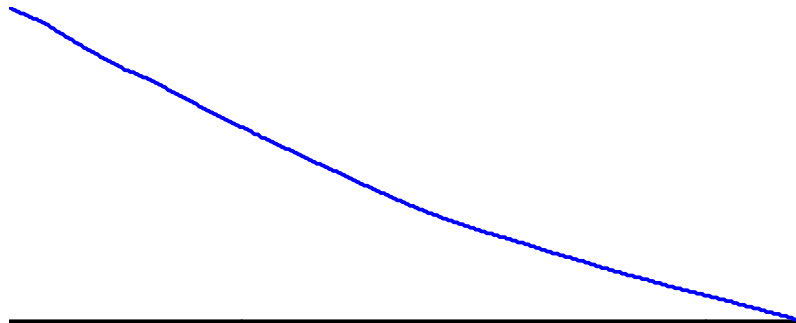
Se entrenó una red de 11 capas, con 10 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

Algoritmo de aprendizaje. Gradiente conjugado.

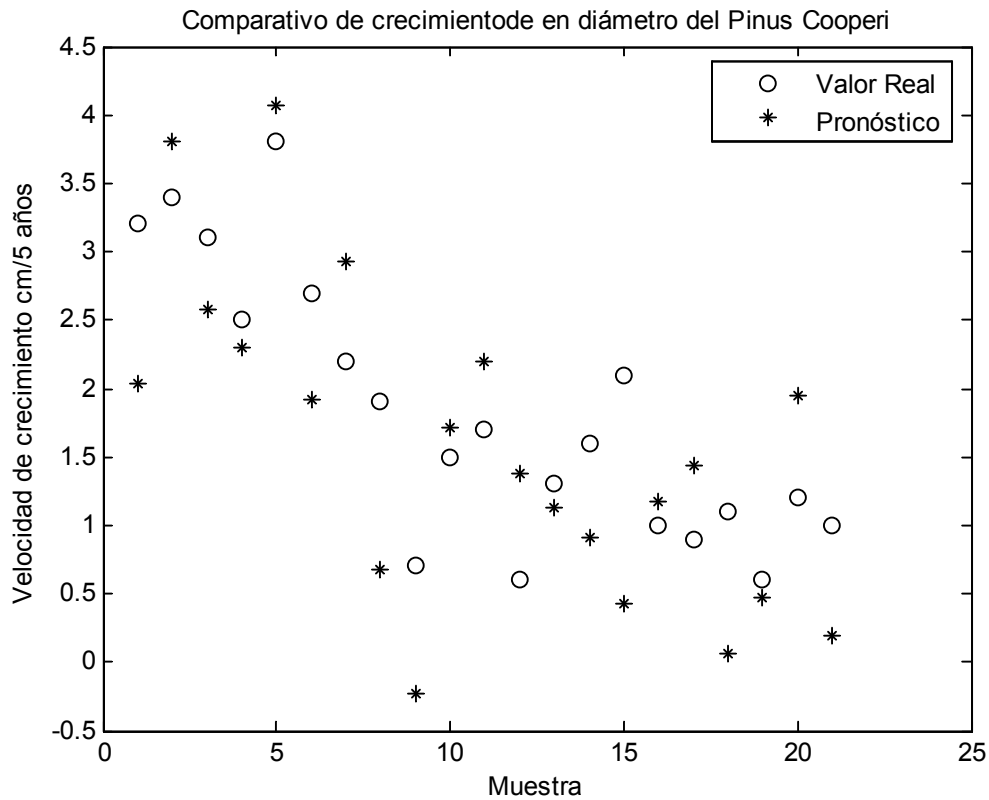
Se entrenó la red hasta alcanzar un error cuadrático medio en el entrenamiento de 0.1, Durante 1877 épocas.

El error se alcanza después de entrenar la red durante 3421 épocas.



Gráfica 23 Entrenamiento de la red (ejemplo 2.3.2.7.)

Una vez se alcanza el error fijado y detener el entrenamiento se analiza la capacidad de predicción de la red y se observan los siguientes resultados:



Gráfica 24 predicción de crecimiento en diámetro (ejemplo 2.3.2.7.)

La capacidad de predicción de la red no mejora por lo que en base a estos resultados y por más entrenamiento que se le aplicó a la red se encontró que no es la morfología adecuada para predecir el crecimiento buscado.

2.3.3. Estimación de la velocidad de crecimiento en altura por medio de una red neuronal backpropagation.

Se tomó una muestra de 30 individuos del total de la población de 2195 para validar la capacidad de predicción de la red neuronal buscada. Y se procedió a seleccionar la arquitectura de red que pudiera predecir la velocidad de crecimiento en altura de forma satisfactoria.

Se buscó la arquitectura. Para lo cual se comparó el funcionamiento de diversas arquitecturas seleccionándose la que se encontró más adecuada. Como ejemplo se muestra los siguientes resultados.

2.3.3.1. Predicción de crecimiento en altura

Número de entradas 21.

Número de salidas 1.

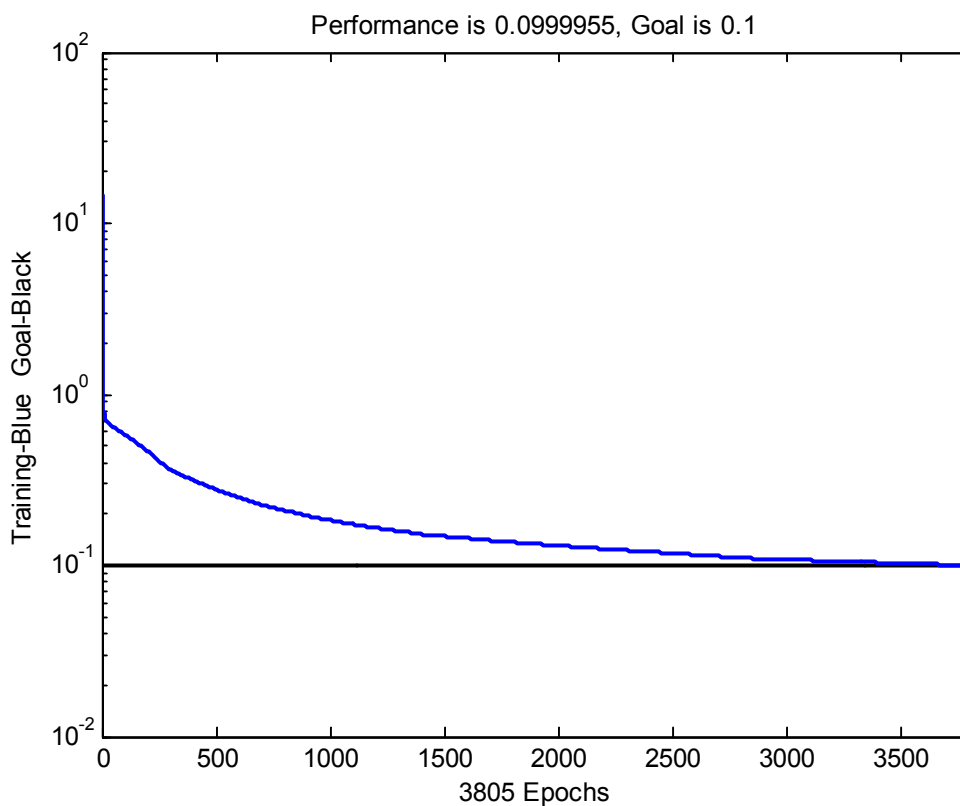
Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

Morfología de la red.

Se entrenó una red de 2 capas, con 1 capas ocultas con 60 neuronas y una capa de salida con una neurona.

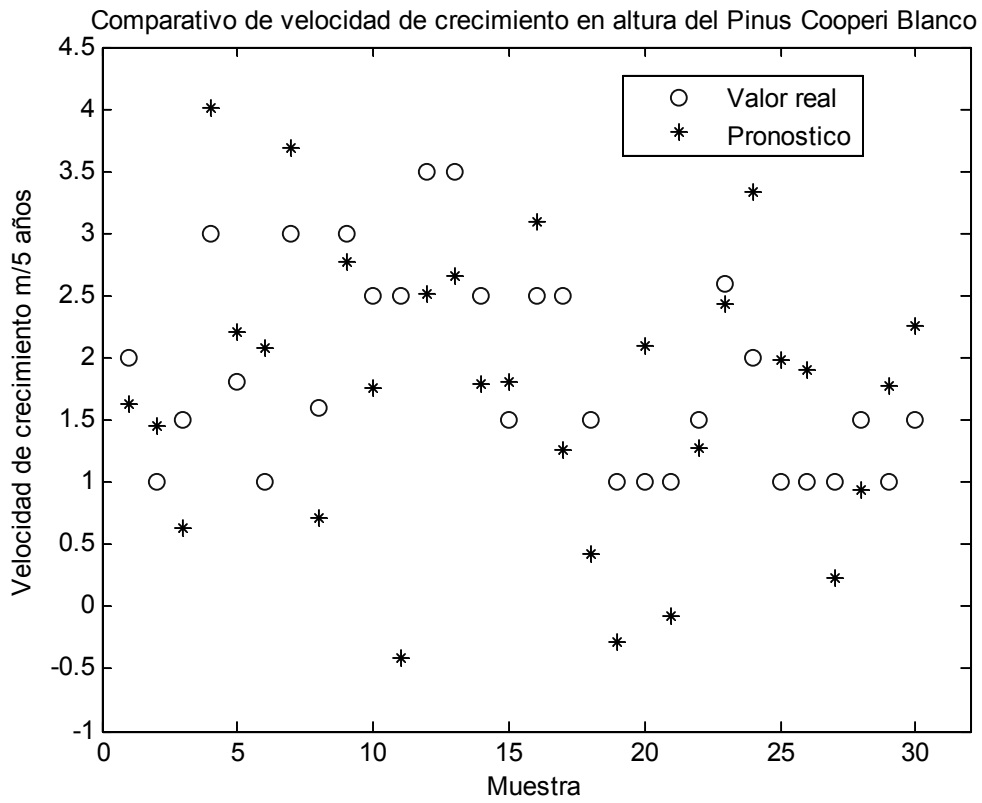
Función de activación. Capa oculta función tangencial capa de salida función lineal.

Algoritmo de aprendizaje. Gradiente conjugado.



Gráfica 25 Entrenamiento de la red (ejemplo 2.3.3.1.)

Se entrenó la red con hasta alcanzar un error medio cuadrático de 0.1 durante 3805 épocas (gráfica 25).



Gráfica 26 predicción de crecimiento en altura (ejemplo 2.3.3.1.)

Los resultados **no** fueron satisfactorios como se muestran en la gráfica 26 y se siguieron implementando otras arquitecturas.

2.3.3.2. Predicción de crecimiento en altura.

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

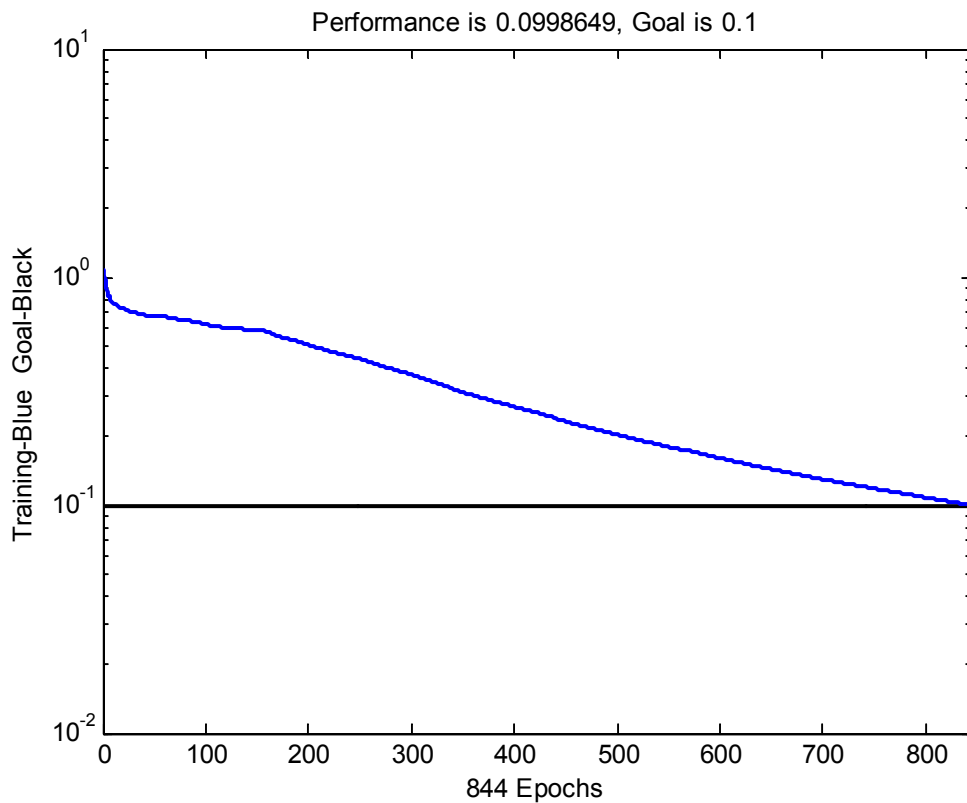
Morfología de la red.

Se entrenó una red de 4 capas, con 3 capas ocultas con 30 neuronas y 1 capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal

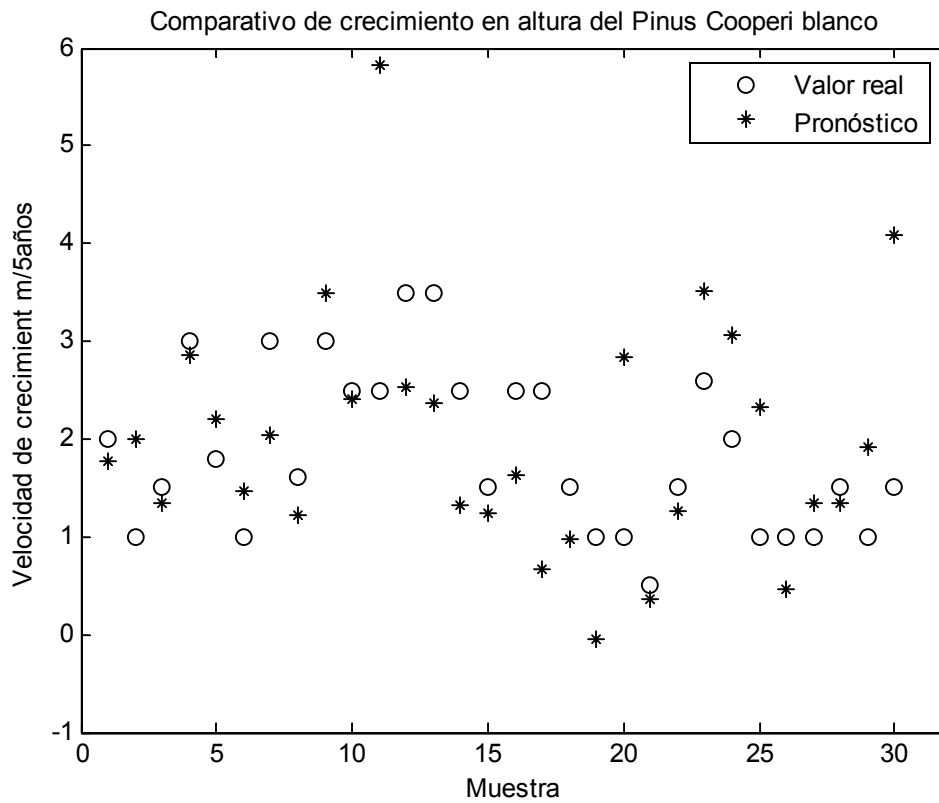
Algoritmo de aprendizaje. Gradiente conjugado.

Se entrenó la red con hasta alcanzar un error medio cuadrático de 0.1 durante 844 épocas.



Gráfica 27 Entrenamiento de la red (ejemplo 2.3.3.2.)

Los resultados que se obtuvieron son:



Gráfica 28 predicción de crecimiento en altura (ejemplo 2.3.3.2.)

Se observa cierta capacidad de la red a predecir por lo que se continua el entrenamiento

2.3.3.3. Predicción de crecimiento en altura.

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

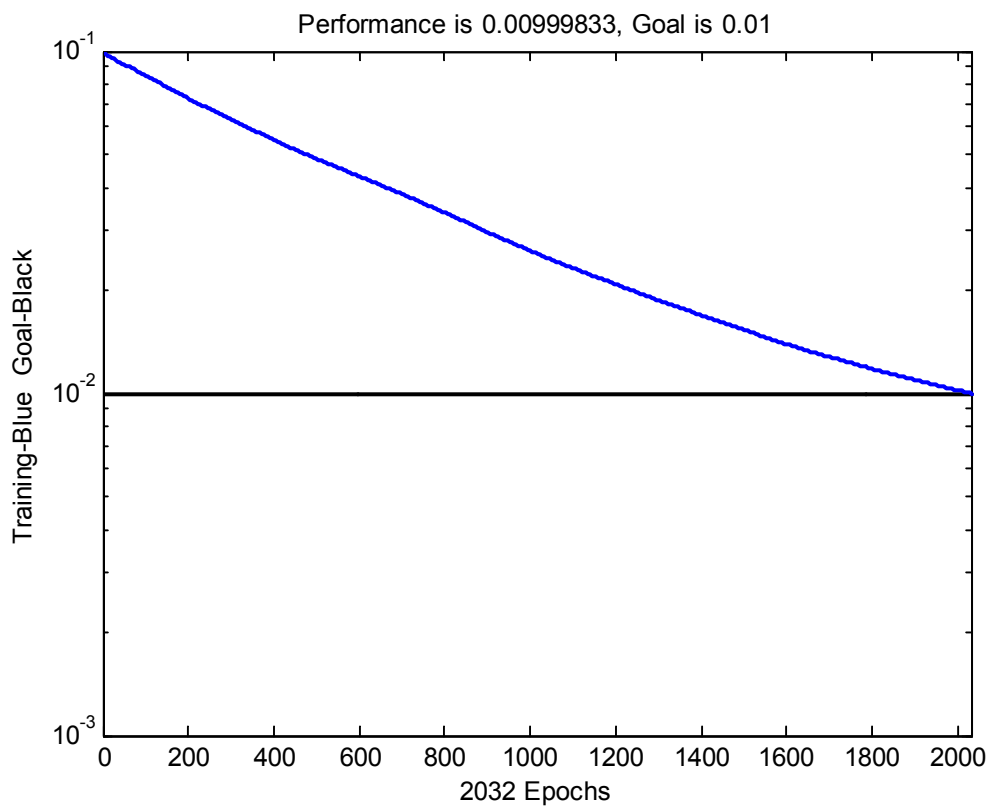
Morfología de la red.

Se entrenó una red de 4 capas, con 3 capas ocultas con 30 neuronas y 1 capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal .

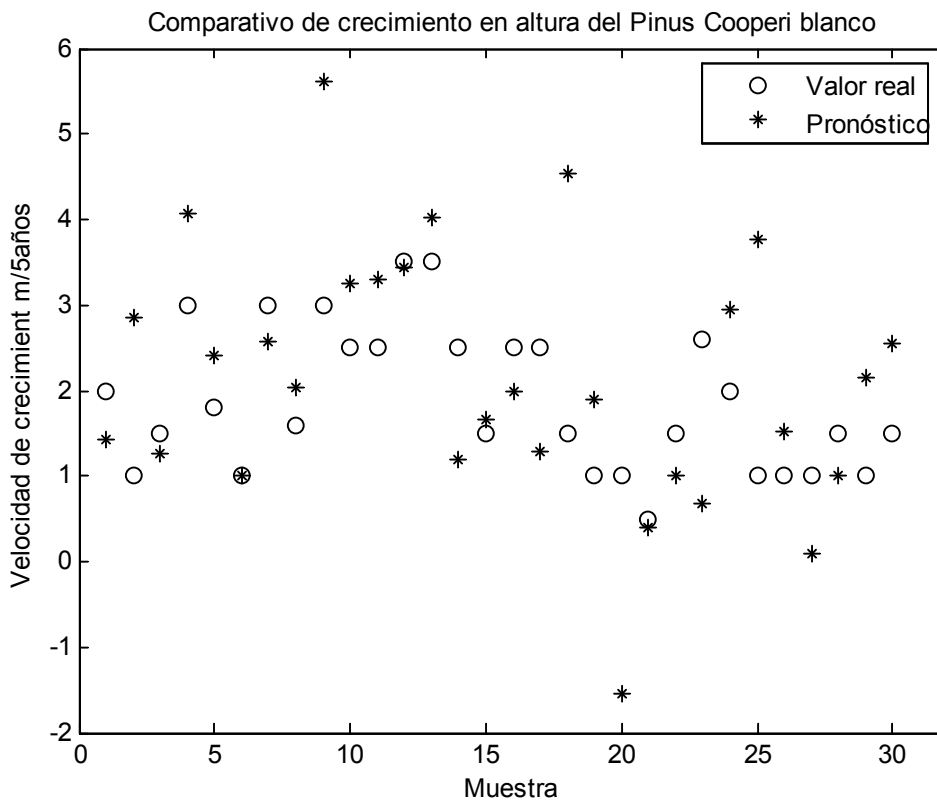
Algoritmo de aprendizaje. Gradiente conjugado.

Se entrenó la red con hasta alcanzar un error medio cuadrático de 0.01 durante 2032 épocas.



Gráfica 29 Entrenamiento de la red (ejemplo 2.3.3.3.)

Los resultados obtenidos son:



Gráfica 30 predicción de crecimiento en altura (ejemplo 2.3.3.3.)

La capacidad de predicción de la red empieza a degradarse en lugar de mejorar. Así habrá de aumentarse el tamaño de la red agregando más capas del mismo tamaño buscando aumentar su capacidad de que sea capaz de reflejar las relaciones entre parámetros.

2.3.3.4. Predicción de crecimiento en altura.

Se siguieron probando diversas arquitecturas hasta que se implemento una red de 6 capas.

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

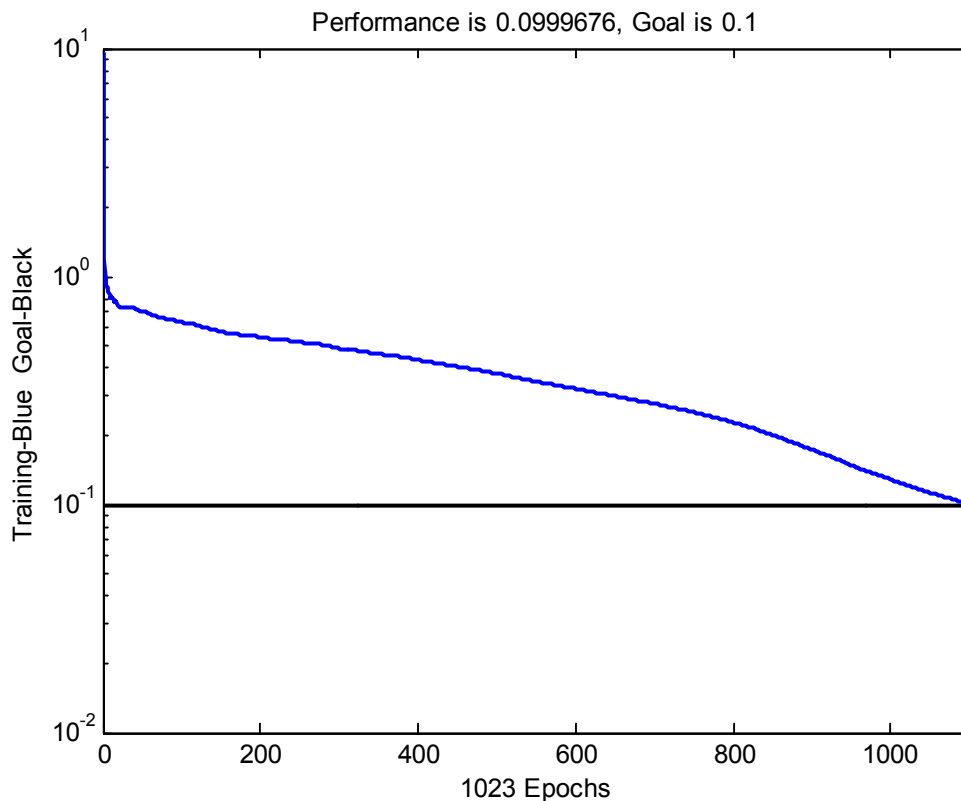
Morfología de la red.

Se entrenó una red de 6 capas, con 5 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

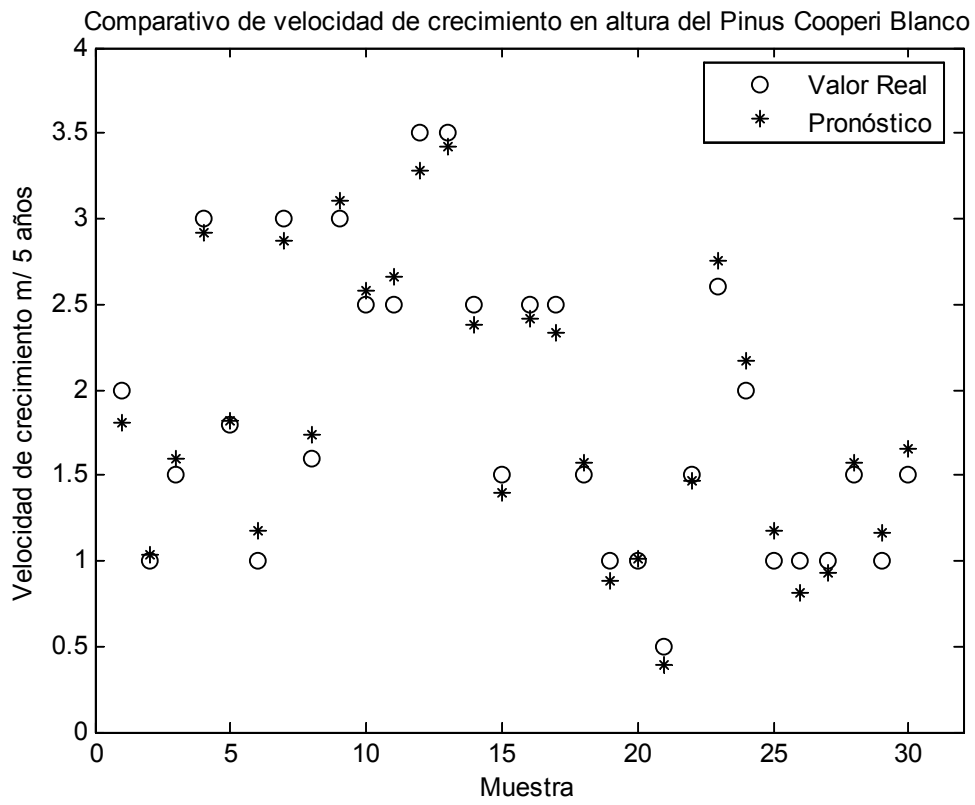
Algoritmo de aprendizaje. Gradiente conjugado.

Se entrenó la red durante 1023 épocas hasta alcanzar un error cuadrático medio de 0.1



Gráfica 31 Entrenamiento de la red (ejemplo 2.3.3.4.)

La red pronostico con excelentes resultados. Obteniendo un **95.5%** de aproximación promedio, para un intervalo de confianza del **90%** como se muestra en la siguiente gráfica:



Gráfica 32 predicción de crecimiento en altura (ejemplo 2.3.3.4.)

2.3.3.5. Predicción de crecimiento en altura.

Se sigue el entrenamiento de la red buscando mejorar el pronóstico.

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

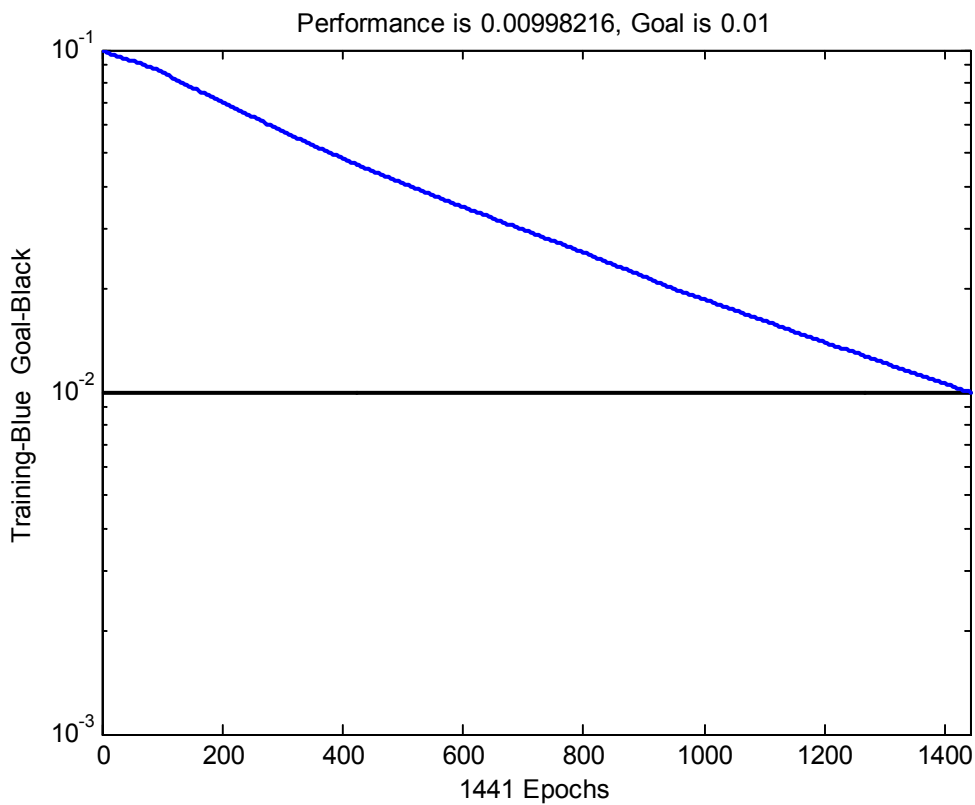
Morfología de la red.

Se entrenó una red de 6 capas, con 5 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

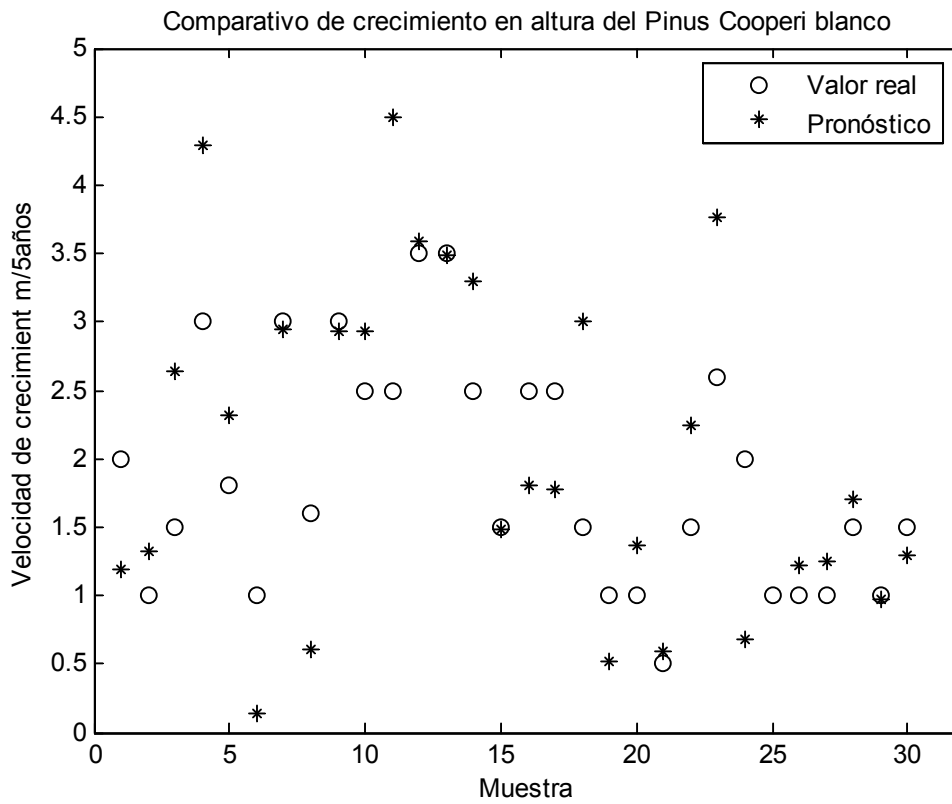
Algoritmo de aprendizaje. Gradiente conjugado.

Así entrenando la red por otras 1441 épocas más, hasta alcanzar un error cuadrático medio de 0.01



Gráfica 33 Entrenamiento de la red (ejemplo 2.3.3.5.)

Se producen los siguientes resultados:



Gráfica 34 predicción de crecimiento en altura (ejemplo 2.3.3.5.)

Los resultados muestran una pérdida de generalización de la red.

2.3.3.6. Predicción de crecimiento en altura.

Se sigue el entrenamiento de la red buscando ver si la tendencia es una pérdida de la generalización de la red.

Número de entradas 21.

Número de salidas 1.

Pre-procesamiento de los patrones de entrenamiento. Media cero (0) desviación uno (1).

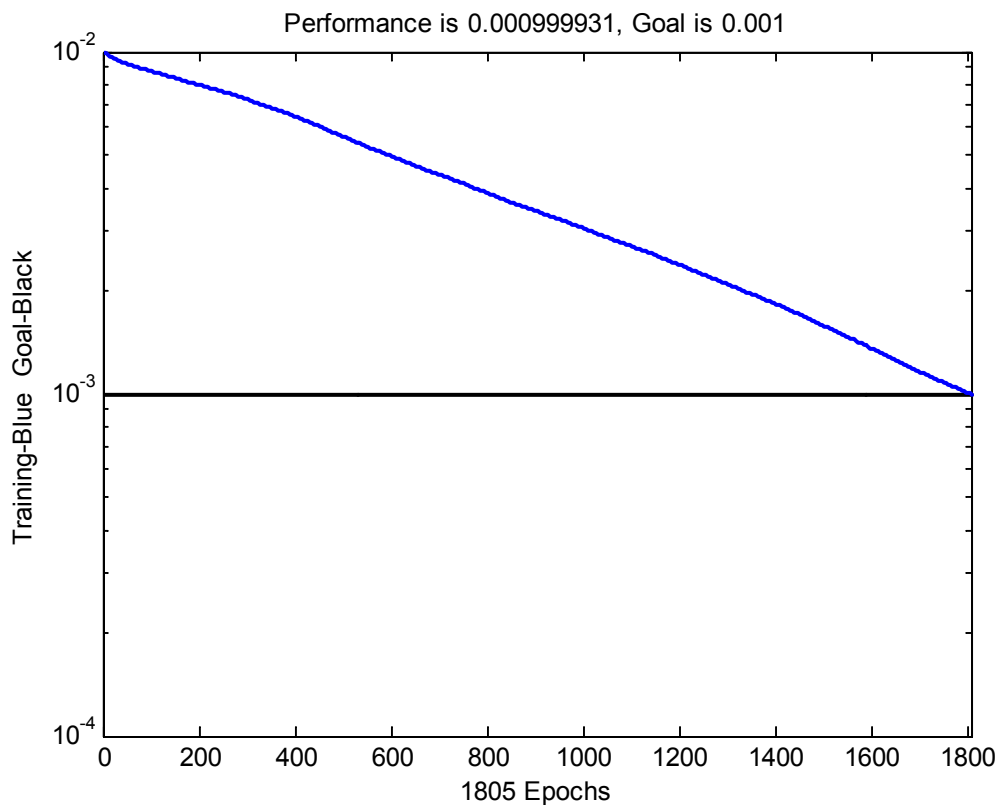
Morfología de la red.

Se entrenó una red de 6 capas, con 5 capas ocultas con 30 neuronas y una capa de salida con una neurona.

Función de activación. Capa oculta función tangencial capa de salida función lineal.

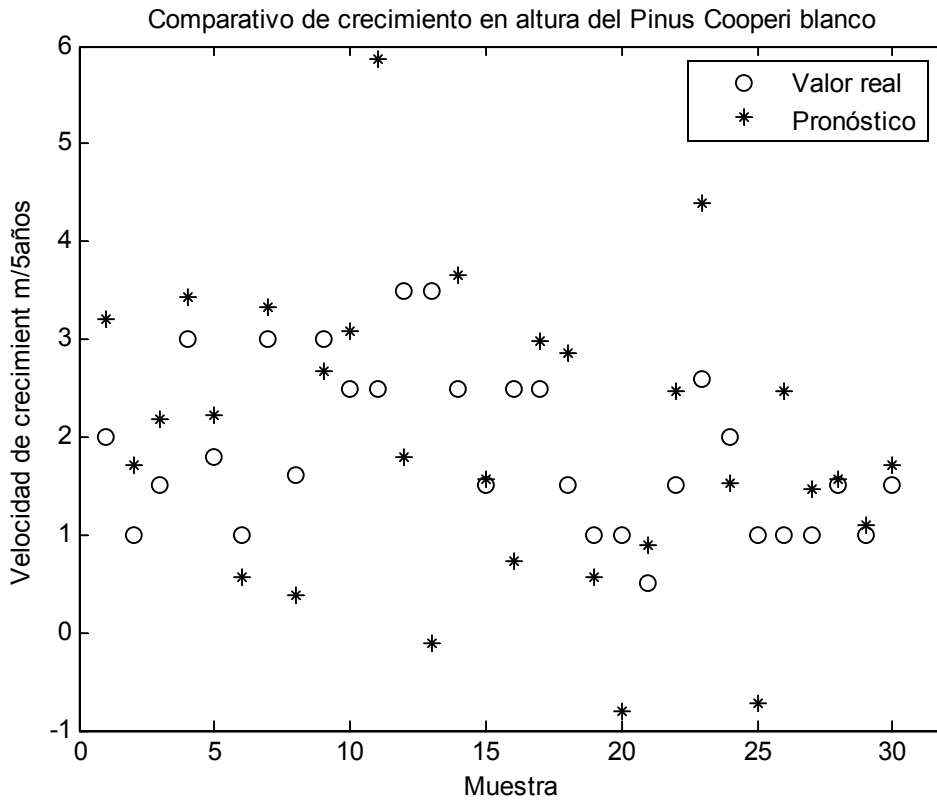
Algoritmo de aprendizaje. Gradiente conjugado.

Así entrenando la red por 1805 épocas más, hasta alcanzar un error cuadrático medio de 0.001



Gráfica 35 Entrenamiento de la red (ejemplo 2.3.3.6.)

Los resultados obtenidos se muestran a continuación:



Gráfica 36 predicción de crecimiento en altura (ejemplo 2.3.3.6.)

Como se puede observar la capacidad de predicción de la red decrece conforme el error de entrenamiento disminuye.

2.3.4. Estimación de la velocidad de crecimiento en diámetro por medio de una red neuro-difusa (ANFIS).

A pesar de los problemas que se presentan por la demanda en la capacidad de memoria se analiza el uso de una red neuro-difusa (ANFIS) en la estimación del crecimiento en diámetro de la especie *Pinus Cooperi* Blanco.

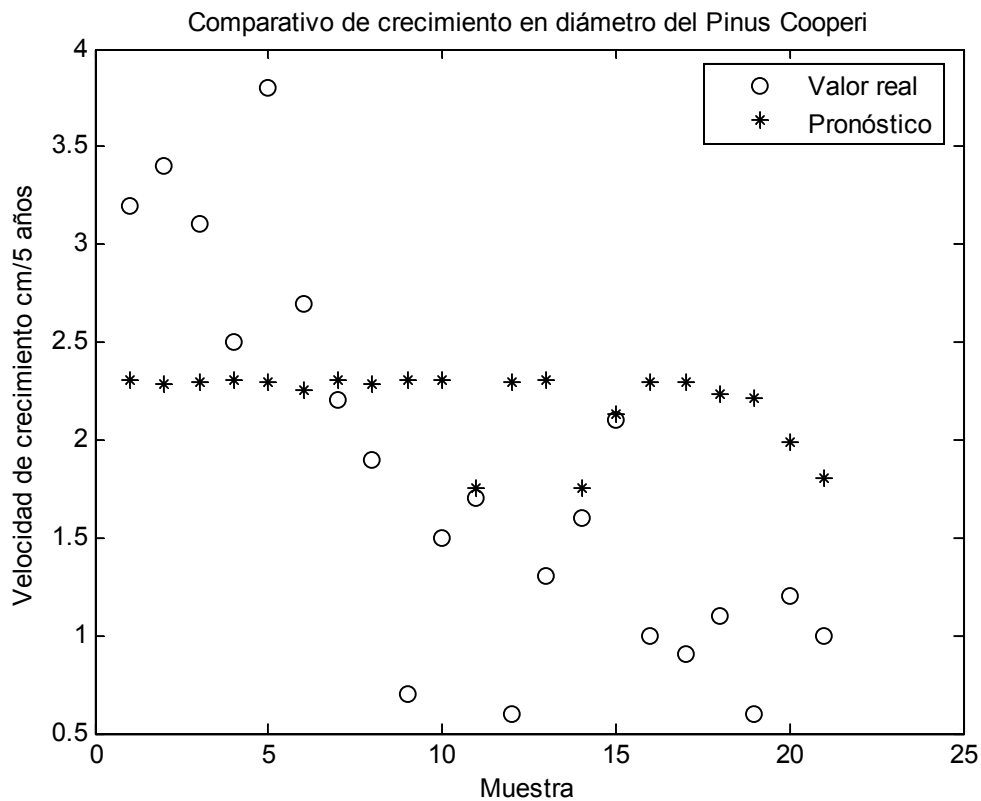
2.3.4.1. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

1 entrada parámetro diámetro 1.3, 1 salida parámetro velocidad de crecimiento en diámetro

3 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 50 épocas; hasta un error de 0.2677



Gráfica 37 predicción de crecimiento en diámetro (ejemplo 2.3.4.1.)

El resultado que se muestra es poco satisfactorio.

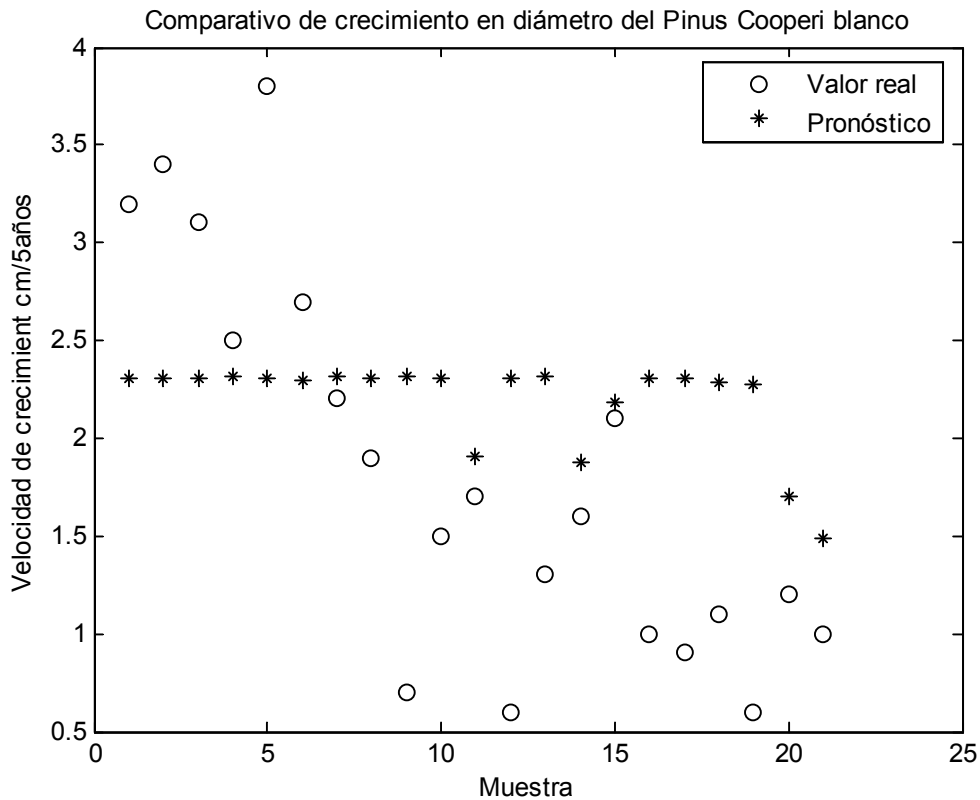
2.3.4.2. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

1 entrada parámetro diámetro 1.3, 1 salida parámetro velocidad de crecimiento en diámetro

5 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 50 épocas



Gráfica 38 predicción de crecimiento en diámetro (ejemplo 2.3.4.2.)

El resultado del el pronóstico no mejora.

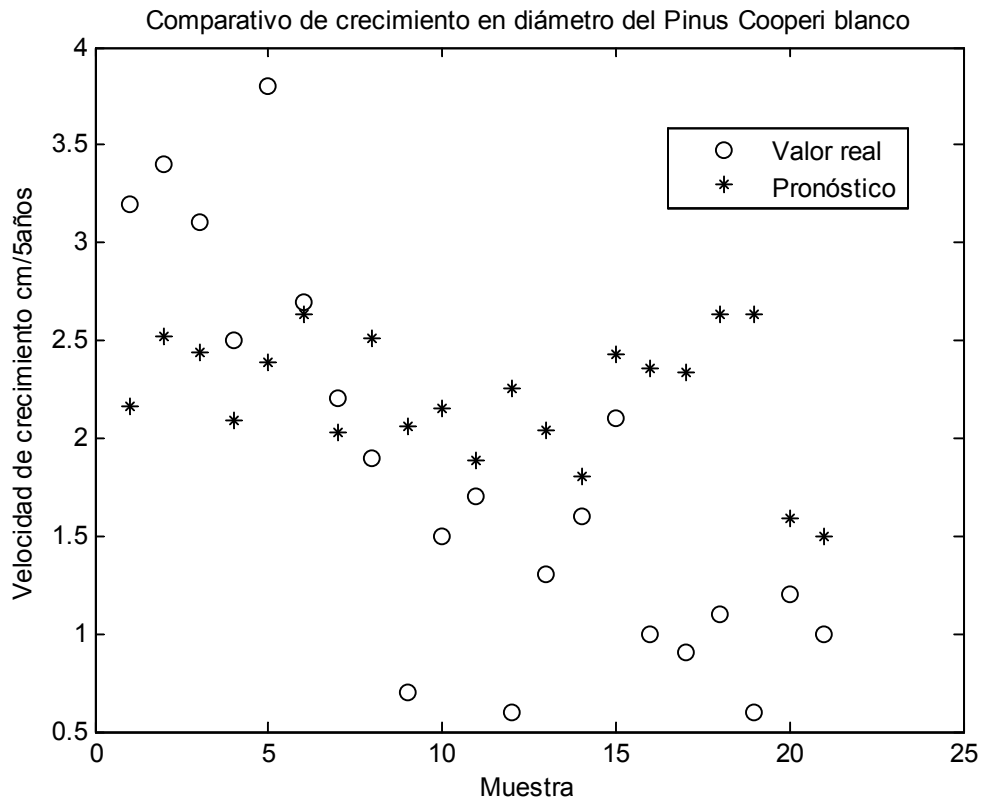
2.3.4.3. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

1 entrada parámetro diámetro 1.3, 1 salida parámetro velocidad de crecimiento en diámetro

11 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 50 épocas



Gráfica 39 predicción de crecimiento en diámetro (ejemplo 2.3.4.3.)

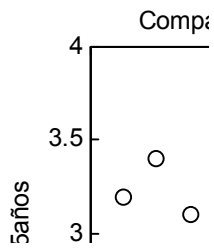
Con 11 membresías el resultado es un poco mejor pero muy lejos de ser satisfactorio.

2.3.4.4. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

2 entradas con los parámetros diámetro 1.3 y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro

3 variables lingüísticas por función de membresía del tipo campana por parámetro.
Entrenamiento 30 épocas



Gráfica 40 predicción de crecimiento en diámetro (ejemplo 2.3.4.4.)

Se obtiene un resultado muy pobre.

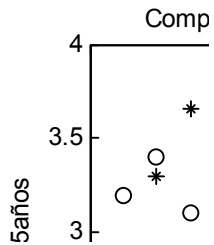
2.3.4.5. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

2 entradas con los parámetros diámetro 1.3 y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro

3 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento **200 épocas**.



Gráfica 41 predicción de crecimiento en diámetro (ejemplo 2.3.4.5.)

Aumentando a 200 épocas de entrenamiento el pronóstico mejora.

2.3.4.6. Predicción de velocidad de crecimiento en diámetro.

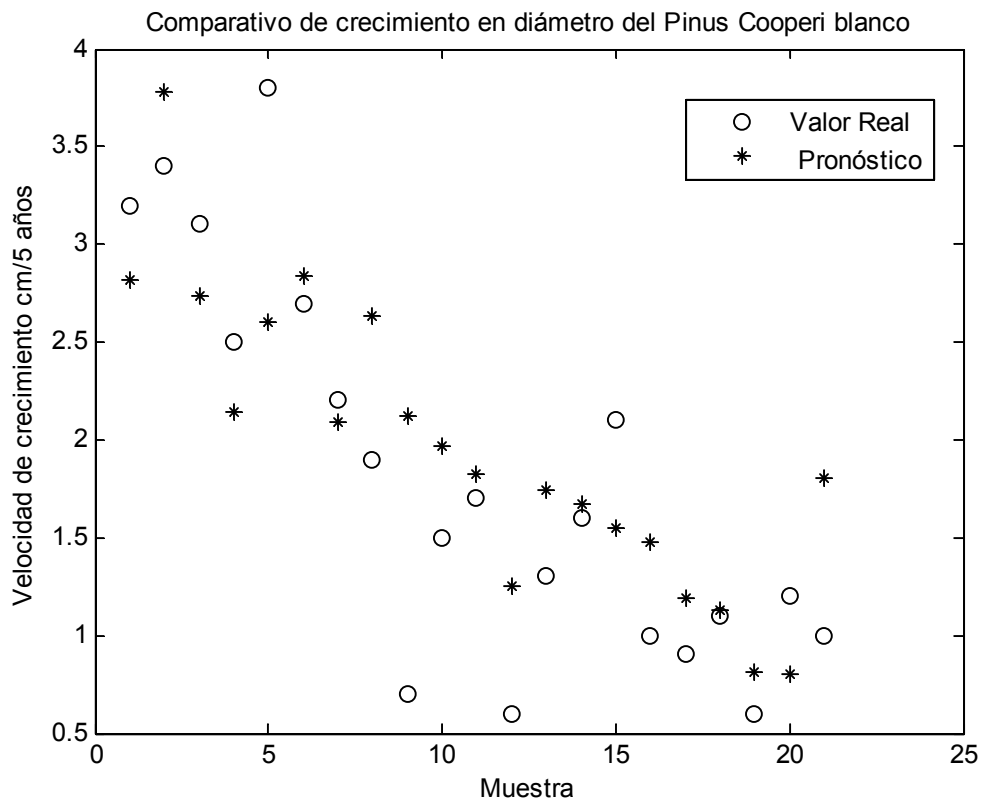
Reduciendo el Entrenamiento de 200 a 100 épocas.

Se entrena ANFIS:

2 entradas con los parámetros diámetro 1.3 y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro

3 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento **100 épocas**.



Gráfica 42 predicción de crecimiento en diámetro (ejemplo 2.3.4.6.)

Con 100 épocas el entrenamiento no mejora la capacidad de predicción.

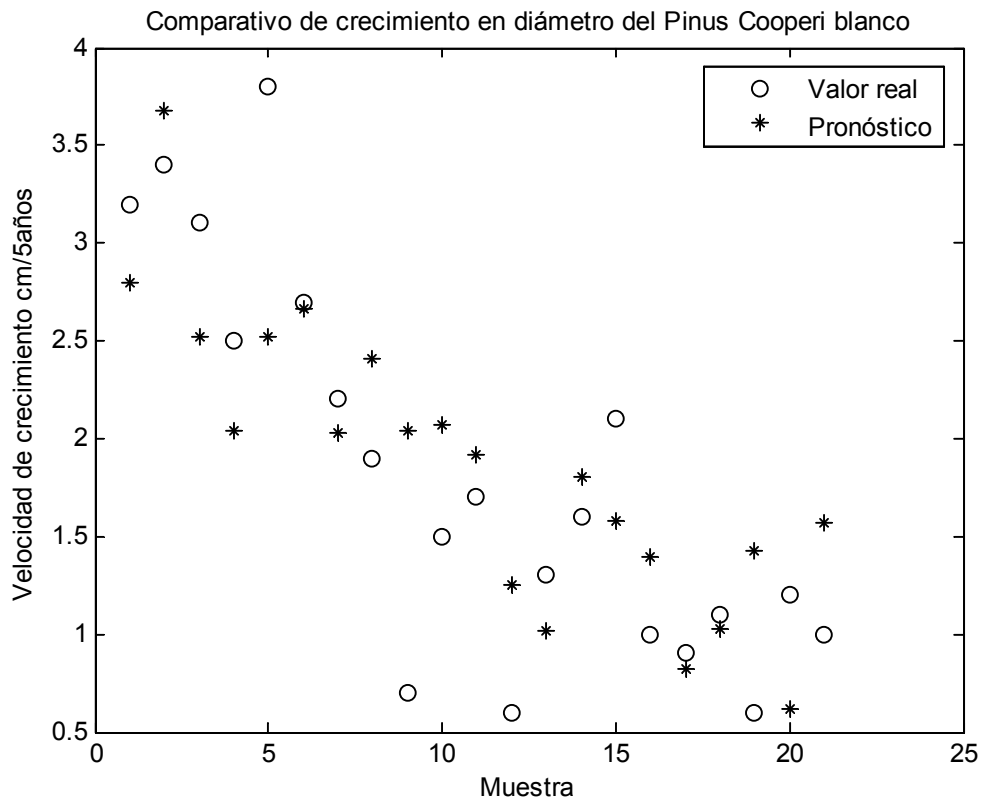
2.3.4.7. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

2 entradas con los parámetros diámetro 1.3 y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro

7 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 30 épocas.



Gráfica 43 predicción de crecimiento en diámetro (ejemplo 2.3.4.7.)

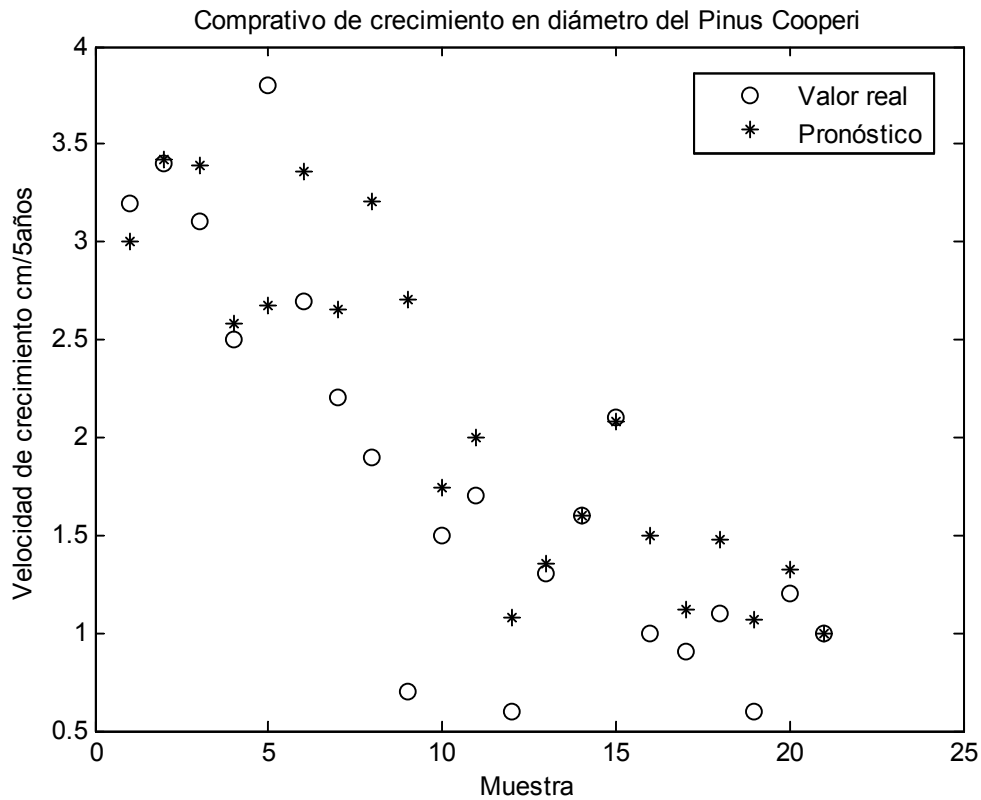
2.3.4.8. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

3 entradas con los parámetros: índice de sitio, diámetro 1.3 y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro

5 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 30 épocas.



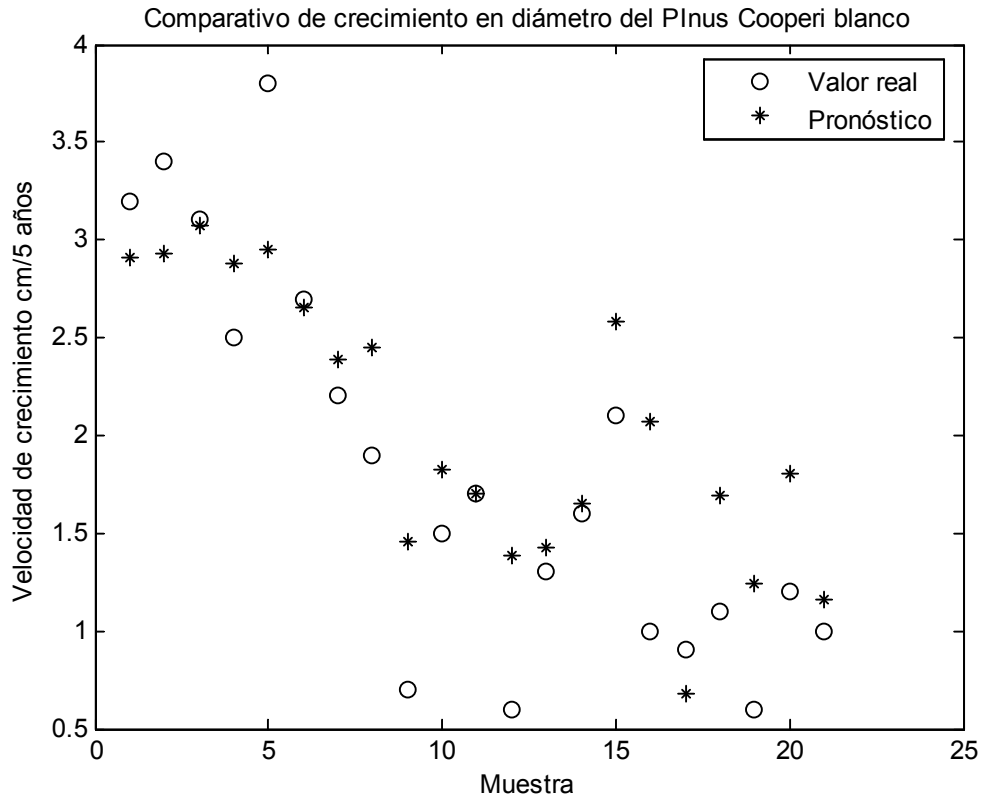
Gráfica 44 predicción de crecimiento en diámetro (ejemplo 2.3.4.8.)

Algunos valores son bien aproximados

2.3.4.9. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

4 entradas con los parámetros: índice de sitio, diámetro 1.3, índice de altura media, y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro
3 variables lingüísticas por función de membresía del tipo campana por parámetro.
Entrenamiento 30 épocas.



Gráfica 45 predicción de crecimiento en diámetro (ejemplo 2.3.4.9.)

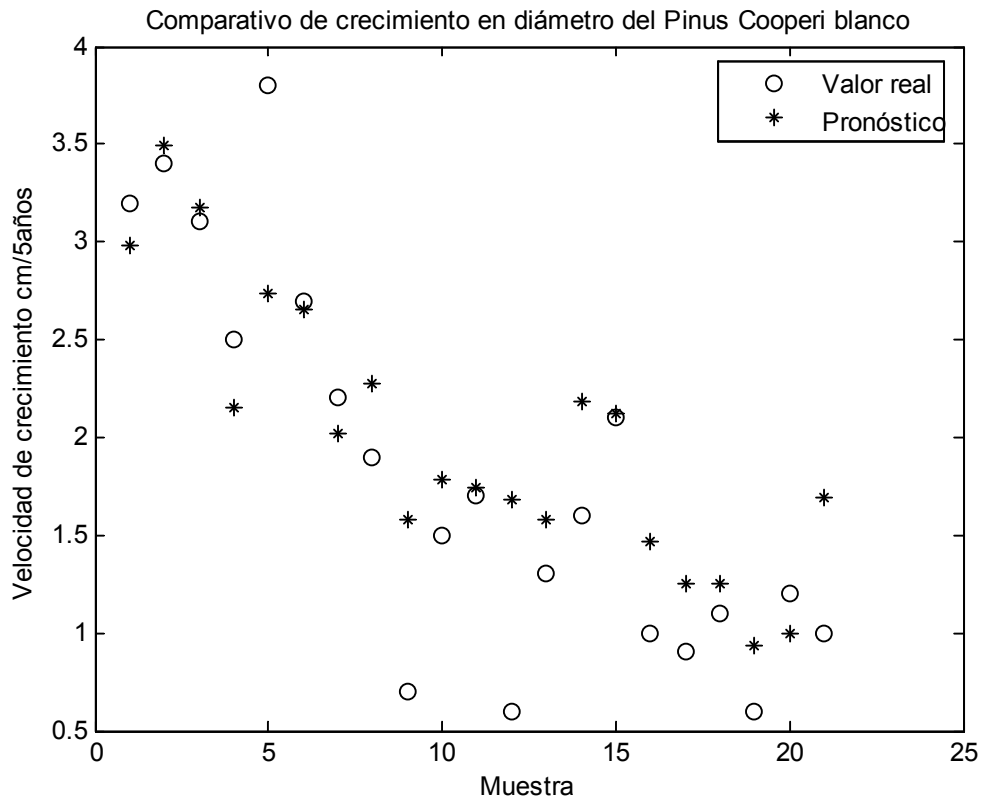
2.3.4.10. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

4 entradas con los parámetros: índice de sitio, diámetro 1.3, índice de altura media, y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro

5 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 30 épocas.



Gráfica 46 predicción de crecimiento en diámetro (ejemplo 2.3.4.10.)

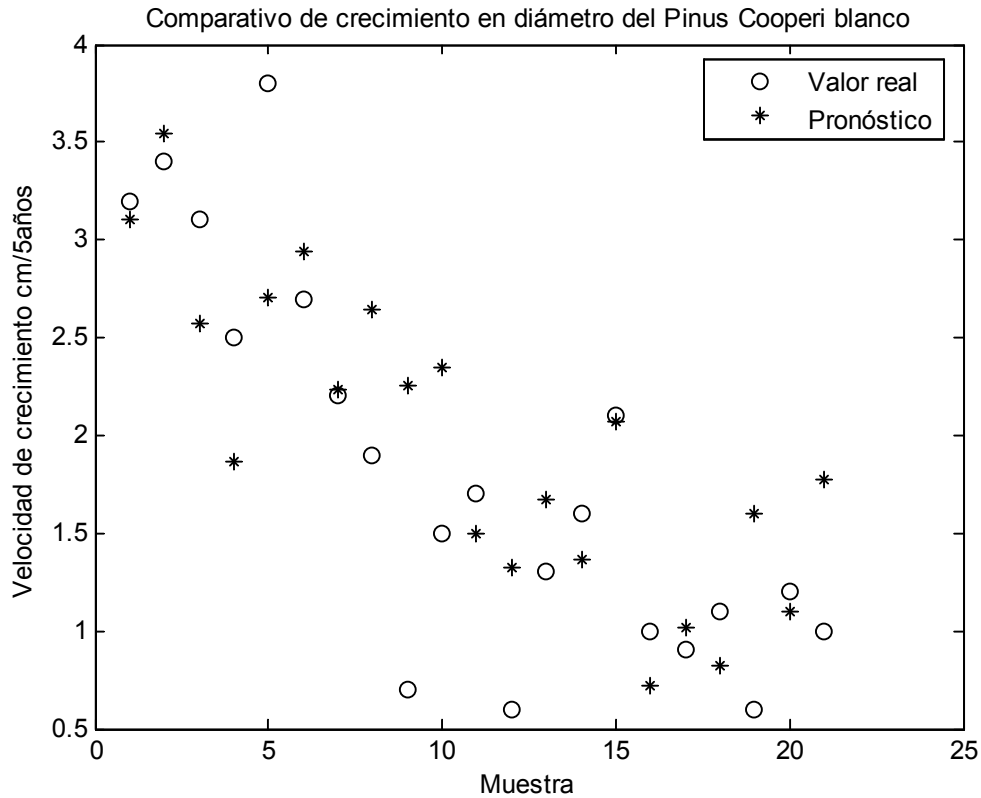
2.3.4.11. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

4 entradas con los parámetros: índice de sitio, diámetro 1.3, índice de altura media, y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro

7 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 100 épocas.



Gráfica 47 predicción de crecimiento en diámetro (ejemplo 2.3.4.11.)

Aumentando a 7 las variables lingüísticas por membresía y entrenando 30 épocas se observa que la red no alcanza a tener la capacidad de predecir el crecimiento en diámetro del *Pinus Cooperi* Blanco.

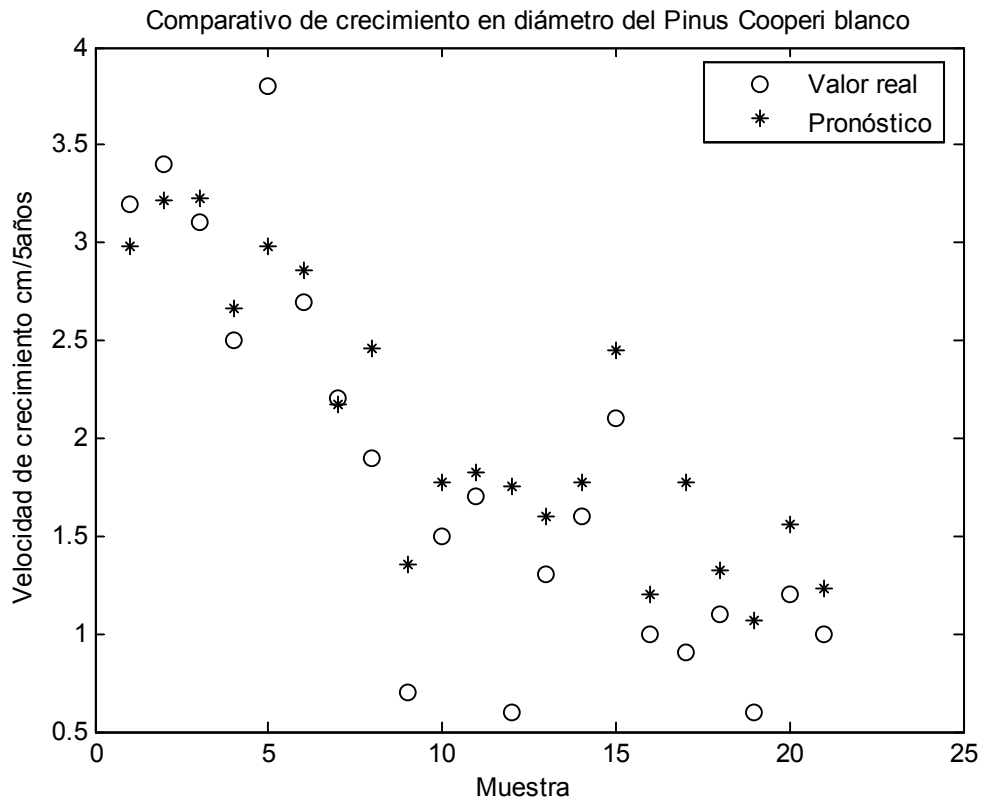
2.3.4.12. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

4 entradas con los parámetros: índice de sitio, diámetro 1.3, índice de altura media, y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro

3 5 7 7 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 3 épocas.



Gráfica 48 predicción de crecimiento en diámetro (ejemplo 2.3.4.12.)

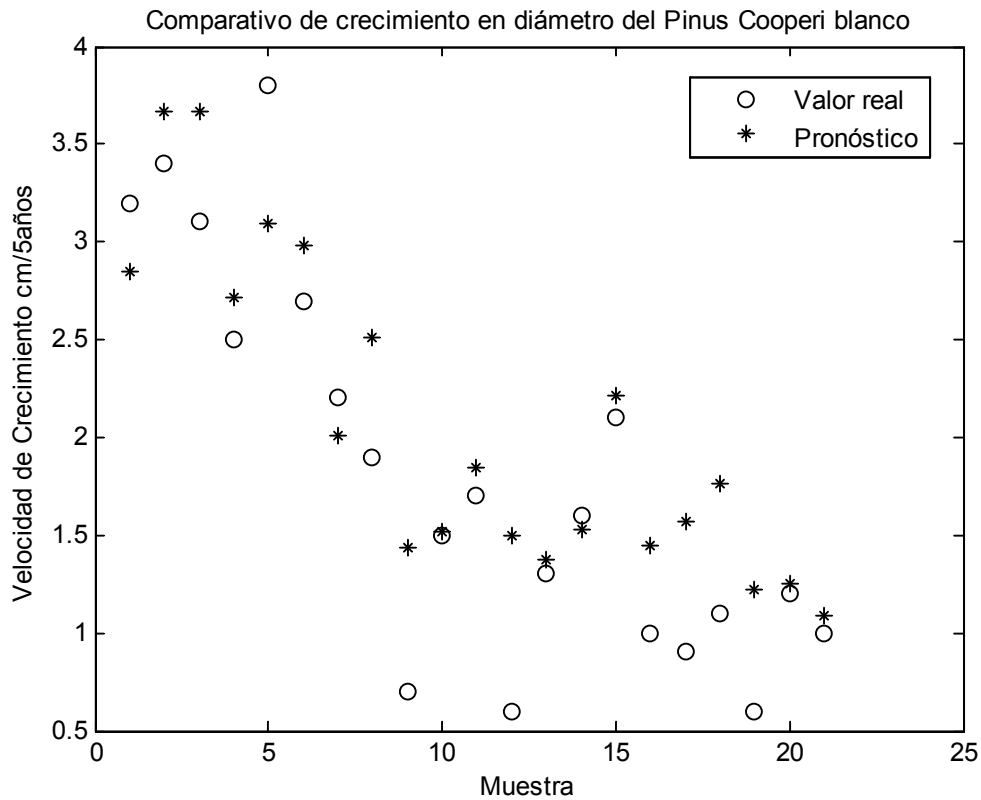
2.3.4.13. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

4 entradas con los parámetros: índice de sitio, diámetro 1.3, índice de altura media, y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro

3 5 7 7 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento **30 épocas**. Hasta un error de 0.22247



Gráfica 49 predicción de crecimiento en diámetro (ejemplo 2.3.4.13.)

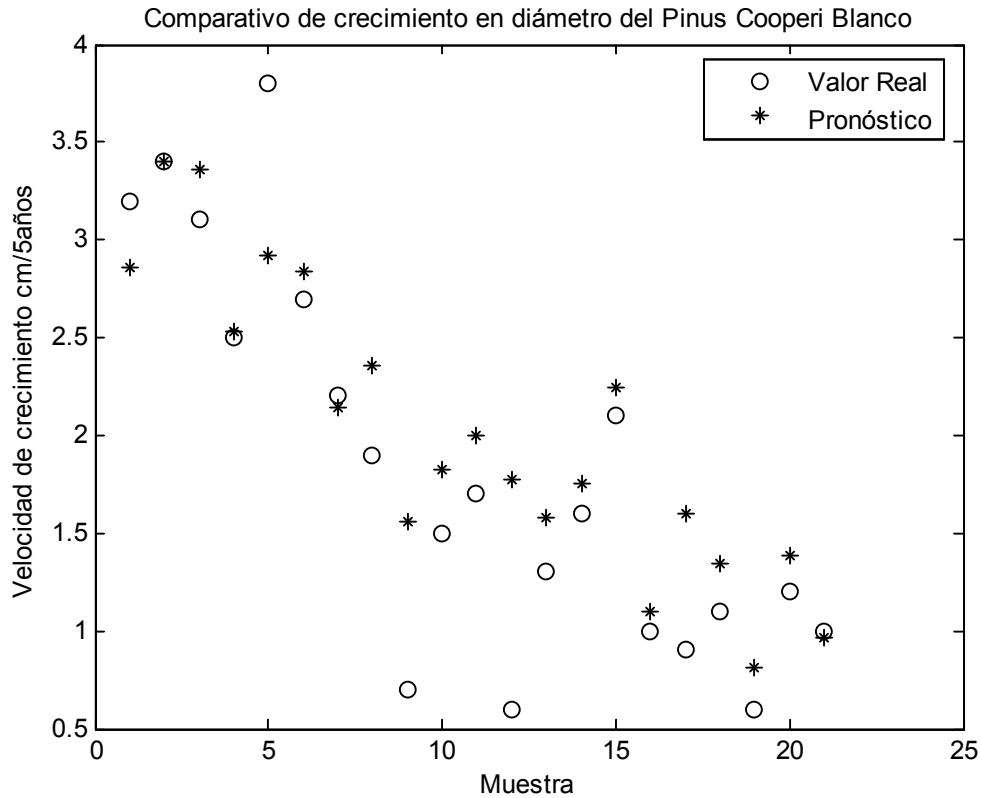
2.3.4.14. Predicción de velocidad de crecimiento en diámetro.

Se entrena ANFIS con:

4 entradas con los parámetros: índice de sitio, diámetro 1.3, índice de altura media, y el índice de Glover y Hool 1 salida parámetro velocidad de crecimiento en diámetro

3 5 7 9 variables lingüísticas por función de membresía del tipo triangular por parámetro.

Entrenamiento 30 épocas. Error 0.22379



Gráfica 50 predicción de crecimiento en diámetro (ejemplo 2.3.4.14.)

Los resultados muestran una tendencia a generalizar de la red pero aún requiere de más parámetros como entradas adicionales a la red.

2.3.5. Estimación de la velocidad de crecimiento en altura por medio de una red neuro-difusa (ANFIS).

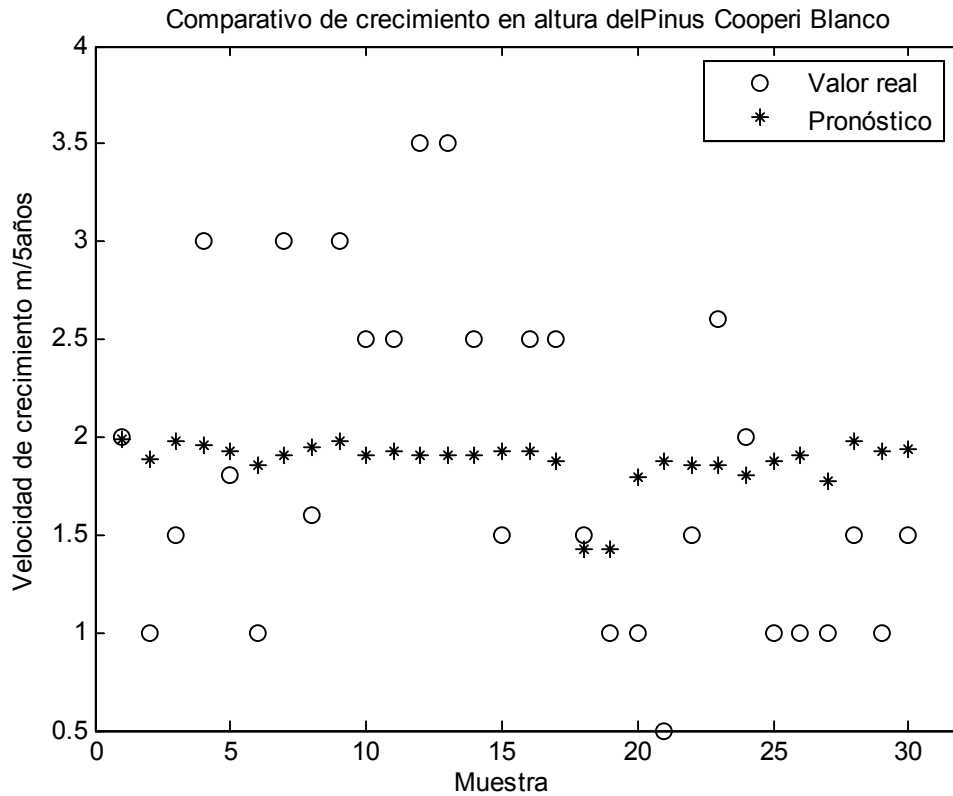
2.3.5.1. Predicción de velocidad de crecimiento en altura.

Se entrena ANFIS con:

1 entrada parámetro altura del pino, 1 salida parámetro velocidad de crecimiento en altura

3 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 30 épocas error 0.2184



Gráfica 51 predicción de crecimiento en altura (ejemplo 2.3.5.1.)

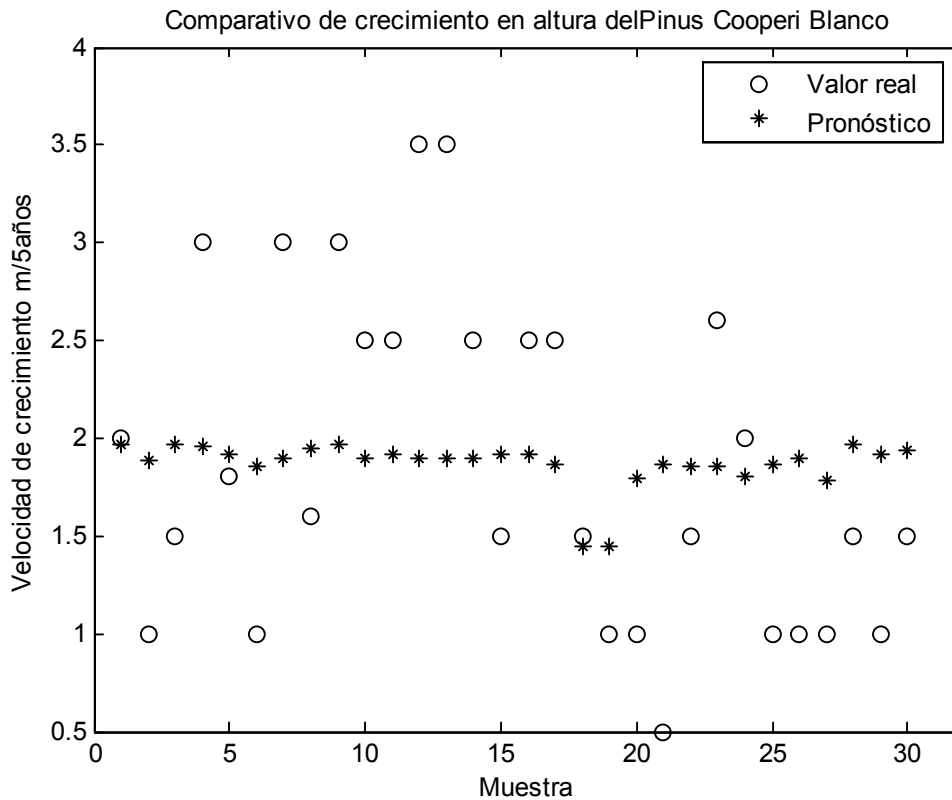
2.3.5.2. Predicción de velocidad de crecimiento en altura.

Se entrena ANFIS con:

1 entrada parámetro altura del pino, 1 salida parámetro velocidad de crecimiento en altura

3 variables lingüísticas por función de membresía del tipo trapezoidal

Entrenamiento 30 épocas, aproximando a un error del 0.21863



Gráfica 52 predicción de crecimiento en altura (ejemplo 2.3.5.2.)

El resultado es muy similar pero la rapidez de cálculo con membresía trapezoidal es mayor.

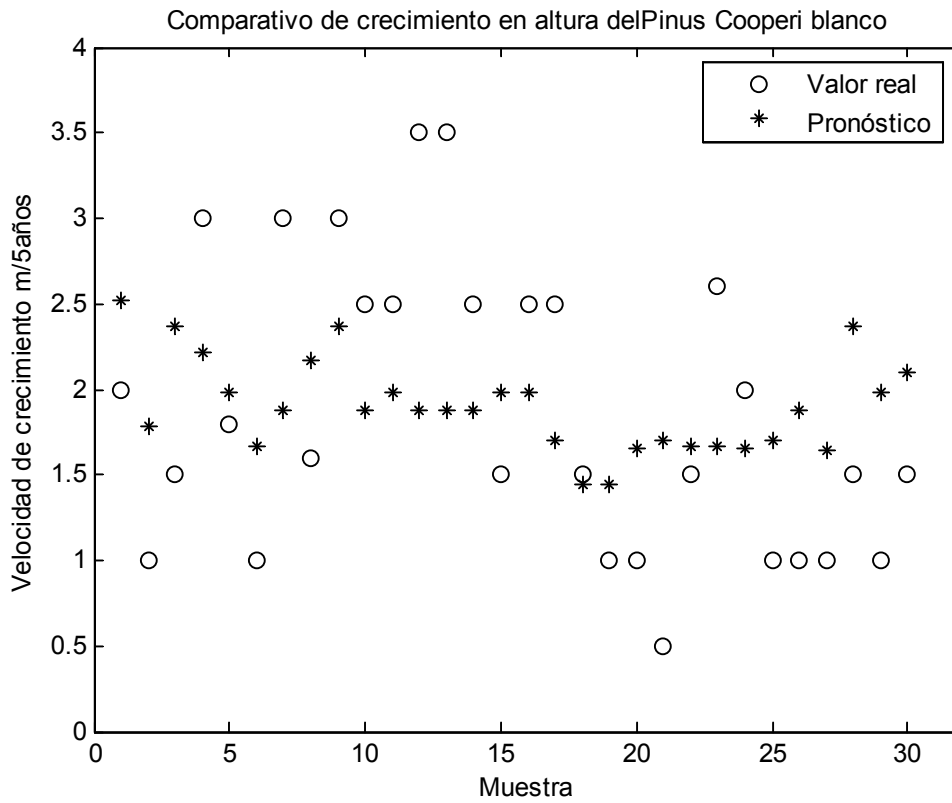
2.3.5.3. Predicción de velocidad de crecimiento en altura.

Se entrena ANFIS con:

1 entrada parámetro altura del pino, 1 salida parámetro velocidad de crecimiento en altura

3 variables lingüísticas por función de membresía del tipo triangular.

Entrenamiento 50 épocas error 0.2184



Gráfica 53 predicción de crecimiento en altura (ejemplo 2.3.5.3.)

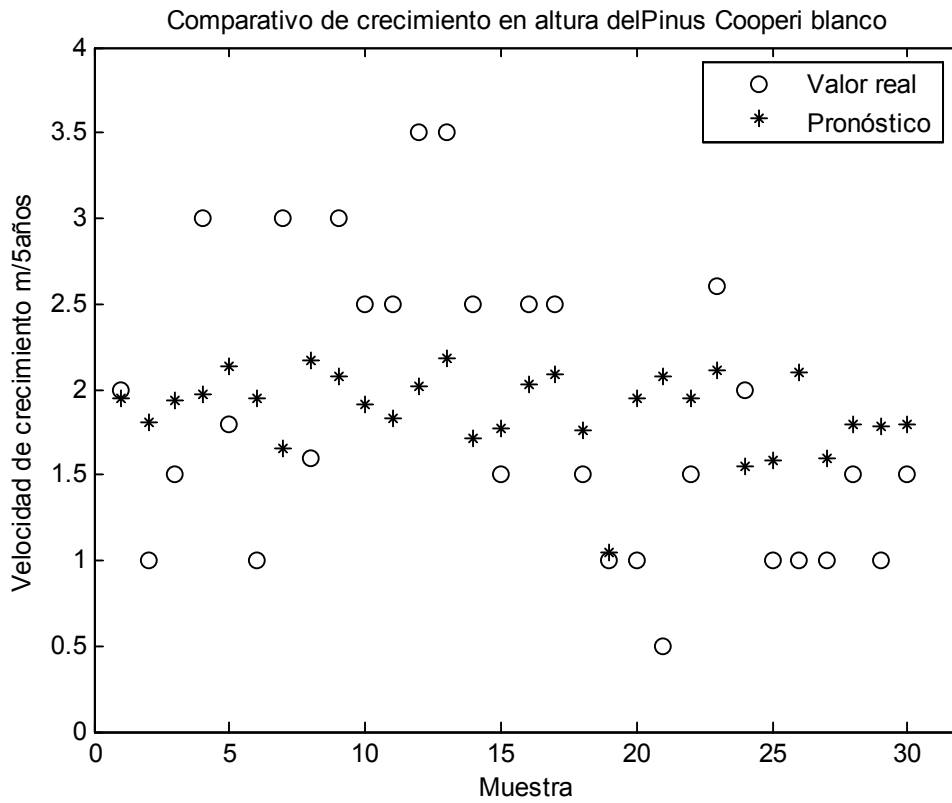
2.3.5.4. Predicción de velocidad de crecimiento en altura.

Se entrena ANFIS con:

2 entradas parámetro altura del pino y el índice de sitio; 1 salida parámetro velocidad de crecimiento en altura

3 variables lingüísticas por función de membresía del tipo triangular

Entrenamiento 50 épocas (hasta un error 0.21378)



Gráfica 54 predicción de crecimiento en altura (ejemplo 2.3.5.4.)

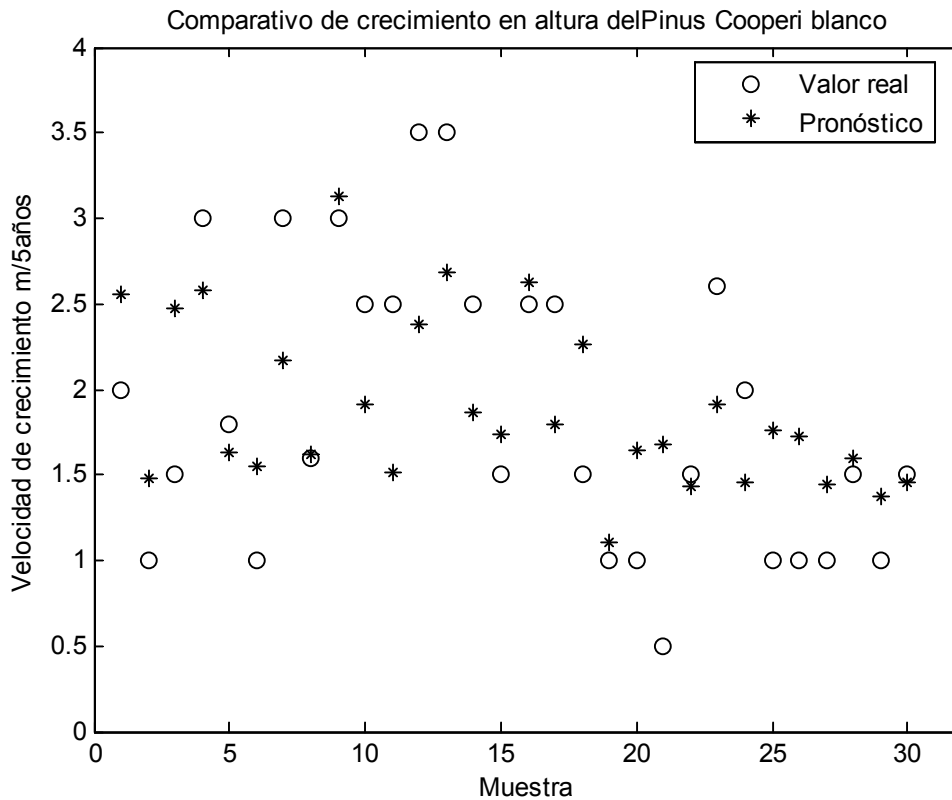
2.3.5.5. Predicción de velocidad de crecimiento en altura.

Se entrena ANFIS con:

2 entradas parámetro altura del pino y el índice de sitio; 1 salida parámetro velocidad de crecimiento en altura

11 variables lingüísticas por función de membresía del tipo triangular por parámetro.

Entrenamiento 50 épocas hasta un error de 0.19104.



Gráfica 55 predicción de crecimiento en altura (ejemplo 2.3.5.5.)

Aproxima bien 7 datos.

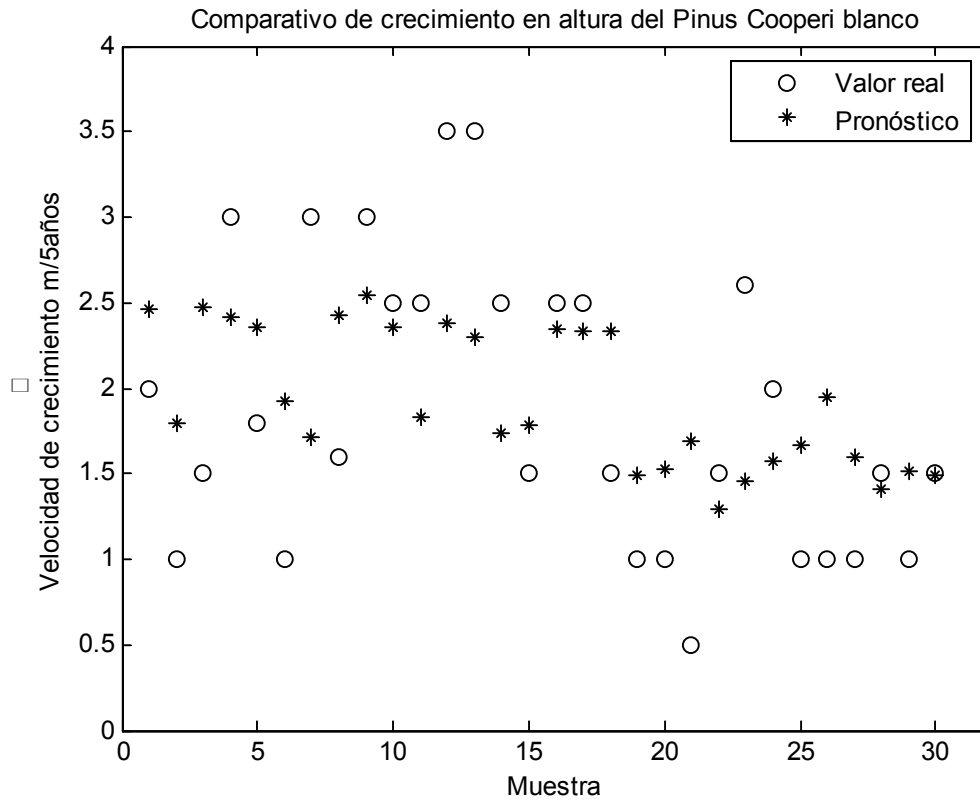
2.3.5.6. Predicción de velocidad de crecimiento en altura.

Se entrena ANFIS con:

3 entradas parámetro altura del pino, el índice de sitio y el índice de Glover y Hool; 1 salida parámetro velocidad de crecimiento en altura

5 variables lingüísticas por función de membresía del tipo triangular por parámetro

Entrenamiento 50 épocas hasta un error de 0.19104.



Gráfica 56 predicción de crecimiento en altura (ejemplo 2.3.5.6.)

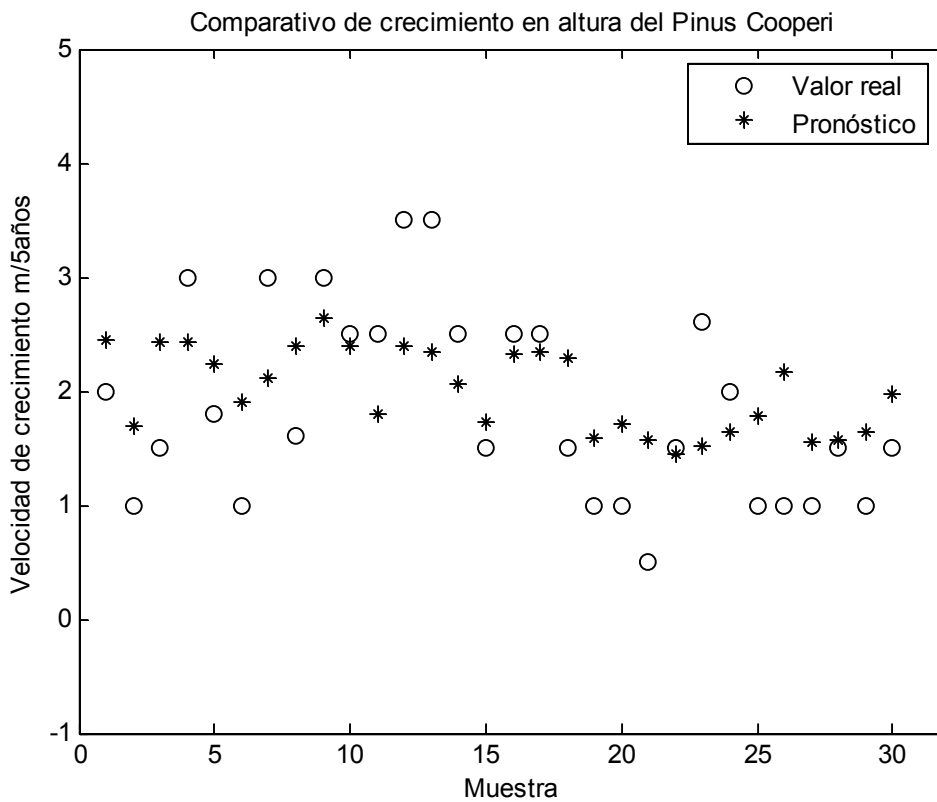
2.3.5.7. Predicción de velocidad de crecimiento en altura.

Se entrena ANFIS con:

4 entradas parámetro altura del pino, el índice de sitio, el índice de Glover y Hool y índice de amplitud número de árboles; 1 salida parámetro velocidad de crecimiento en altura

5 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 30 épocas.



Gráfica 57 predicción de crecimiento en altura (ejemplo 2.3.5.7.)

El resultado muestra que la red no puede predecir de forma correcta

2.3.5.8. Predicción de velocidad de crecimiento en altura.

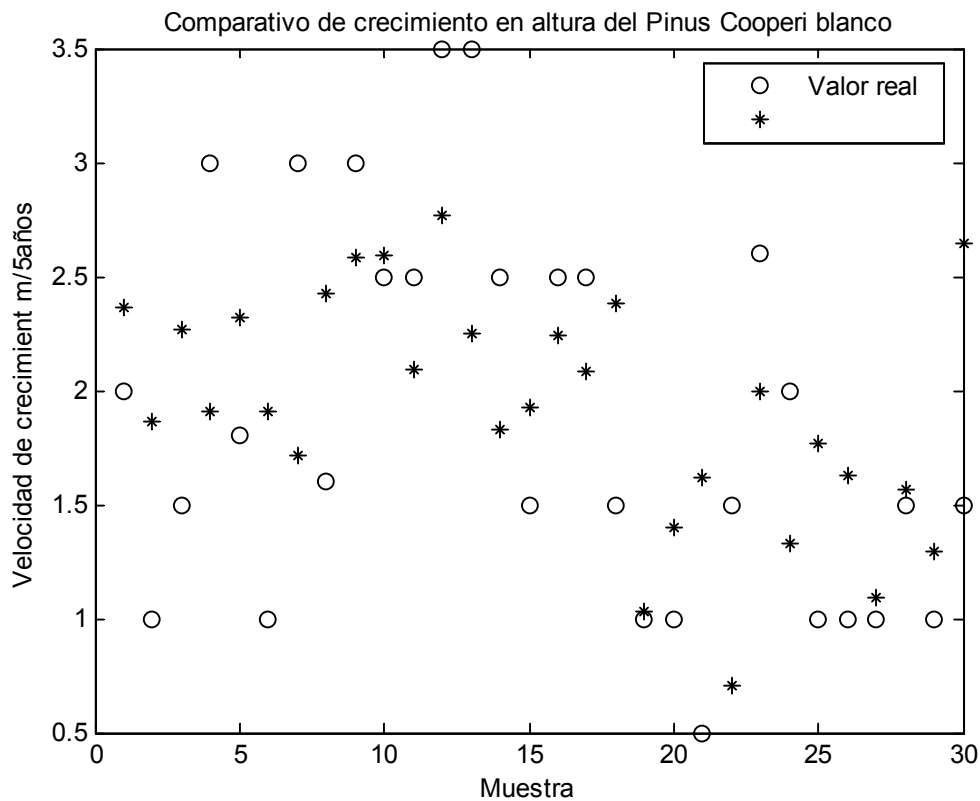
Se intenta mejorar la capacidad de predicción de la red cambiando uno de los parámetros de entrada: el índice de longitud número de árboles por el índice de amplitud número de árboles.

Se entrena ANFIS con:

4 entradas parámetro altura del pino, el índice de sitio, el índice de Glover y Hool y el índice de longitud número de árboles; 1 salida parámetro velocidad de crecimiento en altura

5 variables lingüísticas por función de membresía del tipo campana por parámetro.

Entrenamiento 30 épocas hasta un error de 0.18882



Gráfica 58 predicción de crecimiento en altura (ejemplo 2.3.5.8.)

La red no muestra capacidad de pronóstico como se ha visto en los ejemplos anteriores. Al parecer se necesita de más parámetros para poder describir el crecimiento en altura

CAPITULO III

3. ANÁLISIS DE RESULTADOS, CONCLUSIONES, RECOMENDACIONES Y TRABAJO A FUTURO.

3.1. ANÁLISIS DE RESULTADOS.

De los resultados obtenidos se observa que la red neuronal backpropagation aproximó de forma muy precisa la muestra de validación en crecimiento en diámetro 90% y altura 94.5% en un intervalo de confianza del 90% siendo esta red capaz de predecir este sistema biológico altamente complejo.

Del análisis de los resultados obtenidos con las diferentes morfologías de red analizadas se observó que tiende a existir una red óptima en generalizar. Si la red crece o disminuye en tamaño tiende a ir perdiendo su capacidad de generalizar. Esto es causado en una red pequeña, comparando su tamaño respecto a la red óptima, porque no tiene la capacidad de generalizar de un comportamiento tan complejo (multivariable no lineal) como el del sistema investigado en el presente trabajo por su escaso número de pesos y bias. Y una red grande es muy sensible a vacíos de información entre datos o a datos no bien explicados por la misma poca cantidad de información del comportamiento del sistema y oscilará excesivamente perdiendo capacidad de generalización. Así, pues se puede considerar con base a los resultados obtenidos que existe una región del tamaño de la red para un determinado sistema según lo bien descrito que esté por la base de datos disponible.

Algo similar ocurre con el entrenamiento de una red. Cuando se entrena por ejemplo la red óptima, esta red alcanzará en un determinado número de épocas su mejor capacidad de generalización. Al aumentar el entrenamiento su capacidad de generalización decrecerá. La explicación de que esto ocurra es principalmente causado por los patrones de entrenamiento, pues por lo general como en este trabajo, suelen ser escasos y la influencia de vacíos de información entre datos o de algunos valores de comportamiento no bien explicado por la cantidad insuficiente de información, produciendo en el modelo de la red la degradación de la capacidad de generalización.

En cuanto al funcionamiento de la red neuro-difusa ANFIS, de los resultados obtenidos se observa que la red en crecimiento en diámetro se acerca a hacer una predicción aceptable sin embargo se ve que aún le falta más información para lograrlo, pero su demanda de capacidad de memoria imposibilita el poder seguir buscando la red que pueda predecir correctamente.

En cuanto la predicción en altura la ANFIS nunca logró predecir de forma acertada la muestra de validación, sólo fue capaz de encontrar la región de predicción, al parecer requiere de una descripción más profunda del problema (requiere la red más parámetros de entrada)

Se observó que se puede mejorar la generalización de la red aumentando el número de variables lingüísticas por función de pertenencia (membresía) y/o épocas de entrenamiento, pero si no se tiene el número suficiente de variables que describen un el sistema no se podrá alcanzar la generalización conveniente de la red.

En ANFIS se observó que es una red fácil de utilizar, posee gran potencial, pero que adolece del problema de la explosión en la generación de las reglas y apenas se logró

utilizar de forma adecuada hasta cuatro dimensiones, pues agotó la capacidad de memoria de la máquina.

3.2. CONCLUSIONES:

De la utilización de los índices de competencia independientes de la distancia para predecir el crecimiento del diámetro, y altura de la especie *Pinus Cooperi* Blanco se puede concluir:

a) Una red neuronal del tipo Backpropagation entrenada con índices de competencia independientes de la distancia fue capaz de predecir el crecimiento en diámetro y altura de un bosque natural de *Pinus Cooperi* Blanco a diferente número de épocas de entrenamiento.

b) Una red neuronal backpropagation entrenada hasta un error medio cuadrático de 0.01, alimentada con entradas 21 y con 1 salida.

Con un tratamiento de los patrones de entrenamiento a media cero (0) y desviación uno (1).

Con una Arquitectura:

6 capas, con 5 capas ocultas con 30 neuronas con función tangencial

1 capa de salida con 1 neurona función lineal.

Aplicando el algoritmo de aprendizaje gradiente conjugado.

Fue capaz de generalizar con una aproximación a la muestra del crecimiento en diámetro con un **90%** de exactitud en un intervalo de confianza del 90%

c) Una red neuronal backpropagation entrenada hasta un error medio cuadrático de 0.1, alimentada con entradas 21 y con 1 salida.

Con un tratamiento de los patrones de entrenamiento a media cero (0) y desviación uno (1).

Con una Arquitectura:

6 capas, con 5 capas ocultas con 30 neuronas con función tangencial

1 capa de salida con 1 neurona función lineal.

Aplicando el algoritmo de aprendizaje gradiente conjugado.

Fue capaz de generalizar con una aproximación a la muestra del crecimiento en altura **94.5%** de exactitud en un intervalo de confianza del 90%.

d) La red neuronal backpropagation entrenada con índices de competencia independientes de la distancia es una herramienta adecuada para predecir la velocidad de crecimiento del diámetro y altura del *Pinus Cooperi* Blanco como los resultados lo demuestran.

e) En el desarrollo de esta investigación uso de la red neuronal backpropagation mostró resultados positivos en la predicción del comportamiento de un sistema altamente complejo como lo es el crecimiento de los individuos que integran un bosque.

f) Sin embargo con una red neuro-difusa ANFIS haciendo uso de de cuatro índices fue sólo capaz de seguir la tendencia del pronóstico de la muestra de validación, para predecir el crecimiento en diámetro, es necesario incluir más variables, para lo cual se creará una necesidad de mucha más memoria y en el crecimiento de altura la red resintió aún mas la falta de variables que describen el sistema.

g) La red neuronal backpropagation por su alta eficiencia en el modelado de sistemas, sólo enfrenta como una alternativa viable de uso a las redes neuro-difusas, sin embargo las supera ampliamente en el poco consumo de memoria, por lo cual hace posible resolver el problema de estimación de crecimiento en el mismo campo de explotación maderera, requiriéndose tan solo de una computadora portátil. Mientras una red neuro-difusa hubiera requerido del uso de una gran cantidad de memoria haciendo restrictivo su uso.

3.3. RECOMENDACIONES:

- 1) La capacidad de poder modelar un sistema de alta complejidad por medio de una red neuronal abre la posibilidad de poder extraer el comportamiento dinámico de otros sistemas que presentan igualmente un alto grado de complejidad como son otros sistemas biológicos.
- 2) Por lo general el crecimiento de un árbol requiere de un tiempo mayor al de vida de un ser humano, por lo que se hace difícil el seguir su crecimiento durante las diferentes etapas de su vida. Así, el contar con información en todo el tiempo de vida de los individuos que forman un bosque no es fácil de obtener. Por lo observado en este trabajo de investigación se puede recomendar el realizar mediciones en un mismo tiempo a diferentes individuos de un bosque extrae el comportamiento dinámico de crecimiento de un individuo, siendo una alternativa mucho más económica y rápida para la obtención de información para el desarrollo técnicas de predicción.
- 3) Se recomienda para un trabajo futuro el uso de las redes radial basis, probabilística, regresión generalizada no consideradas en este trabajo.
- 4) Sería recomendable observar la capacidad de predicción de una red neuronal haciendo uso mezclado de índices de competencia dependiente e independiente de la distancia en la predicción del crecimiento de la altura, de la especie *Pinus Cooperi* Blanco.

3.4. TRABAJO A FUTURO:

- a) Ampliar el estudio de la red neuro-difusa ANFIS ya disponiendo de una mayor capacidad de memoria.
- b) En función de disponer de una base de datos con una mayor cantidad de información; ampliar la capacidad de generalización de la red a una más amplia gama de patrones de crecimiento.
- c) Utilizar índices de competencia dependientes de la distancia (Tome, 1989) y analizar su diferencia con respecto a la utilización de índices de competencia independientes de la distancia.

BIBLIOGRAFÍA

BIBLIOGRAFÍA.

- Abhijit S. Pandey, Robert B. Macy;** “Pattern Recognition with Neural Networks in C++”; IEEE Press, 1995.
- Agresti Alan,** Categorical Data Analysis, John Wiley, New York, 1990.
- Aldrich John, y Forrest Nelson,** Linear Probability, Logit, and Probit Models, Sage University Paper, California, 1984.
- Anderson, James A.** "An Introduction to Neural Networks", MIT Press, 1995
- Bishop, Christopher M.** "Neural Networks for Pattern Recognition", Oxford University Press, 1995
- Blum, E. K. and I. k. Li.** Approximation theory and feedforward networks. Neural Networks 4(4):511-515 (1991)
- Boosley, K.M.** (1998). Neurofuzzy Modelling Approaches in System Identification. University of Southampton.
- Braae, M. and D.A. Rutherford** (1979). Theoretical and linguistic aspects of the fuzzy logic controller. *Automatica* 15, pp. 553-577.
- Bormann, F. H and G. E. Likens.** 1979. Pattern and process in a forested ecosystem. Springer-Verlag, NY.
- Bruce, D. and L. C. Wensel.** 1987. Modelling forest growth approaches, definitions, and problems. For. Serv. Gen. Tech. Rep. NC- 120. pp: 1-8
- Caruana Rich, Steve Lawrence, C. Lee Giles.** *Overfitting in Neural Networks: Backpropagation, Conjugate Gradient, and Early Stopping*, Neural Information Processing Systems, Denver, Colorado, November 28–30, 2000.
- Chater, N.** (1995). Neural networks: The new statistical models of mind. En Levy, J.P. y Bairaktaris, D. (eds.), Connectionist models of memory and language. Ucl Press. London.
- Cherkassky, V., Friedman, J.H. y Wechsler, H.** (eds.) (1994). From statistics to neural networks. Springer- Verlag. Berlin.
- Cohen, J. y Cohen, P.** (1983). Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum.
- Clutter J.L. Forston J.C. Pinear L.V. Brister H.G. Bailey R.L.,** Timber management a quantitative approach. New York Jhon Wiley and Sons p 333 (1983).
- Croall, I.F. y Mason, J.P.** (1992). Industrial applications of neural networks. Springer-Verlag. N.Y.
- Cybenko G.**(1989) *Approximation by Superpositions of a Sigmoidal Function*, Mathematics of Control, Signals and Systems, Vol.2, pp. 303–314.
- Dayhoff Judith E., Omid Omidvar** Neural Networks and Pattern Recognition Publisher: Academic Press; 1st edition (January 15, 1998)
- Devijver, P.A. Y Kittler, J.V.** (1982), Pattern Recognition. A Statistical Approach. Prentice Hall- Englewood Cliffs.
- Fausett Laurene** Fundamentals of Neural Networks Publisher: Prentice Hall; 1st edition (January 15, 1994)
- Flexer, A.** (1995). Connectionist and statisticians, friends or foes?. The Austrian Research Institute for Artificial Intelligence. Acceso FTP. Servidor: ai.univie.ac.at.
- Fogelman, F.** (1994). Neural networks architectures for pattern recognition. En Cherkassky, V., Friedman, J.H. y Wechsler, H. (eds.), From statistics to neural networks. Springer-Verlag. Berlin.
- Frasconi P., M. Gori, and G. Soda.** Local Feedback Multilayered Networks. Neural Networks, 4:120-130, 1992.

- Freeman J. A. Skapura D. M.** Neural Networks. Algorithms, Applications, and Programming Techniques. Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1991.
- Friedman J.H.** Introduction to computational learning and statistical prediction. Tutorial Presented at Neural Information Processing Systems, Denver, CO, 1995.
- Garson, G.D.** (1991). A comparison of neural network and expert systems algorithms with common multivariate procedures for analysis of social science data. *Social Science Computer Review*, 9, 399-434.
- Glover, G. R. and J. N. Hool** 1979. A basal area ratio predictor of loblolly pine plantation mortality. *For. Sci.* 25: 275-282
- Guan, B. T. and G. Gertner** 1991a. Using a parallel distributed processing system to model individual tree mortality. *For. Sci.* 37: 871-885.
- Guan, B. T. and G. Gertner** 1991b. Modeling red pine tree survival with an artificial neural network. *For. Sci.* 37: 1429-1440.
- Guan, B. T. and G. Gertner** 1995. Modeling individual tree survival probability with a random optimization procedure: An artificial neural network approach. *AI Application* 9: 39-52.
- Guan, B. T., G. Gertner, and P. Parysow** 1997. A framework for uncertainty assessment of mechanistic forest growth models: A neural network example, *Ecol. Model.* 98: 47-58.
- Hagan M. T., Demouth Howard B.** Neural Network Design. PWS Publishing Company (1996).
- Hagan M. T. and Menhaj M. B.**, "Training feedforward neural networks with the Marquardt algorithm," *IEEE Trans. Neural Networks*, vol. 5, no. 6, November 1994.
- Hagem, O.** 1962. Additional observations on the dry matter increase of coniferous seedlings in winter. *Medd. fra Vestlandets Forstlige Forsøksstasjon*, 37 (11): 5.
- Hasenauer, H. and D. Merkl** 1997. Forest tree mortality simulation in uneven-aged stands using connectionist networks. In: Bulsari, A. B., and S. Kallio (eds.). *Neural Networks in Engineering Proc. Int. Conf. on Engineering Applications of Neural Networks (EANN'97)*, Stockholm, Sweden,
- Hastie y Tibshirani** (1994). Nonparametric Regression and Classification. Part 1- Nonparametric Regression. En Cherkassky, V., Friedman, J.H. y Wechsler, H. (eds.), *From statistics to neural networks*. Springer- Verlag. Berlin.
- Haykin, S.** "Neural Networks: A Comprehensive Foundation". Macmillan, 1994.
- Hecht-Nielsen, R.** (1990). Neurocomputing. Addison-Wesley. Cal.
- Hertz, J. Krogh, A. y Palmer, R.** (1991). Introduction to the theory of neural computation. Addison-Wesley. Cal.
- Hilera, J.R. y Martínez, V.J.** (1995). Redes neuronales artificiales: Fundamentos, modelos y aplicaciones. Ra-ma. Madrid.
- Hilt, D.E and R.M., Teck.** 1987. Individual -tree diameter growth model for Northern new England. Presented at the IUFRO Growth and Yield Modeling and Prediction Conference, Minneapolis, MN. p 86-92.
- Hinton G.E. Rumelhart D.E and R.J. William.** Learning internal representation by back-propagation errors. *Nature*, 323:533-536, 1986.
- Hornik K.M., M. Stinchcombe, H. White,** "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5 pp. 359-366, 1989.
- Howard Demuth, Mark Beale;** "Neural Network Toolbox for Use with Matlab"; User's Guide version 4.0 The Mathworks Inc., 2003.
- Huang, W.Y. y Lippmann, R.P.** (1987). Comparisons between neural nets and conventional classifiers. Proceedings of the IEEE International Conference on

Neural Networks, I, 485-494.

Hwang J.N, S.R. Ray, M. Maechler, D. Martin, and J. Schimert. Regression Modelling in Backpropagation and Projection Pursuit Learning. *IEEE Transactions on Neural Networks*, 5(3):54-64, 1994.

Jamshidi, M. (1997). *Large-Scale Systems: Modelling, Control and Fuzzy Logic*. Prentice Hall. New Jersey.

Jang, J.-S. (1993). ANFIS: Adaptive-Network-based Fuzzy Inference Systems. En: *IEEE Transactions on Systems, Man & Cybernetics* No. 23. pp 665-685.

Jang J.S.R., Sun C.T., Mizutani E. *Neuro-Fuzzy and Soft Computing*. Prentice Hall (1997).

Karampiperis P., N. Manouselis, T. Trafalis, "Architecture selection for neural networks", Proc. of IEEE World Congress on Computational Intelligence, Hawaii, May 2002.

Kerson Huang, 1987. *Statistical Mechanics*, 2d ed. (Wiley: New York,).

Knapp, T.R. (1978). Canonical correlation analysis: A general parametric significance testing system. *Psychological Bulletin*, 85, 410-416.

Kolmogorov, A.N. (1957), *On the Representation of Continuous Functions of Many Variables by Superposition of Continuous Functions of One Variable and Addition*, *Doklady Akademii Nauk SSR*, Vol. 114, pp. 953-956.

Kosko, B. (1994). Fuzzy systems as universal approximators. *IEEE Transactions on Computers*. 43(11), pp. 1329-1333.

Kosko B.; *Neural networks and fuzzy systems: A dynamical systems approach to machine intelligence*. Prentice Hall, 1991.

Kwok T.Y. and D.Y. Yeung. Constructive Algorithms for Structure Learning in Feedforward Neural Networks for Regression Problems. In *IEEE Transactions on Neural Networks*, volume 8, pages 630-645, 1997.

LI, X., van der Wal, A.J., Ruan, D. (1996). Discussion on soft computing at FLINS'96.

Law A. and Kelton D. *Simulation Modelling and Analysis*. Mc-Graw Hill, New York, 1990.

Lev Davidovich Landau and E. M. Lifshitz, 1969, *Statistical Physics (Course of Theoretical Physics*, vol. 5) 2d ed. (Pergamon: Oxford).

Maren, A., C. Harston, and R. Pap, 1990. *Handbook of Neural Computing Applications*, Academic Press, San Diego, California, pp.360.

Manzanilla B. H. (1993). Los sitios permanentes de investigación silvícola un sistema integrado para iniciarse en el cultivo de los ecosistemas forestales. *Boletín Técnico* No 116. SARH-INIFAP. México. 101p.

Martín, B. y Sanz, A. (1997). *Redes neuronales y sistemas borrosos*. Ra-ma. Madrid.

McClelland, J.L. y Rumelhart, D.E., (eds) (1986). *Parallel distributed processing*. Vol. 2. Cambridge, Mass.: MIT Press.

Mehra P. and B. W. Wah, *Artificial Neural Networks: Concepts and Theory*, IEEE Computer Society Press, 1992.

Mehrotra K., C.K. Mohan, and S. Ranka, *Elements of Artificial Neural Networks*, M.I.T. Press, 1997

Michie, D., Spiegelhalter, D.J. y Taylor, C.C. (eds) (1994). *Machine learning, neural and statistical classification*. Londres: Ellis Horwood.

Mohamad H. Hassoun, *Fundamentals of artificial neural networks*, Publisher: MIT Press; (March 27, 1995)

Moller, M. F., "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, vol. 6, pp. 525-533, 1993.

- Montgomery Douglas C.** (1991) Design and Analysis of experiments. John Wiley and Sons, Inc.
- Nauck, D.** (1993). A Fuzzy Neural Network Learning Fuzzy Control Rules and Membership Functions by fuzzy error propagation. En: Proc. IEEE Int. Conference on Neural Networks 1993. pp 1022-1027.
- Nauck, D.** (1995). Beyond Neuro-Fuzzy: Perspectives and Directions. Technical University of Braunschweig, Germany.
- Nebot A. Cueva J., Alquézar R.** Experimental Comparison of Fuzzy and Neural Network Techniques in Learning Models of the Central Nervous System Control. In Proc. of IEUFIT97, 1997
- N. G. van Kampen,** 1981. *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam).
- Ovington, J. D.** 1957. The volatile matter, organic carbon and nitrogen contents of tree species grown in close strands. *New Phytol.*, 56: 289-388.
- Ovington, J. D.** 1958. Some biological considerations of forest production. In *The biological productivity of Britain*. London, Institute of Biology, p. 83-90.
- Ovington, J. D.** 1960. The accumulation of energy in forest plantations in Britain. *J. Ecol.*, 48: 639-646.
- Ovington, J. D.** 1961. The productivity of some British wood lands. In *Recent advances in botany*. Toronto, Univ. of Toronto Press.
- Parisi R., E. D. Di Claudio, G. Orlandi, B. D. Rao,** "A generalized learning paradigm exploiting the structure of feedforward neural networks," *IEEE Trans. Neural Networks*, vol. 7, no. 6, pp. 1450-1459, November 1996.
- Pham D.T., S. Sagioglu,** "Training multilayered perceptrons for pattern recognition: a comparative study of four training algorithms", *International Journal of Machine Tools & Manufacture* 41, 2001.
- Pratt Ian.** Artificial Intelligence. The Macmillan Press LTD.(1994)
- Press, S.J. y Wilson, S.** (1978). Choosing between logistic regression and discriminant analysis. *Journal of the American Statistical Association*, 73, 699-705.
- Reed R.D. and R.J Marks.** *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks*. MIT Press, Cambridge, MA, 1999.
- Riedmiller Martin Heinrich Braun.** A Direct Adaptive Method for Faster Backpropagation Learning The RPROP Algorithm. In *IEEE International Conference on Neural Networks (San Francisco)* vol 1, pp. 586_591. IEEE, New York 1993.
- Ripley, B.D.** (1993). Statistical aspects of neural networks. En O.E. Barndorff-Nielsen, J.L. Jensen y W.S. Kendall (eds), *Networks and chaos: Statistical and probabilistic aspects*. Londres: Chapman and Hall.
- Ripley, B.D.** (1996). *Pattern recognition and neural networks*. Cambridge Univ. Press. Cambridge, G.B.
- Ripley, B.D.** "Can Statistical Theory Use Neural Networks Better?", *Proceedings of the Interface '97, 29th Symposium on the Interface: Computing Science and Statistics*, 1997.
- Rojas, R.** "Neural Networks: A Systematic Introduction". Springer-Verlag, 1995.
- Rosen, B. E.** (1996). Ensemble learning using decorrelated neural networks. *Connection Science: Journal of Neural Computing, Artificial Intelligence and Cognitive Research*, 8, 373-383.
- Rumelhart, D.E., y McClelland, J.L.** (eds) (1986). *Parallel distributed processing*. Vol. 1. Cambridge, Mass.: MIT Press.
- Rumelhart, D.E., Hinton, G.E. y Williams, R.J.** (1986). Learning internal

representations by error propagation. En Rumelhart, D.E., McClelland, J.L. and the PDP Research Group (eds.) (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.

Rumelhart, D.E., Durbin, R., Golden, R. y Chauvin, Y. (1995). Backpropagation: The basic theory. En Y. Chauvin y D.E. Rumelhart (eds.), *Backpropagation: Theory, architectures, and applications*. LEA: Hillsdale, N.J.

Russell D. Reed, Robert J. II Marks, Robert J. MARKS Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks by Publisher: Bradford Book; (March 26, 1999)

Sarle, W.S. (1994). Neural networks and statistical models. Proceedings of the 19th Annual SAS Group conference, Cary, NC. pps. 1538-1550.

Sarle, W.S. (1998). Neural network FAQ. Acceso FTP, parts 1 to 7. Acceso FTP: <ftp://ftp.sas.com/pub/neural/FAQ.html>

Singh, S. (1997). Quantitative classification of conversational language using artificial neural networks. *Aphasiology*. 11(9): 829-844.

Skapura David M. Building Neural Networks. Addison-Wesley Publishing Company (1995).

Smith, M. (1993). Neural networks for statistical modelling. Van Nostrand Reinhold. N.Y.

Soucek B. e IRIS group (1991). Neural and Intelligent Systems Integration. John Wiley & Sons, Nueva York.

Stigler Stephen M., The History of Statistics: The measurement of Uncertainty before 1990, Belknap, Cambridge, Massashusetts, 1986.

Tabachnick, B.G. y Fidell, L.S. (1989). *Using multivariate statistics*. N.Y.: Harper & Row.

Tarassenko Lionel A Guide to Neural Computing Applications Publisher: Butterworth-Heinemann; (January 30, 1998)

Tatsuoka, M.M. (1989). Multivariate analysis (2a ed.). Macmillan: N.Y.

Thrun, S., Mitchell, T. y Cheng, J. (1991). The MONK's comparison of learning algorithms. Introduction and survey. En S. Thrun, J. Bala, E. Bloedorn e I. Bratko (eds.), The MONK's problem. A performance comparison of different learning algorithms. Pittsburg: Carnegie-Mellon Univ.

Tome, M., and H. E. Burkhardt. 1989. Distance dependent competition measures for predicting growth of individual trees. *For. Sci.* 35: 816-831.

Tsoukalas Lefteri H. Uhrig R. E. Fuzzy and Neural Approaches in Engineering. Joh Wiley & Sons Inc.(1997).

Valles Gándara A.G., Torres Rojo J.M., Velásquez Mtz. A., Rodríguez Franco C., Relación de nueve índices de competencia con el crecimiento en diámetro de *Pinus Cooperi* Blanco. *Agrociencia* Vol. 32, Num. 3, Julio-Septiembre 1998.

Valles Gándara A.G., Gonzáles Laredo R.F., Gallegos Infante A., Torres Rojo J.M., Nívar Chaidez J.J., Rocha Fuentes M. Nuevos índices de competencia independientes de la distancia para predecir el crecimiento en diámetro y altura del pinus cooperi blanco. RECURSOS FORESTALES_AGROFAZ VOLUMEN 5 NÚMERO 1, 2005.

Valentine, H. 1988. *A carbon balance model of stand growth: A derivation employing pipe-model theory and the self-thinning rule*. *Ann. Bot.* 62: 389-396.

Valles Gándara A. G., Islas Gutiérrez F. Sistema de ecuaciones del simulador del crecimiento de maderable para la región de San Dimas, Dgo. INIFAP-SAGAR publicación especial num. 14 febrero 2000.

- Wasserman, P.D.** (1989). *Neural computing: Theory and practice*. Van Nostrand Reinhold. N.Y.
- Waterman, D. A.** (1985), *A Guide to Expert Systems*. Addison-Wesley, Reading, MA.
- Watson, D. J.** 1958. Factors limiting production. In *The biological productivity of Britain*. London, Institute of Biology p. 25-32.
- Watson, D. J.** 1956. Leaf growth in relation to crop yield. In *The growth of leaves*. London, Butterworth, p. 178-194.
- Weiss, S. M. and Kulikowski, C. A.** (1984), *A Practical Guide to Designing Expert Systems*. Rowman and Allanheld, Totowa, N.J.
- Weiss, S.M. Y Kulikowsky, C.** (1991), *Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning, and Expert Systems*. Morgan Kaufmann Publishers.
- Werbos P.** *Beyond Regression: New tools for prediction and analysis in the behavioral sciences*. Doctoral thesis, Harvard, Cambridge, MA, August 1974.
- Werbos, P.** (1993). Neurocontrol and Elastic Fuzzy Logic: capabilities, concepts, and applications. En: *IEEE Transactions on industrial electronics*. Vol 40. No 2.
- Weisbuch Gérard,** (1991). *Complex Systems Dynamics* (Addison-Wesley, Reading, Mass.).
- Widrow B. and M. E. Hoff.** Adaptive switching circuits. *1960 IRE WESCON Convention Record*. New York IRE, pages 96–104, 1960.
- White, H.** (1994). Parametric statistical estimation with artificial neural networks. En Smolensky, Mozer y Rumelhart (eds.), *Mathematical perspectives on neural networks*. Hillsdale, NJ: Erlbaum.
- Whittaker, R.** (1975) *Communities and Ecosystems*. New York: MacMillan Press, 2nd edition.
- Zurada J. M.,** *Introduction to Artificial Neural Systems*, West Publishing, 1992.

ANEXO

2.3.2. Estimación de la velocidad de crecimiento en diámetro con una Red Neuronal Backpropagation.

Valores de los gráficos presentados en tablas.

2.3.2.1. Predicción de crecimiento en Diámetro

Comparativo de predicción de la velocidad de crecimiento en diámetro del
Pinus Cooperi Blanco
cm/ 5 años

Valor de la muestra de validación	Valor Pronosticado
3.20	4.0669
3.40	6.0583
3.10	2.7702
2.50	4.5814
3.80	1.5565
2.70	4.0162
2.20	1.5976
1.90	2.4699
0.70	15.6831
1.50	0.0822
1.70	4.6951
0.60	0
1.30	0
1.60	0
2.10	0.1260
1.00	0.2392
0.90	1.4001
1.10	4.1960
0.60	5.6851
1.20	3.8686
1.0	2.7854

2.3.2.2. Predicción de crecimiento en Diámetro

Comparativo de predicción de la velocidad de crecimiento en diámetro
del
Pinus Cooperi Blanco
cm/ 5 años

Valor de la muestra de validación	Valor Pronosticado
3.20	4.2063
3.40	3.3004
3.10	2.0103
2.50	2.4921
3.80	3.3618
2.70	3.0056
2.20	1.7240
1.90	1.6246
0.70	0
1.50	0
1.70	0.0318
0.60	0.6790
1.30	0.3212
1.60	0.7473
2.10	0
1.00	0.5428
0.90	1.8671
1.10	0
0.60	0.6237
1.20	0.9712
1.0	0.7786

2.3.2.3. Predicción de crecimiento en Diámetro

Comparativo de predicción de la velocidad de crecimiento en diámetro
del
Pinus Cooperí Blanco
cm/ 5 años

Valor de la muestra de validación	Valor Pronosticado
3.20	2.7647
3.40	3.3630
3.10	2.7112
2.50	3.1527
3.80	4.5049
2.70	2.3000
2.20	2.2009
1.90	1.7312
0.70	1.2968
1.50	2.6403
1.70	0.4325
0.60	0.6436
1.30	0.9445
1.60	1.1304
2.10	1.1545
1.00	1.4650
0.90	0.1355
1.10	1.0425
0.60	0.8805
1.20	0.5998
1.0	0.9302

2.3.2.4. Predicción de crecimiento en Diámetro.

Comparativo de predicción de la velocidad de crecimiento en diámetro del
Pinus Cooperi Blanco
cm/ 5 años

Valor de la muestra de validación	Valor Pronosticado
3.20	3.3258
3.40	3.5467
3.10	2.9636
2.50	2.6528
3.80	3.7553
2.70	2.4557
2.20	2.2266
1.90	1.9606
0.70	0.7371
1.50	1.7041
1.70	1.5542
0.60	0.7265
1.30	1.3409
1.60	1.7309
2.10	2.0870
1.00	1.0600
0.90	1.0572
1.10	0.9765
0.60	0.6185
1.20	1.2131
1.0	0.9641

2.3.2.5. Predicción de crecimiento en Diámetro.

Comparativo de predicción de la velocidad de crecimiento en diámetro del
Pinus Cooperi Blanco
cm/ 5 años

Valor de la muestra de validación	Valor Pronosticado
3.20	3.2822
3.40	5.0818
3.10	3.1429
2.50	2.6894
3.80	3.3189
2.70	3.1254
2.20	1.8337
1.90	2.2226
0.70	3.6433
1.50	2.8653
1.70	2.6125
0.60	1.9438
1.30	1.3291
1.60	1.7543
2.10	2.4484
1.00	1.0869
0.90	1.3190
1.10	0.9018
0.60	0
1.20	0.9880
1.0	0.9879

2.3.2.6. Predicción de crecimiento en Diámetro.

Comparativo de predicción de la velocidad de crecimiento en diámetro
del
Pinus Cooperí Blanco
cm/ 5 años

Valor de la muestra de validación	Valor Pronosticado
3.20	2.7692
3.40	3.4242
3.10	2.9690
2.50	2.4326
3.80	5.2585
2.70	1.8747
2.20	2.4962
1.90	1.4536
0.70	0.1864
1.50	4.1687
1.70	1.2483
0.60	1.8864
1.30	1.0899
1.60	1.4894
2.10	0.4479
1.00	1.2838
0.90	0.7997
1.10	1.1413
0.60	0.4021
1.20	1.1644
1.0	0.7001

2.3.2.7. Predicción de crecimiento en Diámetro.

Comparativo de predicción de la velocidad de crecimiento en diámetro
del
Pinus Cooperí Blanco
cm/ 5 años

Valor de la muestra de validación	Valor Pronosticado
3.20	2.0338
3.40	3.8074
3.10	2.5771
2.50	2.2953
3.80	4.0741
2.70	1.9212
2.20	2.9313
1.90	0.6808
0.70	0
1.50	1.7126
1.70	2.1947
0.60	1.3730
1.30	1.1245
1.60	0.9161
2.10	0.4220
1.00	1.1786
0.90	1.4302
1.10	0.0580
0.60	0.4708
1.20	1.9498
1.0	0.1945

2.3.3. Estimación de la velocidad de crecimiento en altura por medio de una red neuronal backpropagation.

Valores de los gráficos presentados en tablas.

2.3.3.1. Predicción de crecimiento en altura.

Comparativo de predicción de la velocidad de crecimiento en altura
del
Pinus Cooperi Blanco
m/ 5 años

Valor de validación	Valor pronosticado
2.00	1.6282
1.00	1.4551
1.50	0.6310
3.00	4.0132
1.80	2.2019
1.00	2.0734
3.00	3.6800
1.60	0.7083
3.00	2.7783
2.50	1.7546
2.50	0
3.50	2.5121
3.50	2.6526
2.50	1.7875
1.50	1.8058
2.50	3.0864
2.50	1.2654
1.50	0.4230
1.00	0
1.00	2.1023
1.00	0
1.50	1.2810
2.60	2.4287
2.00	3.3389
1.00	1.9813
1.00	1.9024
1.00	0.2257
1.50	0.9306
1.00	1.7732
1.50	2.2619

2.3.3.2. Predicción de crecimiento en altura.

Comparativo de predicción de la velocidad de crecimiento en altura
del
Pinus Cooperi Blanco
m/ 5 años

Valor de validación	Valor pronosticado
2.00	1.7699
1.00	2.0023
1.50	1.3393
3.00	2.8568
1.80	2.1973
1.00	1.4693
3.00	2.0369
1.60	1.2306
3.00	3.5000
2.50	2.4046
2.50	5.8310
3.50	2.5306
3.50	2.3684
2.50	1.3226
1.50	1.2343
2.50	1.6318
2.50	0.6716
1.50	0.9692
1.00	0
1.00	2.8473
1.00	0.3619
1.50	1.2599
2.60	3.5157
2.00	3.0653
1.00	2.3252
1.00	0.4678
1.00	1.3357
1.50	1.3393
1.00	1.9219
1.50	4.0956

2.3.3.3. Predicción de crecimiento en altura.

Comparativo de predicción de la velocidad de crecimiento en altura
del
Pinus Cooperi Blanco
m/ 5 años

Valor de validación	Valor pronosticado
2.00	1.4249
1.00	2.8516
1.50	1.2651
3.00	4.0708
1.80	2.4089
1.00	0.9966
3.00	2.5638
1.60	2.0460
3.00	5.6063
2.50	3.2587
2.50	3.2956
3.50	3.4431
3.50	4.0210
2.50	1.1821
1.50	1.6536
2.50	1.9789
2.50	1.2855
1.50	4.5438
1.00	1.8876
1.00	0
1.00	0.3870
1.50	1.0048
2.60	0.6827
2.00	2.9551
1.00	3.7700
1.00	1.5269
1.00	0.1021
1.50	1.0106
1.00	2.1585
1.50	2.5423

2.3.3.4. Predicción de crecimiento en altura.

Comparativo de predicción de la velocidad de crecimiento en altura
del
Pinus Cooperi Blanco
m/ 5 años

Valor de validación	Valor pronosticado
2.00	1.8023
1.00	1.0338
1.50	1.5965
3.00	2.9197
1.80	1.8168
1.00	1.1812
3.00	2.8768
1.60	1.7337
3.00	3.1084
2.50	2.5755
2.50	2.6582
3.50	3.2756
3.50	3.4153
2.50	2.3806
1.50	1.3997
2.50	2.4126
2.50	2.3331
1.50	1.5684
1.00	0.8869
1.00	1.0169
0.50	0.3945
1.50	1.4730
2.60	2.7559
2.00	2.1739
1.00	1.1791
1.00	0.8076
1.00	0.9317
1.50	1.5753
1.00	1.1636
1.50	1.6522

2.3.3.5. Predicción de crecimiento en altura.

Comparativo de predicción de la velocidad de crecimiento en altura
del
Pinus Cooperi Blanco
m/ 5 años

Valor de validación	Valor pronosticado
2.00	1.1895
1.00	1.3289
1.50	2.6434
3.00	4.2976
1.80	2.3141
1.00	0.1391
3.00	2.9463
1.60	0.6049
3.00	2.9347
2.50	2.9281
2.50	4.4993
3.50	3.5942
3.50	3.4879
2.50	3.3025
1.50	1.4847
2.50	1.8105
2.50	1.7777
1.50	3.0067
1.00	0.5213
1.00	1.3668
0.50	0.5914
1.50	2.2509
2.60	3.7620
2.00	0.6762
1.00	0
1.00	1.2272
1.00	1.2535
1.50	1.7052
1.00	0.9746
1.50	1.2982

2.3.3.6.) *Predicción de crecimiento en altura.*

Comparativo de predicción de la velocidad de crecimiento en altura
del
Pinus Cooperi Blanco
m/ 5 años

Valor de validación	Valor pronosticado
2.00	3.1966
1.00	1.7099
1.50	2.1833
3.00	3.4238
1.80	2.2220
1.00	0.5668
3.00	3.3360
1.60	0.3801
3.00	2.6672
2.50	3.0881
2.50	5.8620
3.50	1.8039
3.50	0
2.50	3.6628
1.50	1.5669
2.50	0.7386
2.50	2.9786
1.50	2.9786
1.00	0.5574
1.00	0
1.00	0.8867
1.50	2.4714
2.60	4.3966
2.00	1.5304
1.00	0
1.00	2.4749
1.00	1.4679
1.50	1.5613
1.00	1.1082
1.50	1.7150