



COLEGIO DE POSTGRADUADOS
INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO FORESTAL

**PREDICCIÓN DE PUNTOS DE CALOR COMO
PRECURSORES DE INCENDIOS FORESTALES**

ADALBERTO GONZÁLEZ RAMOS

T E S I S

PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS

MONTECILLO, TEXCOCO, ESTADO DE MÉXICO

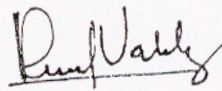
2008

La presente tesis titulada: **Predicción de puntos de calor como precursores de incendios forestales**, realizada por el alumno: **Adalberto González Ramos**, bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

**MAESTRO EN CIENCIAS
FORESTAL**

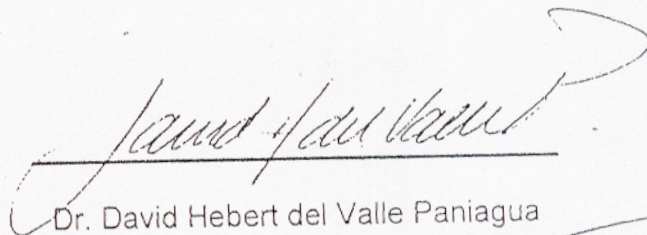
CONSEJO PARTICULAR

CONSEJERO:



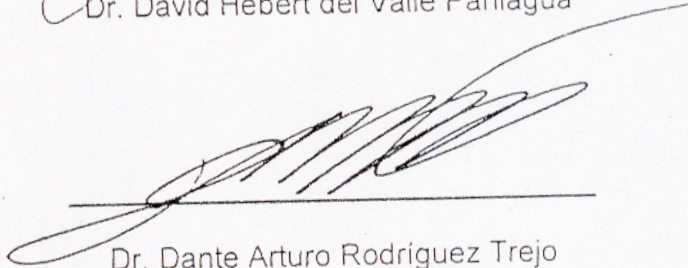
Dr. José René Valdez Lazalde

ASESOR:



Dr. David Hebert del Valle Paniagua

ASESOR:



Dr. Dante Arturo Rodríguez Trejo

Montecillo, Texcoco, Estado de México; a Julio de 2008.

PREDICCIÓN DE PUNTOS DE CALOR COMO PRECURSORES DE INCENDIOS FORESTALES

Adalberto González Ramos, M. C.

Colegio de Postgraduados, 2008.

Con el propósito de determinar la relación entre factores humanos y factores geográficos con la presencia de puntos de calor que pudieran constituir posibles incendios forestales, se definió una metodología para evaluar tres tipos de arquitecturas de redes neuronales backpropagation, empleándose como datos de entrada cuatro modelos que resultaron de combinar seis factores humanos con siete factores geográficos. Como datos de salida se usaron los puntos de calor detectados en el territorio que comprende las zonas susceptibles a incendios forestales del Estado de México durante el período del año 2000 al 2005 para las etapas de entrenamiento y prueba. Asimismo, datos del año 2006 se destinaron para la comprobación del funcionamiento de la red. Se usó información cartográfica proveniente del sensor AVHRR así como de otras fuentes. Para una muestra de 2,800 observaciones tomada de las 8,938 originales que conforman las celdas de vegetación susceptible a incendios forestales, se determinó una red de tipo Jordan-Elman, con la que se obtuvo un coeficiente de correlación (R^2) de 0.8852 empleando el modelo que incluyó la totalidad de factores humanos y geográficos. Sin embargo no se obtuvo el resultado esperado para todos los datos-objetivo, el cual solo alcanzó un valor de R^2 de 0.2181, y la aplicación a los datos del 2006 presentó una R^2 de solamente 0.1052. Se determinó con base en los resultados de la muestra, que factores humanos en combinación con factores geográficos sí pueden emplearse para determinar la incidencia de puntos de calor cuando las condiciones y características físicas del sitio de estudio son más homogéneas entre sí.

Palabras clave: algoritmos genéticos, redes neuronales retroalimentadas, predicción de incendios forestales, modelación, puntos de calor.

PREDICTION OF POINTS OF HEAT LIKE PRECURSORS OF WILDFIRES

Adalberto González Ramos, M.C.

Colegio de Postgraduados, 2008.

With the purpose of determining the existent relationship between human and geographical factors with the presence of points of heat than they could constitute possible wildfires, was defined the methodology to evaluate three types of architectures of nets neuronales backpropagation using as input data four models than they were of combining six human factors with seven geographical factors. As output data the points of heat were used detected in the susceptible areas to wildfires in the State of Mexico, during the period of the year 2000 at the 2005 for the stage of training and it proves. Also, data of the year 2006 were employees for the confirmation of the operation of the net. It was used cartographic information coming from the sensor AVHRR and other sources. For a taken sample of 2,800 observations of the 8,938 originals that they conform the cells of susceptible vegetation to wildfires, a net Jordan-Elman was determined, a coefficient of correlation (R^2) 0.8852 was obtained for the sample using the entirety of human factors and geographical. However the prospective result was not obtained for all the data of the process, the one which alone it reached a value of R^2 of 0.2181, and the application to the data of the year 2006 only a R^2 of 0.1052 it presented. It was determined based on the results of the sample that human factors in combination with geographical factors they can be used to determine the incidence of points of heat when the conditions and physical characteristics of the study place are more homogeneous to each other.

Keywords: genetic algorithm, backpropagation neural nets, wilfires prediction, modeling, fire points.

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico y la confianza recibida para la realización de este grado.

A los integrantes de mi Consejo Particular por su esfuerzo, confianza y demasiada paciencia brindada en estos años de formación.

A la Comisión Nacional para la Biodiversidad y muy especialmente al personal de la Delegación de la Procuraduría Federal de Protección al Ambiente en el Estado de México por el material proporcionado para la realización de este trabajo.

A mis innumerables amigos, compañeros, personal docente y administrativo, que deambulan como almas sin pena entre los corredores y pasillos del Programa Forestal, del Colegio de Postgraduados y de la Universidad Autónoma Chapingo.

A la Educación Pública, que en este país ha permitido a miles de personas superarse aún a costa de múltiples esfuerzos y privaciones, en aras de alcanzar una satisfacción personal.

DEDICATORIA

A mi finado Padre, el Comandante del H. Cuerpo de Bomberos de Naucalpan, Adalberto González Ángeles, mi ejemplo.

A toda mi familia, hermanos, tíos y primos; y muy especialmente a mi madre, Alicia Ramos Reyes por su apoyo incondicional y paciencia a toda prueba que me tuvo en la realización de mis estudios de Maestría.

A Gris, esposa y compañera, quien es mi razón para superarme todos los días.

CONTENIDO

1.	INTRODUCCIÓN.....	1
2.	OBJETIVOS.....	4
2.1.	Objetivo general	4
2.2.	Objetivos particulares	4
3.	HIPÓTESIS.....	4
4.	REVISIÓN DE LITERATURA	5
4.1.	Caracterización de los incendios forestales	5
4.1.1.	Los incendios forestales en el Estado de México	6
4.1.2.	Variables que definen el riesgo de incendio forestal	9
4.1.3.	Función ecológica del fuego.....	11
4.1.4.	Causas de los incendios forestales.....	13
4.1.5.	Condiciones socioeconómicas	15
4.1.6.	Conflictos en la zona de transición rural	15
4.1.7.	Conflictos en la zona de transición urbana	18
4.1.8.	Conflictos no relacionados con el uso de la tierra.....	18
4.2.	Sistemas de información geográficos (SIG).....	19
4.3.	Estudios relacionados en detección y predicción de incendios forestales.....	20
4.4.	Caracterización de las redes neuronales artificiales (RNA)	27
4.4.1.	Aspectos generales.....	27
4.4.2.	Definición de redes neuronales artificiales.....	29
4.4.3.	Características de las redes neuronales artificiales.....	29
4.4.4.	Redes neuronales basadas en algoritmos genéticos.....	31
4.4.4.1.	PNN (Probabilistic Neural Network)	33
4.4.4.2.	GRNN (General Regression Neural Net).....	34
4.4.5.	Ventajas de las redes neuronales artificiales	36
4.4.6.	Estructura de las redes neuronales artificiales.....	37
4.4.7.	Diseño de la red neuronal artificial	40
4.4.8.	Diseño de la arquitectura de la red	42
4.4.9.	Funciones de activación (FA).....	44
4.4.10.	Modo de actualización de pesos (MAP).....	47
4.4.11.	Neuronas en la capa oculta.....	50
4.4.12.	Calibración	51
4.4.13.	Secuencia para la entrada de los datos	51
4.4.14.	Selección de variables de entrada	52
4.5.	Criterios para la evaluación de redes	53
4.5.1.	Error mínimo promedio (EMP)	53
4.5.2.	Coeficiente de determinación múltiple (R^2).....	55
4.5.3.	Coeficiente de correlación lineal de Pearsons (CCr)	55
4.6.	Procesamiento de imágenes del sensor AVHRR.....	57
5.	MATERIALES Y EQUIPO.....	60
5.1.	Caracterización de la zona de estudio	60
5.2.	Obtención del material cartográfico.....	66
5.2.1.	Información obtenida de la CONABIO	66
5.2.2.	Material obtenido de la PROFEPA.....	68

5.3.	Software y equipo utilizado.....	69
6.	METODOLOGÍA.....	71
6.1.	Preprocesamiento del material cartográfico.....	71
6.1.1.	Límite del Estado de México (EDO_MEX).....	72
6.1.2.	Distancia a cabeceras municipales (CAB_MUN).....	72
6.1.3.	Distancia a la carretera más cercana (DIST_CARR).....	73
6.1.4.	Mapa de zonas de campamento (ZC).....	73
6.1.5.	Áreas naturales protegidas (ANP).....	74
6.1.6.	Aprovechamientos forestales (APROV).....	74
6.1.7.	Líneas de conducción eléctrica (LIN_COND).....	75
6.1.8.	Vegetación susceptible a incendios (US_SUELO).....	75
6.1.9.	Pendientes categorizadas (PEND_CAT).....	76
6.1.10.	Altitudes categorizadas (ALT_CAT).....	77
6.1.11.	Régimen de humedad del suelo (REHUSUEL).....	77
6.1.12.	Temperatura media (TEMP_MED).....	78
6.1.13.	Precipitación media (PREC_MED).....	78
6.1.14.	Calidad ecológica (CAL_ECO).....	78
6.2.	Adaptación de los datos para la RNA.....	79
6.3.	Selección de las variables de entrada y salida.....	83
6.4.	Definición de cuatro modelos para entrenar la RNA.....	86
6.5.	Alimentación de la base de datos.....	95
6.6.	Definición de la muestra para el entrenamiento.....	97
7.	RESULTADOS.....	99
7.1.	Determinación de las variables para el MODELO 1 con GRNN.....	99
7.2.	Aplicación de la muestra.....	101
7.2.1.	Redes lineales.....	102
7.2.2.	Redes recurrentes (Jordan-Elman Nets).....	106
7.2.3.	Redes múltiples (Ward nets).....	110
7.3.	Aplicación de la mejor red en los datos de 2006.....	119
7.4.	Discusión de resultados.....	121
8.	CONCLUSIONES Y RECOMENDACIONES.....	123
8.1.	Conclusiones.....	123
8.2.	Recomendaciones.....	125
9.	LITERATURA CITADA.....	127

ÍNDICE DE FIGURAS

Figura 4.1. Dos neuronas biológicas interconectadas	30
Figura 4.2. Equivalente artificial de dos neuronas interconectadas.....	31
Figura 4.3. Esquema del funcionamiento de la arquitectura GRNN	36
Figura 4.4. Funcionamiento de una neurona artificial.....	38
Figura 4.5. Componentes de una red neuronal backpropagation.....	39
Figura 4.6. Búsqueda del mejor diseño de arquitectura en un espacio 3D	41
Figura 4.7. Funciones en una red backpropagation	42
Figura 4.8. Función logística.....	45
Figura 4.9. Función gaussiana.	45
Figura 4.10. Función gaussiana complementaria.....	46
Figura 4.11. Función tangente hiperbólica.....	47
Figura 5.1. Ubicación de la zona de estudio.....	60
Figura 5.2. Imagen satelital en falso color del Estado de México.....	61
Figura 5.3. Cobertura por tipo de uso de suelo	66
Figura 5.4. Aprovechamiento potencial por tipo de bosque.....	66
Figura 6.1. Transformación a código ASCII para la alimentación de la red.....	80
Figura 6.2. Diagrama de flujo para el procesamiento de la base de datos de la red.....	81
Figura 6.3. Puntos de calor diurnos y nocturnos del 2000 al 2006 en vegetación susceptible	88
Figura 6.4. Mapa de cabeceras municipales (CAB_MUN)	88
Figura 6.5. Mapa de distancia a la carretera principal más cercana (DIST_CARR).....	89
Figura 6.6. Mapa de zonas de campamento (ZC)	89
Figura 6.7. Mapa de áreas naturales protegidas (ANP)	90
Figura 6.8. Mapa de aprovechamientos forestales autorizados (APROV)	90
Figura 6.9. Mapa de líneas de conducción eléctrica (LIN_COND).....	91
Figura 6.10. Mapa de vegetación susceptible a incendios forestales (US_SUELO).....	91
Figura 6.11. Mapa de pendientes categorizadas (PEND_CAT)	92
Figura 6.12. Mapa de altitudes categorizadas (ALT_CAT).....	92
Figura 6.13. Mapa del régimen de humedad en el suelo (REHUSEUL).....	93
Figura 6.14. Mapa de temperatura media anual (TEM_MED).....	93
Figura 6.15. Mapa de precipitación media anual (PREC_MED).....	94
Figura 6.16. Mapa de calidad ecológica (CAL_ECO).....	94
Figura 6.17. Datos que intervienen en la alimentación de la red.....	96
Figura 6.18. Conformación de la muestra de entrenamiento.....	98
Figura 7.1. Resultados del entrenamiento de la mejor red para los datos-objetivo.....	117

INDICE DE GRÁFICAS

Gráfica 4.1. Error Promedio vs. Épocas/Intervalos evaluados del conjunto de datos de entrenamiento (izquierda) y del conjunto de datos de prueba (derecha) en NeuroShell 2© durante un aprendizaje normal.....	54
Gráfica 4.2. Gráficas de Error Promedio vs. Épocas/Intervalos evaluados del conjunto de datos de entrenamiento (izquierda) y del conjunto de datos de prueba (derecha) en NeuroShell 2© donde no hay aprendizaje.....	54
Gráfica 7.1. Error mínimo promedio con FA [-1,1], obtenido en el procedimiento GRNN para MODELO 1 en la muestra de 2,800 observaciones.....	101
Gráfica 7.2. Valores de R^2 obtenidos para MODELO 1 en arquitecturas de redes lineales.....	103
Gráfica 7.3. Valores de R^2 obtenida para MODELO 2 en arquitecturas de redes lineales.....	104
Gráfica 7.4. Valores de R^2 obtenidos para MODELO 3 en arquitecturas de redes lineales.....	105
Gráfica 7.5. Valores de R^2 obtenidos para MODELO 4 en arquitecturas de redes lineales.....	106
Gráfica 7.6. Valores de R^2 obtenidos para MODELO 1 en arquitecturas de redes recurrentes.....	107
Gráfica 7.7. Valores de R^2 obtenidos para MODELO 2 en arquitecturas de redes recurrentes.....	108
Gráfica 7.8. Valores de R^2 obtenidos para MODELO 3 en arquitecturas de redes recurrentes.....	109
Gráfica 7.9. Valores de R^2 obtenidos para MODELO 4 en arquitecturas de redes recurrentes.....	110
Gráfica 7.10. Valores de R^2 obtenidos para MODELO 1 en arquitecturas de redes múltiples.....	111
Gráfica 7.11. Valores de R^2 obtenidos para MODELO 2 en arquitecturas de redes múltiples.....	112
Gráfica 7.12. Valores de R^2 obtenidos para MODELO 3 en arquitecturas de redes múltiples.....	113
Gráfica 7.13. Valores de R^2 obtenidos para MODELO 4 en arquitecturas de redes múltiples.....	114
Gráfica 7.14. Mejores valores de R^2 obtenidos para cada arquitectura de red.....	115
Gráfica 7.15. Comparativa de valores de R^2 obtenidos en la muestra y en la aplicación de la red a los datos-objetivo.....	116
Gráfica 7.16. Comportamiento de R^2 con la mejor red para los datos-objetivo.....	118
Gráfica 7.17. Etapa de prueba que muestra el nivel de EMP en relación a las épocas transcurridas en la aplicación de la red a los datos de 2006.....	120

ÍNDICE DE TABLAS

Tabla 4.1. Entidades con mayor número de incendios forestales en 2006.	8
Tabla 4.2. Resoluciones espaciales y espectrales del sensor AVHRR.	58
Tabla 5.1. Principales especies y utilidad por uso de suelo	65
Tabla 5.2. Información digital obtenida de CONABIO.	67
Tabla 5.3. Puntos de calor detectados en el período 2000 a 2006 en el Estado de México.	68
Tabla 5.4. Información digital obtenida de PROFEPA-EDOMEX.	69
Tabla 6.1 Tipo de vegetación y su régimen de fuego.	75
Tabla 6.2. Variables seleccionadas como factores humanos.	84
Tabla 6.3. Variables seleccionadas como factores geográficos.	85
Tabla 6.4. Valores de salida que representan incidencia de puntos de calor.	96
Tabla 6.5. Identificación por tipo de arquitectura backpropagation.	98
Tabla 7.1. Resultados del proceso GRNN para la determinación del MODELO 1.	99
Tabla 7.2. Valores obtenidos en la aplicación de la mejor arquitectura.	118
Tabla 7.3. Contribución relativa para cada variable en la mejor arquitectura.	119
Tabla 7.4. Valores obtenidos en la aplicación de la red a los datos de 2006.	120

1. INTRODUCCIÓN

Además de diversos bienes directos, el bosque aporta una serie de beneficios indirectos para el hombre y el ambiente, entre los que destacan el hábitat de numerosas especies animales y vegetales, la producción de oxígeno, el almacenamiento y fijación de carbono y la disminución del bióxido de carbono de la atmósfera y por ende, la reducción del efecto invernadero en el mundo. Por otra parte, es fuente de riqueza, bienestar y una de las bases fundamentales del paisaje natural que identifica y proporciona valor estético al territorio. Contribuye además a la conservación de la humedad ambiental controlando el ciclo del agua, favoreciendo la cantidad y calidad del recurso hídrico. Por estos y otros beneficios, el mantenimiento de nuestros bosques adquiere importancia fundamental para mantener una buena calidad de vida. Así, el fuego cumple una función vital en el mantenimiento de muchos ecosistemas y de las comunidades que dependen de ellos. No obstante, la alta incidencia de incendios forestales como parte de regímenes de fuego alterados, disminuye de forma considerable la superficie de los bosques, con los perjuicios que ello acarrea. Este tipo de acontecimientos provoca una gran diversidad de impactos en los ámbitos económico, ecológico, político, social, operativo y legal. (Rodríguez, 1996; Shlisky *et al.* 2007).

Se puede destacar el desequilibrio de los flujos y ciclos naturales, el incremento de la erosionabilidad de los suelos y la consecuente pérdida de los mismos, disminución de la recarga de acuíferos y de los recursos hídricos y por ende, aumento de los riesgos de

avenidas de agua, así como la contaminación atmosférica e incremento del efecto invernadero.

Si bien el efecto del fuego en los bosques suele considerarse como un impacto negativo, es de resaltar que el fuego cumple una función vital en el mantenimiento de muchos ecosistemas y de las comunidades que dependen de ellos, es especial, aquellos que son dependientes del fuego, cuya supresión o introducción inadecuada de este fenómeno puede perturbar el sistema de manera importante. Un buen manejo del fuego puede producir más beneficios al ecosistema que deterioros. Esto último sucede cuando son alterados los regímenes naturales del fuego, principalmente como consecuencia de diversas actividades antropogénicas como se ha reconocido ampliamente a nivel mundial.

Debido a la problemática que significa la alta incidencia de incendios forestales en los que se implica la acción del hombre en las zonas boscosas de nuestro país, es de importancia llevar a cabo estudios que determinen formas efectivas de control y prevención en aquellos lugares donde el fuego signifique un mayor riesgo en la vegetación de lugares destinados para aprovechamiento, recreación, conservación y actividades de trabajo.

Para caracterizar de forma cualitativa y cuantitativa este fenómeno, es necesario aplicar conocimientos surgidos a partir de bases de datos geográficos que contengan en sus campos aquella información de importancia para la evaluación física, social, económica y cultural del factor humano en aquellas zonas donde sus actividades se encuentren presentes. Mediante el uso de los Sistemas de Información Geográfica (SIG) y de

técnicas de Percepción Remota (PR) es posible que a partir del manejo de imágenes de satélite y de la sobreposición de diferentes capas de información (datos digitales topológicos sobre una característica en particular dispuestos espacialmente en una imagen) se generan imágenes que faciliten la toma de decisiones sobre una región en particular.

En el presente estudio se evalúan factores humanos y geográficos relacionados a la génesis de incendios forestales en el territorio que comprende el Estado de México. Para ello, fueron usadas herramientas de Sistemas de Información Geográficos (SIG), técnicas de Percepción Remota (PR) y se aplicaron distintas arquitecturas de Redes Neuronales Artificiales (RNA) a fin de llevar a cabo el procesamiento y la evaluación de dichos factores.

2. OBJETIVOS

2.1. Objetivo general

Determinar la relación existente entre diversos factores humanos y geográficos con respecto a la incidencia de puntos de calor como precursores de incendios forestales en el Estado de México.

2.2. Objetivos particulares

- Definir la metodología para evaluar puntos de calor mediante redes neuronales artificiales usando factores humanos y geográficos.
- Determinar si existe relación entre la actividad humana con los puntos de ignición de posibles incendios forestales.
- Aplicar diferentes arquitecturas de redes neuronales a fin de determinar cuál es la que mejor modela la actividad humana en la génesis de incendios forestales en el Estado de México.

3. HIPÓTESIS

Es posible establecer una relación ente los puntos de ignición (detonadores de incendios forestales) con características humanas y geográficas para predecir su ocurrencia.

4. REVISIÓN DE LITERATURA

4.1. Caracterización de los incendios forestales

En Shlisky *et al.* (2007) se afirma que el fuego es un proceso natural que ha jugado un papel importante en la formación del ambiente y en el mantenimiento de la biodiversidad en todo el mundo. Debe tomarse en cuenta que los beneficios e impactos del fuego son amplios. Se menciona que la mayoría de los hábitats terrestres dependen del fuego para su sostenibilidad ecológica.

Un incendio forestal es la propagación libre y no programada del fuego sobre la vegetación en los bosques selvas y zonas áridas y semiáridas. El fuego es la liberación y desprendimiento de energía en forma de luz y calor producido por la combustión de vegetación forestal cuya ignición no estaba prevista, lo que obliga a realizar su extinción (CONAFOR, 2001).

Suelen ser fuegos de origen natural o inducidos que queman la vegetación de un bosque ó pastizales. Se distinguen tres tipos de incendios forestales: los fuegos subterráneos, que queman la capa de humus del suelo del bosque; los fuegos de superficie, que queman el sotobosque y los residuos superficiales; y los fuegos de copa, que avanzan a través de las copas de los árboles. Es frecuente que se produzcan dos y hasta tres de estos tipos de incendio al mismo tiempo (Pyne *et al.*, 1996).

Las condiciones climatológicas, del ambiente, así como características del medio físico influyen en la susceptibilidad que un área determinada presenta frente al fuego; factores

como la pendiente, la temperatura, la humedad y la cantidad de lluvia determinan la velocidad y el grado al que se seca el material inflamable y, por tanto, la combustibilidad del bosque. El viento tiende a acelerar la desecación y a aumentar la severidad de los incendios avivando la combustión. Estableciendo la correlación entre los diversos elementos climatológicos y la susceptibilidad de ignición de los residuos de ramas y hojas, es posible predecir el riesgo de incendio de un día cualquiera en cualquier localidad. En condiciones de riesgo extremo, los bosques pueden cerrarse al público.

Aunque las organizaciones relacionadas con el control del fuego combaten todos los incendios, el fuego debido a causas naturales siempre ha estado presente en el ecosistema. La supresión total de los incendios podría producir cambios indeseables en los patrones de vegetación, además de permitir la acumulación de materiales combustibles, aumentando las posibilidades de que se produzcan incendios catastróficos. En algunos parques y reservas naturales de América del Norte, donde el objetivo es mantener las condiciones naturales, normalmente se deja que los incendios provocados por los rayos sigan su curso bajo una meticulosa vigilancia. Sin embargo, son más las causas de los regímenes de fuego alterados a nivel mundial: el desarrollo urbano, la cría de animales de granja, la ganadería y la agricultura, el uso y supresión del fuego, la extracción de recursos (incluidas la producción de energía, la minería y la tala) y el cambio climático global (Rodríguez, 1996; Shlisky *et al.* 2007).

4.1.1. Los incendios forestales en el Estado de México

De acuerdo al Programa de Prevención y Combate de Incendios Forestales 2006-2007 realizado por el Gobierno Estatal, el Estado de México mantuvo el primer lugar durante el año pasado en cantidad de incendios forestales, seguido por los Estados de

Michoacán, Distrito Federal, Chihuahua y Chiapas. Con base a la información estadística estatal, durante el periodo 2003-2007 se presentaron anualmente 1,342 incendios en promedio, afectando 5,436 ha. (4.1 hectáreas por incendio en promedio). Sólo durante el año 2007 se presentaron en el territorio Estatal 902 incendios forestales que afectaron una superficie total de 2,750.30 ha. Estos incendios corresponden aproximadamente al 17% de los que se han registrado en el país durante el mismo periodo. La superficie estatal con riesgo de presentar incendios forestales es de 894 mil ha., mismas que se distribuyen de la siguiente forma: 558 mil de bosques, 88 mil de selvas, 16 mil de vegetación de zonas áridas, 6 mil de vegetación hidrófila y halófila y 226 mil de áreas perturbadas (PROBOSQUE, 2007).

La temporada de incendios en nuestro país tiene una época bien determinada. Estos suelen presentarse entre diciembre y agosto en diferentes regiones. La incidencia aumenta en los meses de marzo, abril y mayo, aunque suelen existir años atípicos en que la temporada de lluvias comienza antes o se retrasa, afectando con ello la época de incendios. Estos datos son indicativos de que el número de incendios y superficie afectada a nivel nacional no tienen un patrón determinado.

El tiempo que transcurre entre la detección de un incendio forestal y su extinción es también factor a tomar en cuenta. Cuando los incendios son detectados y combatidos oportunamente los daños son menores. De lo contrario, o si el incendio es incontrolable a pesar de ser poco extenso, es decir, un solo incendio que afecta grandes cantidades de terreno (Rodríguez, 1996).

A nivel nacional el Estado de México es el que realiza más actividades preventivas. En el período 2000 a 2005 se elaboraron un total de 1,774 km de brechas cortafuego, 1,281 hectáreas de quemas controladas y se han capacitado a 1,950 elementos para combatirlos. El Estado de México tiene en promedio una afectación de 3.4 ha/incendio en comparación con el 26,3 ha/incendio que es el promedio nacional. Con ello ha alcanzado el 5to. lugar nacional en la eficiencia en el combate de incendios forestales, y el 11avo. lugar en superficie afectada.

Datos estadísticos indican que hasta 1998 (que hasta hoy sigue siendo el año con mayor número de incendios y superficie afectada en nuestro país) en promedio se presentaban en el Estado de México, poco más de 3,500 incendios por año, que afectaron una superficie del orden de 15,000 hectáreas, que en comparación a 2006, se ha logrado disminuir esta incidencia, pero a nivel estatal, continúa siendo uno de los más afectados (Tabla 4.1).

Tabla 4.1. Entidades con mayor número de incendios forestales en 2006.

Entidad federativa	Número de incendios	Superficie afectada (ha)	Superficie afectada por incendio (ha)
Estados Unidos Mexicanos	8,725	243,864.01	27.95
México	1,511	7,072.76	4.68
Distrito Federal	1,078	1,694.06	1.57
Michoacán de Ocampo	1,062	13,175.55	12.41
Chihuahua	1,057	18,504.66	17.51
Jalisco	815	15,810.00	19.4
Puebla	350	3,306.30	9.45
Chiapas	329	10,485.02	31.87
Oaxaca	239	13,096.75	54.8
Tlaxcala	222	945.5	4.26
Baja California	195	13,737.70	70.45
Otros	1,867	146,035.71	78.22

Fuente: CONAFOR, 2006.

De tal forma que el Estado de México en la entidad a nivel nacional con mayor número de incendios forestales, y el undécimo con relación a la superficie afectada.

4.1.2. Variables que definen el riesgo de incendio forestal

Aunque los términos *peligro* y *riesgo* de incendio se usan genéricamente, el término *peligro* se refiere a todo factor que facilita el inicio y la propagación del fuego, como una temperatura alta, humedad relativa baja, combustibles forestales secos, vientos, altas cargas de combustibles forestales, entre otras. A su vez, el *riesgo* de incendio se refiere a todo factor que puede iniciar el fuego directamente, tal como un rayo en seco, una quema agropecuaria, una fogata o una colilla de cigarro.

La aparición de incendios forestales es producto de diversos factores; sin embargo, la manifestación de éstos generalmente obedece a un patrón determinado. Con frecuencia las áreas de más alta posibilidad de incendio se encuentran próximas a caminos, carreteras o poblados, así como en áreas de fuertes pendientes, zonas de pastoreo, lugares de recreo, áreas en aprovechamiento maderable, etc. (CONAFOR, 2007).

El riesgo de un incendio forestal estará definido por las circunstancias permanentes y transitorias que hacen posible su ocurrencia. Cuando se conocen dichas circunstancias, su magnitud y las funciones que las relacionan, se pueden realizar predicciones para obtener un indicador de peligrosidad de que ocurran incendios forestales en áreas boscosas específicas.

De acuerdo a CONAFOR (2006), existen dos tipos de factores de riesgo relacionados de manera directa con la ocurrencia de incendios forestales: los permanentes, que son aquellos que no varían continuamente y que permanecen estables durante largos períodos de tiempo; y los transitorios, que están en cambio constante. También son factores permanentes:

1. La composición de los combustibles
2. Las especies vegetales
3. La topografía.

Estos factores no varían continuamente, sino que permanecen estables durante períodos más o menos largos, por lo que su influencia es constante.

Datos de CONAFOR (2006) indican que en el período entre 1998 y 2005, los factores permanentes (donde se incluyen a las actividades humanas) causantes de incendios forestales destacan las labores agropecuarias (44%), causas intencionales como litigios y/o rencillas (19%), fogatas (12%), fumadores (11%) y en general otras causas que en su conjunto suman el restante 14% del total.

La misma fuente indica que son considerados factores transitorios:

1. Los factores meteorológicos.
2. Temperatura.
3. Humedad relativa.
4. Velocidad y dirección del viento.
5. Precipitación.
6. Cantidad y tipo de combustible.

Generalmente el combustible, es el factor principal que determina si se inicia o no un incendio, así como la dificultad para controlarlo y el tipo de comportamiento. Otros factores que también son considerados en el riesgo de la presencia de un incendio forestal son:

Calor: Una fuente de calor externa es necesaria para iniciar un incendio forestal, en muchos casos, dicha fuente de ignición requiere una temperatura mayor a los 200°C.

Época: Temporada de peligro de incendios en función de los factores meteorológicos y uso del fuego por parte de las actividades humanas: mes, día de la semana, hora del día.

Lugar: Que define la zona de peligro en una región por las vías de transporte; centros de población y por mayor acceso de la población a las áreas forestales (relación entre espacios forestales y los asentamientos humanos) (CONAFOR, 2006).

4.1.3. Función ecológica del fuego

Shlisky *et al.* (2007) menciona que los ecosistemas pueden ser clasificados en términos de su relación con las características del régimen de fuego, como los combustibles, inflamabilidad, igniciones y condiciones de propagación del fuego. Se clasifican en:

- Ecosistemas dependientes del fuego

Aquellas donde las especies han evolucionado en presencia del fuego y donde el fuego es un proceso esencial para conservar la biodiversidad (por ejemplo las

sabanas y los bosques templados de coníferas). Si se excluye el fuego de estos sistemas o si se introduce un fuego ecológicamente inadecuado -con frecuencia, severidad o época del año inadecuadas- se puede alterar sustancialmente estos sistemas.

- Ecosistemas sensibles al fuego

Aquellos donde la mayoría de las especies no ha evolucionado considerablemente en presencia del fuego. Si bien el fuego puede tener una función secundaria en el mantenimiento de la estructura y el funcionamiento natural del ecosistema en sistemas sensibles al fuego, la introducción de un fuego ecológicamente inadecuado puede tener un impacto negativo amplio sobre la biodiversidad (por ejemplo, en bosques húmedos latifoliados tropicales). Demasiado fuego en bosques sensibles al fuego también puede crear un ciclo de retroalimentación negativa, volviéndolos más sensibles al fuego en el futuro y degradando rápidamente los ecosistemas ecológicos más intactos.

- Ecosistemas independientes del fuego

Aquellos que carecen naturalmente de combustible o fuentes de ignición que permitan al fuego actuar como una fuerza evolutiva (por ejemplo, desiertos y tundras).

Asimismo, los ecosistemas, dependientes, sensibles e independientes del fuego pueden clasificarse además en términos de su condición. Por ejemplo, incluso los sistemas independientes del fuego pueden sufrir más incendios a los usos humanos del suelo

que los que habrían ocurrido naturalmente mediante la introducción de especies exóticas invasoras o excesivas igniciones de origen humano.

- Regímenes del fuego intactos

Aquellos en los que las características del régimen del fuego (por ejemplo, frecuencia, severidad, extensión y estación del fuego se encuentran dentro de su rango naturalidad de variabilidad.

- Regímenes de fuego degradado

Son las condiciones en las que los expertos consideran que se encuentran fuera de su rango natural de variabilidad, pero que se estima que son restaurables.

- Regímenes de fuego muy degradado

Son condiciones que están muy fuera de su rango natural de variabilidad, y probablemente no sean restaurables (Shlisky *et al.*, 2007).

4.1.4. Causas de los incendios forestales

Por su origen, las causas de un incendio forestal pueden clasificarse como causas naturales y causas no naturales. Estas se definen de la siguiente forma:

1. Son causas naturales aquellas en las que no existe intervención del hombre en su génesis (Por ejemplo: la caída de un rayo sobre un material combustible, como en el caso de pastizales secos o de un árbol).

2. Son causas no naturales aquellas en las que sí existe por cualquier motivo intervención o actividad derivada de la presencia hombre o alguna de sus actividades. Estas pueden a su vez clasificarse como:

- Accidentales o fortuitos.
- Intencionales.
- Por omisión ó negligencia.

Dentro de las causas naturales, las más comunes son los rayos en seco. En la actualidad, se estima que sólo el cinco porciento de los incendios forestales son atribuibles a dichas causas en México.

Refiriéndose al factor humano, las causas varían en función de sus necesidades y objetivos. Sin embargo, las más frecuentes tienden a estar relacionadas a actividades agropecuarias, como su uso para la limpia de terrenos con vegetación para promover la emisión de rebrotes tiernos para alimentar al ganado y a la apertura de nuevos campos de cultivos, entre otros.

Por otro lado, el uso planificado del fuego en condiciones controladas, pronto se reconoció como de utilidad para lograr objetivos como la reducción de los combustibles y la preparación del suelo para la germinación de las semillas. De esta forma, la política actual de mayor aceptación sobre los incendios forestales es la administración de los incendios, no simplemente el control de los mismos. Esta política aprovecha los efectos benéficos de algunos incendios forestales, al tiempo que lleva a cabo la extinción de éstos (Young, 1991).

Entre las causas consideradas como accidentales, se encuentran aquellas que a su vez se derivan de otras actividades humanas: las chispas de ferrocarril o vehículos, maquinaria o equipo con motor de combustión interna, fugas en gaseoductos u oleoducto en caso de explotar, ruptura de líneas de transmisión eléctrica, chispas de chimeneas de industrias, fábricas o casas en el bosque, entre otras (Rodríguez, 1996).

4.1.5. Condiciones socioeconómicas

Estudios realizados por el Colegio de Ingenieros de Montes (2002) menciona que grandes extensiones de tierra suelen quedar abandonadas por la baja rentabilidad de sembrarlas o por la migración, por lo que su superficie suele ser colonizada por diversas especies de matorrales, que con el tiempo, al secarse pueden constituir fuentes de combustibles para un incendio forestal.

La situación actual en el campo puede caracterizarse por los siguientes hechos:

- Despoblación de las áreas rurales atraída por mayores incentivos en las zonas urbanas.
- Abandono de los usos tradicionales en el área rural.
- Tendencia del uso forestal a la desaparición y/o a reducción de ser empleado como productor de materias primas.
- Tendencia de los usos tradicionales a considerar el aprovechamientos de pastos y leñas, en vez de quedar éstos como productos residuales.
- Tendencia de los usos recreativos a crecer (excursionismo, caza y pesca) debido al crecimiento continuo de la interfaz bosque-terreno urbano.

4.1.6. Conflictos en la zona de transición rural

Asimismo, el Colegio de Ingenieros de Montes (2002) determinó diversas causas que pueden estar relacionadas al ámbito rural y que son similares en el país, que son:

Cambio de uso de suelo forestal

El conflicto sucede por el uso del fuego para eliminación de la vegetación forestal y su posterior sustitución por cultivos agrícolas. En general, las legislaciones establecen la inalienabilidad del terreno forestal. Esta limitación legal al cambio de uso puede generar conflictos en los lugares en los que exista demanda de tierras para cultivo. Pero la tendencia actual es el agotamiento del conflicto por falta de esta demanda y por tanto, el conflicto tiende a desaparecer.

Abandono de tierras

El conflicto se plantea como consecuencia del cese de actividades rurales en las tierras marginales. El abandono da lugar a la invasión por las especies invasoras. Esa invasión genera en breve plazo acumulaciones de combustibles ligeros muy peligrosos, en los que los fuegos iniciados adquieren velocidades e intensidades altas.

La tendencia, por tanto, es la regulación del cambio de uso de agrícola a forestal, con fondos para hacerlo viable. Sin embargo, el proceso de abandono es más rápido y extenso que las actuales políticas y además, el abandono hacen desaparecer a los propietarios que podrían solicitar subsidios al gobierno municipal, estatal o federal.

Pastoreo con empleo del fuego

El uso del fuego para el mantenimiento del pastizal es posible, como una forma de permitir la introducción de animales en sitios donde se lleve a cabo quemas controladas. En general, la legislación de todos los países prohíbe el fuego en las superficies arboladas y en una franja a su alrededor. La incentivación del pastoreo podría ser una solución ante el conflicto, pero esta solución puede plantear problemas. La tendencia es la de promover la técnica de quema controlada.

Quemas sistemáticas de restos agrícolas

El conflicto se origina al emplearse el fuego para la eliminación de restos de cosechas (quema de rastrojos) y preparación del terreno para nuevas siembras. Es práctica común la utilización del fuego como un método sencillo y rápido para eliminar desechos agrícolas en vez de reintegrarlos al suelo, como compostas. En dichas quemas suelen quemarse además otros desechos como plásticos, ropa usada y materiales similares. Dependiendo de su dimensión, el fuego puede salirse de control y extenderse.

Declaración de zonas de reserva ecológica ó protegidas

El conflicto se plantea acerca de las limitaciones que dicha declaración acarrea a las poblaciones locales. Cuando se hace una declaración de este tipo, se producen determinadas restricciones encaminadas a la conservación o restauración de los recursos naturales. Ello tiene influencia inmediata en la vida de los habitantes de las zonas locales y próximas, y puede chocar con sus usos y costumbres, apareciendo enfrentamientos, de los que el incendio será un síntoma. La tendencia en la política proteccionista es a reconocer estos conflictos potenciales y a arbitrar medidas compensatorias, que deben extenderse a áreas de influencia de los espacios bajo algún tipo de protección. Aun así, el conflicto no está cerrado por la lentitud administrativa de

los sistemas compensatorios, necesaria para evitar fraudes, y por los recelos que despierta (Colegio de Ingenieros de Montes, 2002).

4.1.7. Conflictos en la zona de transición urbana

El conflicto se plantea al extenderse aceleradamente las zonas de uso urbano, ocupando primero las tierras agrícolas limítrofes y después las forestales más alejadas. Debido a la tecnología actual se hace potencialmente urbanizable una proporción cada vez mayor del territorio.

Expansión de los usos recreativos en el área forestal

El conflicto se plantea como consecuencia de la invasión de los espacios forestales para el desarrollo de actividades recreativas como el excursionismo y la caza. La presencia creciente de personas en el monte se traduce en mayor probabilidad de utilización negligente del fuego. La legislación de todos los países es análoga. Las fogatas están reglamentadas. También está prohibido arrojar cigarros encendidos.

4.1.8. Conflictos no relacionados con el uso de la tierra

Venganzas

Pueden ser venganzas contra individuos o contra la sociedad. En las primeras, el fuego es uno de los medios para hacer daño a una persona por una motivación privada. En las segundas, la causa de la venganza puede ser muy variada (como por ejemplo, el cazador que no obtiene un permiso para cazar y expresa su protesta quemando el bosque). En cualquier caso, el incendio intencional es un delito en todas las legislaciones.

Delincuencia

En este caso, el incendio se utiliza bien para ocultar otro delito, bien para ayudar a realizarlo. Se pueden señalar muchos ejemplos: cazadores furtivos, incendios para provocar futuras inversiones en reforestación, para distraer a la policía en operaciones de contrabando, e incluso para deforestar y poder plantar después diferentes especies de droga. Estos casos son objeto igualmente de la sanción de las legislaciones penales (Colegio de Ingenieros de Montes, 2002).

4.2. Sistemas de información geográficos (SIG)

Los SIG son una tecnología utilizada para analizar y reportar los recursos del planeta, siendo ellas complementarias unas con otras. Los procedimientos de percepción remota (PR) proporcionan mucha de la información que requieren los SIG. Los análisis de PR pueden ser mejorados a través de información externa proveniente de los SIG. Esto permite concluir que la integración de métodos y tecnologías de SIG ahunados a los de PR mejoran significativamente el análisis de datos espaciales.

Un SIG es un sistema informático que graba, almacena y analiza la información sobre los elementos que componen la superficie de la Tierra (Johnston, 1988). Un SIG puede generar imágenes de un área en dos o tres dimensiones, representando elementos naturales como colinas o ríos, junto a elementos artificiales como carreteras, tendidos eléctricos, núcleos urbanos o estaciones de metro mediante puntos, líneas, figuras geométricas u otras. Los expertos utilizan las imágenes del SIG como modelos, realizan mediciones precisas, recogen datos y corroboran sus teorías con la ayuda de una computadora.

Es obvio que en este acercamiento a nivel global, se requieren del uso de nuevas técnicas para obtener, procesar y presentar información digital de una manera rápida y

a un costo razonable. Es éste el objetivo del uso de los SIG. Un SIG toma ventaja de las habilidades para guardar y procesar grandes volúmenes de datos. Además, es posible la actualización o corrección de la información espacial, tal como derivar modelos cartográficos de diferentes formas, con las capas de información incluidas en las bases de datos.

En el caso de los incendios forestales, la diversidad de factores que afectan el inicio y desarrollo del fuego en el bosque dicta el uso de un análisis integral. Considerando el dinamismo intrínseco del fenómeno, el conjunto de imágenes obtenidas a través de técnicas de percepción remota es invaluable para este tipo de estudios. Se provee una rápida evaluación del estado de la vegetación, tal como el estudio de los efectos del fuego sobre el ambiente. En este papel radica una gran importancia la aplicación de SIG para el mapeo de riesgos de incendios (Chuvieco y Rusell, 1989).

4.3. Estudios relacionados en detección y predicción de incendios forestales

Diversos estudios se han realizado en lo concerniente al uso de redes neuronales y algoritmos genéticos para la modelación y predicción de incendios forestales.

En principio, Giroux y Dean (2002), realizaron un estudio con el objeto de comprobar que el uso de redes neuronales es una alternativa al uso del modelamiento estadístico, en particular en el análisis de procedimientos basados en *co-Kriging*. Este es un procedimiento geoestadístico basado en semi-variogramas.

En Valencia (2005) se reportaron avances en el desarrollo de algoritmos y redes neuronales con el propósito de evaluar la incidencia de incendios en el Parque Nacional

de Monfragüe, en la región española de Extremadura. Usando un modelo matemático del fuego (*Fire Mathematical Model*) aplicado a imágenes de la plataforma CHRIS/PROBA, (*Compact High Resolution Imaging Spectrometer / Project For On Board Autonomy*) a fin de evaluar la probabilidad de ocurrencia de incendios forestales.

Federico (2005) realizó un método alternativo para el entrenamiento de redes *backpropagation*, donde una vez determinada la topología de la red neuronal, empleó un algoritmo genético a fin de realizar el ajuste de los pesos en las conexiones, de tal manera que se evalúan diferentes variantes de los operadores genéticos para el entrenamiento de las redes. Una vez obtenidos los resultados, éstos fueron comparados con los obtenidos en una red con propagación hacia atrás (denominada comúnmente como *backpropagation*). El trabajo determinó que el método sí tiene la capacidad de entrenar redes *backpropagation* de las mismas características que aquellas redes en las que son empleados los métodos de entrenamiento tradicionales (prueba y error). El estudio demostró también que el dicho método propuesto puede incluso mejorar el rendimiento del entrenamiento en aquellos aspectos en que los métodos tradicionales suelen tener dificultades (Federico, 2005).

El mismo estudio estableció que los tiempos de entrenamiento promedio permitieron una marcada superioridad de los algoritmos genéticos por sobre las redes con un entrenamiento bajo el algoritmo *backpropagation*, siendo menor en un 40% aproximadamente. Asimismo, hay una mayor variabilidad del error mínimo promedio conforme transcurre el tiempo de entrenamiento.

Dos técnicas fueron empleadas por Blackard y Dean (1999) para comparar la precisión en la predicción de tipos de coberturas forestales en el Bosque Nacional Roosevelt en

el norte del Estado de Colorado, E.U., a partir de emplear variables cartográficas (elevación, exposición, pendiente, distancia horizontal a un cuerpo de agua, distancia horizontal al camino más cercano, tipo de suelo, área protegida, entre otras variables) y usando una red neuronal y un modelo estadístico basado en un análisis discriminante de tipo Gausiano. El estudio demostró que se obtiene mayor precisión en la predicción de estas coberturas, y por otra parte, que existe una relación entre los tipos de cobertura forestal con las actividades humanas.

En el mismo estudio se llevó a cabo una línea de investigación relacionada con la extracción y clasificación de imágenes hiperespectrales. A partir de ellas se desarrolló un algoritmo para obtener automáticamente el espectro de las especies presentes en la escena, y capaz de alcanzar mediante técnicas de mezclas espectrales hasta un nivel de sub-píxel, obteniéndose información incluso del nivel de humedad en el sitio. Con las imágenes de CHRIS es posible alcanzar una alta precisión en la predicción y evaluación de los incendios forestales.

Davis y Dean (2002) usaron análisis topográfico, investigaciones estadísticas y métodos de inteligencia artificial en varias zonas del Parque Nacional de Yosemite, a fin de identificar puntos de ignición de incendios forestales. Determinaron que dichos puntos de ignición se relacionan con características cartográficas (pendiente, exposición, elevación y modelos de combustibles), Por otra parte, los resultados demostraron que no había diferencia estadística con el tamaño del incendio forestal. El estudio menciona que el crecimiento de la población humana hacia las zonas boscosas, y que en dichas zonas la cantidad de combustible, prácticas de excursionismo, y el incremento de otras actividades humanas pueden contribuir al riesgo de incendios.

Un estudio realizado por Pew y Larsen (2001) consistió en llevar a cabo análisis con sistemas de información geográficos de patrones temporales de ocurrencia de incendios causados por el hombre en los bosques lluviosos de la isla de Vancouver en Canadá en un amplio período de tiempo (entre los años de 1950 a 1992), mediante una regresión logística, la cual fue utilizada para construir un modelo predictivo usando variables climáticas y relacionadas con distancias a poblaciones, sitios de acampada, veredas y caminos principales.

McCormick (2002) realizó modelaje de incendios forestales mediante el uso de meso-escalas y redes neuronales artificiales, utilizando variables climáticas, topográficas y modelos de combustibles. Asumió que el comportamiento del fuego es un sistema que puede ser descrito mediante ecuaciones simples. Asimismo establece que las redes neuronales son capaces de determinar las ecuaciones que describen tal comportamiento por medio de interacciones.

En relación a estudios forestales en nuestro país, éstos se han enfocado al impacto que se provoca en el equilibrio ecológico. A través de los sensores remotos se obtiene información de extensas áreas forestales, en algunos casos inaccesibles, con relativa rapidez (Mora y Hernández-Cardenas, 1999).

Los estudios forestales con imágenes de satélite se han generalizado principalmente en inventarios forestales, incluyendo la caracterización de masas vegetales, usos de suelo, niveles de mortandad, y otros aspectos similares (Soria *et al*, 1998).

Chuvieco y Russell (1989) realizaron estudios relacionados con la aplicación de técnicas de clasificación supervisada y sistemas de información geográficos a fin de determinar zonas con alto riesgo de incendios en la región Mediterránea de España. Las variables empleadas fueron especies vegetales, clasificadas de acuerdo a clases de combustibles, la elevación, al pendiente, exposición y la proximidad a caminos y senderos, lugares para acampar y viviendas. Para ello fueron utilizadas imágenes de alta resolución del Mapeador Temático (TM) de Landsat de los años 1984 y 1985. La modelación del riesgo de incendio se llevó a cabo mediante un modelo cualitativo, aplicado mediante cinco capas de información (vegetación, pendiente, exposición, elevación y proximidad a caminos), en virtud de que algunas capas tendrán más influencia que otras.

Los resultados reportaron que más del 22% de los píxeles evaluados como de alto riesgo de incendios se incendiaron, en tanto que solo 3.74% de las que fueron consideradas con bajos valores de riesgo, fueron afectadas por el fuego (Chuvieco y Russell, 1989).

Kourtz (1990) menciona que los incendios forestales de causalidad humana no son sucesos aleatorios. Dtermina que a menudo se localizan cerca de núcleos habitados y tienden a seguir patrones estacionales, semanales e incluso diarios.

Lymberopoulos *et al.* (1996) desarrolló un sistema de información geográfica para el manejo de incendios forestales para las regiones de Grecia, Italia y Portugal a partir de datos geográficos y de vegetación obtenidos por el sensor AVHRR, en donde determinaron que la precisión de los estudios realizados varían acorde a la fuente y calidad de datos utilizada y a su resolución espacial. Determinaron en su estudio que en el manejo de incendios forestales, la utilización de imágenes LANDSAT y SPOT y la inclusión de información en tiempo real de detección de puntos de calor pueden aportar nuevas alternativas para la investigación de este fenómeno.

En Vega-García *et al.* (1996) fueron empleadas redes neuronales *backpropagation* para predecir ocurrencia de incendios forestales causados por el hombre en Whitecourt Provincial Forest, en las cercanías de Alberta, Canada. Un modelo *logit* fue empleado para seleccionar las mejores variables para la red. Los modelos fueron desarrollados utilizando los datos del período 1986 a 1990 para la etapa de entrenamiento; en tanto que para la etapa de prueba se usaron los datos de los años 1991 y 1992. Del total de las observaciones (8,009), fue escogida una muestra aleatoria de 157 observaciones de puntos donde NO se tuvo presencia de incendios, más otras 157 observaciones donde al contrario, SÍ hubo presencia de incendio.

Variables como distancia a caminos, sitios de campamento y a núcleos de población, elevación topográfica, combustibles, valor comercial, humedad relativa, velocidad del viento y mes de ocurrencia, fueron empleadas en este estudio. A partir de éstas, se crearon otras variables como distancias promedios a zonas de campamento, a caminos y centros de población, área, elevación, entre otras. Los resultados de este estudio

demonstraron una precisión del 85% en las predicciones de sitios que no se incendiarían, y del 78% en las predicciones donde sí existió presencia de fuego.

La información relativa a características geográficas cambia con el tiempo debido a la construcción de nuevas carreteras e infraestructura, por lo que es importante utilizar un período de tiempo corto para el estudio, porque la precisión de la información se va reduciendo con el tiempo.

Posteriormente, en el estudio realizado por Vega-García *et al.* (1999) se utilizaron dos métodos a fin de predecir los días de ocurrencia de incendios forestales también en Whitecourt Provincial Forest. Uno de estos métodos fue basado completamente en la aplicación de un modelo de regresión logarítmica (modelo logit) y que fue comparado contra los resultados de un segundo método, el cual fue constituido por una red neuronal artificial de tipo *backpropagation* con una regla de aprendizaje *delta generalizada*. Se utilizaron dos tipos de factores: *geográficos* (muy variables en el espacio pero poco variables en el tiempo) y *temporales* (muy cambiantes en el tiempo durante el período del estudio pero con poca variabilidad espacial). Los resultados fueron similares a los obtenidos por Vega-García *et al.* (1996), ya que la red clasificó correctamente el 84% y casi el 78% de las observaciones de NO presencia y SI presencia de días sin incendios respectivamente.

Buendía (2000) evaluó la habilidad de las redes neuronales artificiales en la predicción de tipos de coberturas vegetales, basándose en información geográfica (altitud, exposición, pendiente, distancia a los escurrimientos, geología y edafología como variables independientes) empleando imágenes de satélite transformadas en análisis de

componentes principales como la variable dependiente (vegetación). En este estudio realizó el entrenamiento de una red *backpropagation* con dos capas ocultas con funciones de activación TANGENTE HIPERBÓLICA y GAUSIANA en cada capa antes mencionada, y empleando para tal fin el programa NEUROSHHELL 2©. En su estudio obtuvo en esta arquitectura de red una R^2 de 0.8617 en la fase de entrenamiento y una R^2 de 0.8514 en la fase de prueba, alcanzando un 83% de sitios predichos correctamente.

Finalmente, Miller *et al.* (2005) probó y desarrollo técnicas de RNA para analizar imágenes multicanal de las plataformas MODIS, GOES y AVHRR, entrenando las redes para las firmas espectrales por incendios forestales para automatizar el proceso de detección. En este estudio se utilizó en conjunto de incendios forestales causados por el hombre en la zona norte de Florida, Estados Unidos. Se utilizó un arreglo de siete por siete píxeles para cada plataforma satelital, alcanzando una precisión de entre el 80 al 92% dependiendo de la modificación de algunos parámetros de la red neuronal y del satélite empleado.

4.4. Caracterización de las redes neuronales artificiales (RNA)

4.4.1. Aspectos generales

Las RNA forman parte de los sistemas basados en la inteligencia artificial (IA). Este término fue acuñado en 1956 por John McCarthy, del Instituto de Tecnología de Massachussets. En ese año se celebró la conferencia de Dartmouth, en Hanover (Estados Unidos), y en ella, McCarthy, Marvin Minsky, Nathaniel Rochester y Claude E. Shannon establecieron las bases de la inteligencia artificial como un campo

independiente dentro de la informática. Previamente, en 1950, Alan M. Turing había publicado un artículo en la revista *Mind*, titulado *Computing Machinery and Intelligence* (“Ordenador e inteligencia”), en el que reflexionaba sobre el concepto de inteligencia artificial y establecía lo que luego se conocería como el “Test de Turing”, una prueba que permite determinar si una computadora se comporta conforme a lo que se entiende como artificialmente inteligente o no (Microsoft, 2007). Con el avance de la ciencia moderna la búsqueda de la IA ha tomado dos caminos fundamentales: la investigación psicológica y fisiológica de la naturaleza del pensamiento humano, y el desarrollo tecnológico de sistemas informáticos cada vez más complejos. La IA es una nueva ciencia con la cual se representa, se adquiere automáticamente y se utiliza el conocimiento (Kourtz, 1990).

En este sentido, el término IA se ha aplicado a sistemas y programas informáticos capaces de realizar tareas complejas, simulando el funcionamiento del pensamiento humano, aunque todavía muy lejos de éste. En esta esfera los campos de investigación más importantes son el procesamiento de la información, el reconocimiento de modelos, los juegos y las áreas aplicadas, como el diagnóstico médico. Un ejemplo de los logros alcanzados fue la partida de ajedrez que el superordenador de IBM denominado *Deep Blue* ganó, en mayo de 1997, al campeón del mundo Gari Kaspárov. Actualmente el desarrollo tecnológico de los sistemas informáticos de la IA se centra en los sistemas expertos y las redes neuronales. Los sistemas expertos intentan reproducir el razonamiento humano de forma simbólica, en tanto que las redes neuronales lo hacen desde una perspectiva más biológica (recreando la estructura de un cerebro humano mediante algoritmos genéticos) (Microsoft, 2007).

4.4.2. Definición de redes neuronales artificiales

Una red neuronal (biológicamente) es un procesador distribuido masivamente y en paralelo que es propenso por naturaleza a almacenar conocimiento experimental y hacerlo disponible para su uso. Este mecanismo se funda en dos aspectos:

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

Una red neuronal artificial es un modelo computacional con un conjunto de propiedades específicas, como son la habilidad de adaptarse o aprender, generalizar u organizar la información, todo ello basado en un procesamiento eminentemente paralelo. Esto es, que a diferencia del algoritmo computacional tradicional, las operaciones de la red se realizan simultáneamente en cada neurona que constituyen la capa oculta.

Las RNA constituyen una estructura de procesamiento de información paralela y distribuida, que intenta emular las funciones computacionales elementales de la red nerviosa del cerebro humano, con base a la interconexión de multitud de elementos de procesamiento, cada uno de los cuales presenta un comportamiento completamente local.

4.4.3. Características de las redes neuronales artificiales

Las RNA se desarrollaron como un intento de emular la sofisticada capacidad de procesamiento del cerebro humano. A excepción de las tareas basadas en operaciones

aritméticas simples, el cerebro es sin duda, superior a cualquier equipo de cómputo concebido hasta hoy por el hombre.

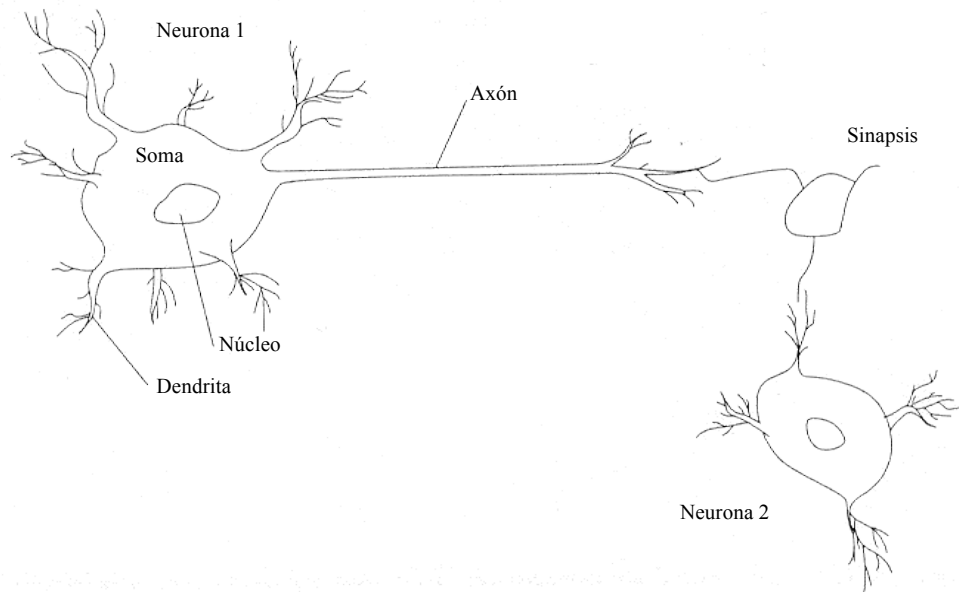


Figura 4.1. Dos neuronas biológicas interconectadas

Estas redes combinan modelos matemáticos de las células nerviosas y modelos de arquitecturas que describen las interconexiones que existen en esas células, en similitud al funcionamiento de dos neuronas biológicas (Figura 4.1). El *software* (programas o paquetería) convencional utilizado en los sistemas computacionales actuales procesa datos en base a la ejecución de un algoritmo establecido, que no es sino una secuencia ordenada de pasos y reglas preestablecidas. Las redes neuronales en cambio, solucionan problemas en base a un gran número de elementos, asociando los datos de entrada con los datos de salida. Las RNA además, generan ellas mismas sus propias "reglas", para asociar la respuesta a su entrada; es decir, aprende por ejemplos y de sus propios errores. Así, el conocimiento de una RNA se encuentra en la función de activación utilizada y en los valores de sus ponderaciones (ISA-UMH, 2007).

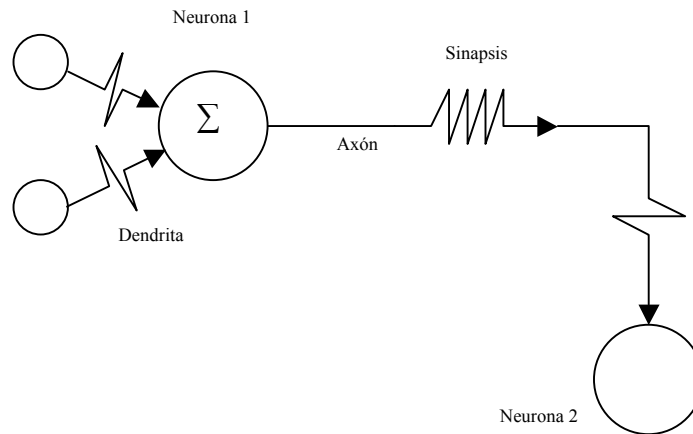


Figura 4.2. Equivalente artificial de dos neuronas interconectadas

Las RNA están formadas por perceptrones (unidad básica de la red artificial donde se procesa la información y se aplican los pesos y que es equivalente a la neurona humana) que se interconectan en forma muy similar a la forma como están dispuestas en la corteza cerebral humana (en la Figura 4.2 puede observarse ésta interconexión artificial) y lo que es más importante, son capaces de asociar y generalizar sin reglas. Han sido utilizadas con gran éxito para reconocer retornos de sonar bajo el agua, escritura a mano, voz, topografía de terrenos, controlar brazos de robots, evaluar datos personales, modelar fenómenos cognoscitivos y predecir tendencias financieras. La clase de problemas que mejor se resuelven con las redes neuronales son los mismos que el ser humano resuelve con mayor facilidad, como son la asociación, la evaluación y el reconocimiento de patrones.

4.4.4. Redes neuronales basadas en algoritmos genéticos.

Otros tipos de redes diferentes al algoritmo tradicional *backpropagation* son aquellas basadas en algoritmos genéticos. Un algoritmo genético funciona seleccionando de un conjunto de individuos a uno, el cual es una solución potencial para el problema. En

este caso, una solución potencial es un conjunto de factores de atenuamiento o efecto *smooth*, y el algoritmo genético busca un individuo que minimiza el Error Mínimo Promedio (EMP) del conjunto de datos de prueba. Después de probarlos todos, una nueva generación de observaciones es generada para ser probada nuevamente.

Los factores de atenuamiento individuales son presentados para cada variable de entrada. El factor de atenuamiento es un ajuste empleado para modificar los pesos que permiten proveer un nuevo valor para cada variable de entrada. Al final de la etapa del entrenamiento, los factores individuales de atenuamiento son usados en un análisis de sensibilidad: los valores de salida más altos son aquellos que tienen mayor importancia para el modelo, en tanto que los factores con valores más bajos son candidatos a ser eliminados del modelo para pruebas posteriores. Dentro del programa NeuroShell 2© se tienen dos arquitecturas basadas en algoritmos genéticos: *Probabilistic Neural Network* (PNN) y *General Regression Neural Network* (GRNN).

Los algoritmos genéticos constituyen una de las formas más importantes de los algoritmos evolutivos, y se consideran herramientas de optimización muy poderosas. Al igual que las redes neuronales, los algoritmos genéticos están inspirados en la biología: se basan en la teoría de la evolución genética y en el concepto de la supervivencia del más apto. Entre las características más importantes se destacan su naturaleza estocástica, su capacidad de considerar simultáneamente una población de soluciones y su adaptabilidad ante un rango amplio de problemas. Actualmente constituyen un paradigma en inteligencia artificial que ha recibido mucha atención. Estos algoritmos que emplean modelos idealizados del código genético, la recombinación de información genética, las mutaciones y la selección, han producido formas muy genéricas y robustas

de encontrar soluciones para problemas de búsqueda computacionalmente difíciles. Uno de estos problemas es la adaptación (o entrenamiento) de las arquitecturas y parámetros de las redes neuronales (Fiszelew y García, 2000).

Una de las principales funciones del algoritmo genético es su capacidad de permitir que la intervención del programador humano sea mínima, permitiendo que la computadora y el programa resuelvan el problema, aprendiendo de sus propios errores y optimizando los resultados conforme va entrenándose.

4.4.4.1. PNN (Probabilistic Neural Network)

Las redes neuronales probabilísticas o *Probabilistic Neural Network* (o PNN por sus siglas en inglés) son conocidas por su capacidad para entrenar rápidamente datos dispersos. Las PNN separa los datos dentro de un número específico de datos de salida categorizados.

Está conformada por una red de tres capas donde las observaciones de entrenamiento son presentadas en una capa de entrada y la capa de salida está conformada por una neurona por cada categoría en que se desea clasificar dichas observaciones. La capa oculta estará conformada por un número de neuronas idéntico al número de observaciones presentes en el estudio. Esta red clasifica de acuerdo a la definición de probabilidad estimada para cada categoría. Los datos de salida más altos representan la categoría más probable. Dado que es esencialmente una red clasificatoria, no se empleó para este estudio.

4.4.4.2. GRNN (General Regression Neural Net)

El algoritmo genético en el programa NeuroShell 2© es la red denominada *General Regression Neural Network* o GRNN (por sus siglas en inglés), el cual utiliza un procedimiento basado en factores de atenuamiento con el que el algoritmo intenta adaptar los resultados del proceso a una forma gaussiana en la capa oculta. Básicamente, consiste en que el *software* emplea una red con numerosas neuronas a fin de determinar en cada variable cual de ellas requiere ser ajustada.

La arquitectura GRNN es una red integrada por tres capas: una capa de entrada con una neurona por cada dato de entrada, una capa oculta integrada por una neurona por cada observación, y una capa de salida con una neurona por cada observación. Éste método no requiere que se ajusten parámetros como tasa de aprendizaje, momentum o pesos iniciales.

Esencialmente, las GRNN son redes de aprendizaje genético adaptativo, es decir, basan su aprendizaje a partir de un algoritmo genético el cual se basa en la búsqueda selectiva en una población de patrones individuales, cada uno de los cuales constituye una solución potencial al problema. En este caso, dicha solución es un conjunto de factores denominados factores de atenuamiento (efecto *smoothing*) mismos que se encuentran dentro de una capa denominada Conjunto de Patrones Genéticos (*Genetic Breeding Pool*), siendo ahí donde se aplican en cada época de la fase de entrenamiento a los datos de entrada y los compara con los datos de salida a fin de minimizar el error mínimo promedio en el conjunto de prueba, determinando cuales valores deben ser ajustados mediante la función de suavizamiento. Este módulo se inicia con un factor

tomado de forma aleatoria, y se va perfeccionando conforme se realizan nuevos ciclos (Figura 4.3).

Es posible seleccionar en el programa Neuroshell 2© el tamaño de la población de este conjunto de posibles soluciones, cuyo tiempo de entrenamiento aumentará proporcionalmente conforme lo hace dicho conjunto de patrones, pero al mismo tiempo es más grande la posibilidad de encontrar una solución potencial al problema. Después de probar todos los patrones individuales, se genera un nuevo conjunto de patrones para ser probados en la siguiente época. Es por ello que éste modelo es capaz de entrenarse aún con pocos datos de entrada en los ciclos iniciales.

Con las variables obtenidas por el método anteriormente referido, se probaron diferentes tipos de arquitecturas, con el fin de diseñar una red neuronal de algoritmo genético que tuviera el mejor desempeño para tales variables, en base a criterios de decisión ya establecidos.

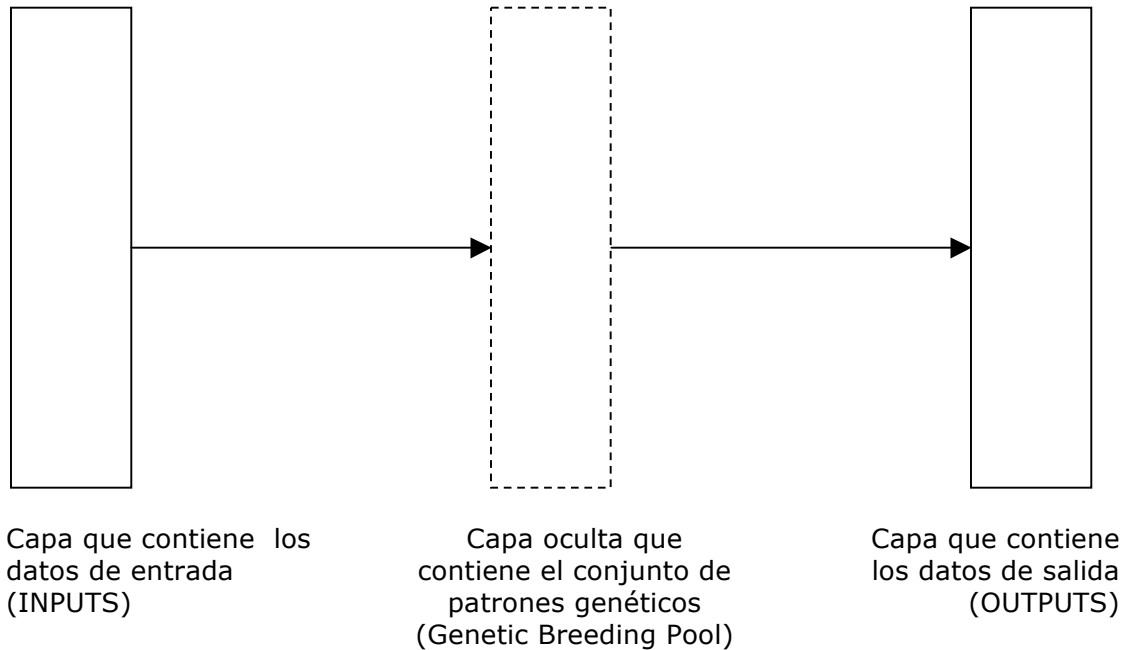


Figura 4.3. Esquema del funcionamiento de la arquitectura GRNN

4.4.5. Ventajas de las redes neuronales artificiales

En Fiszlelew y García (2000) se menciona que la diferencia sustancial entre una RNA y una aplicación de *software* diseñado por ejemplo para determinar patrones de comportamiento consiste en que una red no se programa, sino que se entrena. Las redes neuronales trabajan de forma completamente distinta a los algoritmos con los cuales una computadora es programada. Las RNA ofrecen un paradigma atractivo para el diseño y análisis de sistemas adaptativos inteligentes. Entre estas se encuentran:

- Flexibilidad para la adaptación y aprendizaje.
- Robustez ante la presencia de ruido (datos erróneos o incompletos).
- Habilidad para generalizar.
- Capacidad de recuperación ante fallas.

- Su computación es masiva y paralela.
- Guardan semejanza con las redes neuronales biológicas.

4.4.6. Estructura de las redes neuronales artificiales

El funcionamiento de una red neuronal se relaciona estrechamente con sus componentes y su disposición tal como se muestra en la Figura 4.4. El componente básico de las redes neuronales es la neurona artificial ó *neuron*, la cual se encuentra interconectada con otras neuronas mediante enlaces denominados *links*. Neuronas y enlaces conforman la red. Estas neuronas interconectadas poseen pesos denominados *weights*, que son aplicados cuando los valores de entrada pasan de una neurona a la siguiente.

Los pesos adaptables emplean leyes de aprendizaje para ajustar el valor de la fuerza de una interconexión con otras neuronas. Si las neuronas utilizan pesos fijos, entonces su tarea deberá estar previamente definida. Los pesos serán determinados a partir de una descripción completa del problema. Por otra parte, los pesos adaptables son esenciales si no se conoce previamente cual deberá de ser su valor correcto (Vega-García *et al.*, 1996).

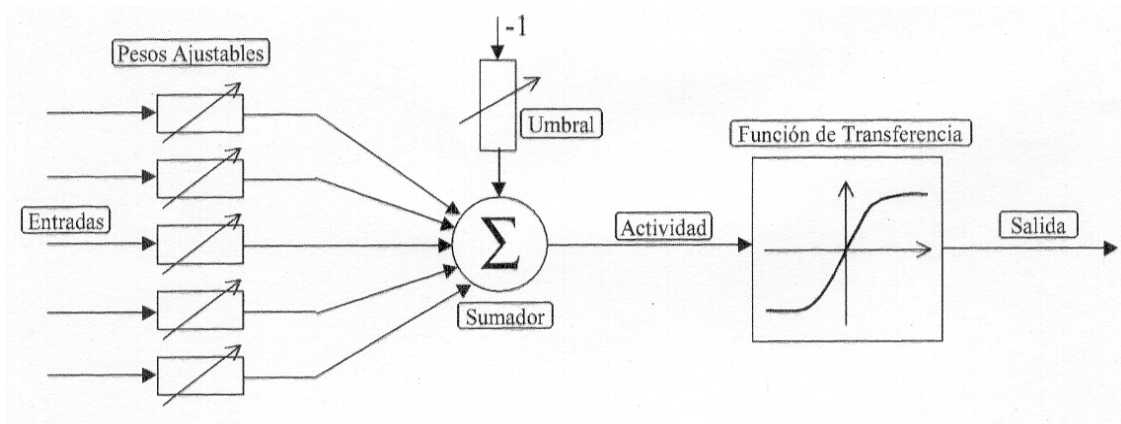


Figura 4.4. Funcionamiento de una neurona artificial
(Tomado de ISA-UMH, 2007)

Las neuronas se agrupan en tres tipos diferentes de capas (un grupo de neuronas se denomina como *slack*), de acuerdo a su conexión con el exterior de la red. Con base a lo anterior, son neuronas de la capa de entrada (también llamada *input layer*) aquellas que reciben datos del exterior de la red. Si las neuronas contienen las predicciones o las clasificaciones de la red, entonces corresponden a neuronas de la capa de salida, denominada también como *output layer*.

Las neuronas que se encuentran entre las capas de entrada y de salida corresponden a la capa oculta ó *hidden layer*, pudiendo haber más de una capa oculta en una red neuronal formando varios *slacks*. En la Figura 4.5, se muestran los componentes más comunes de las RNA.

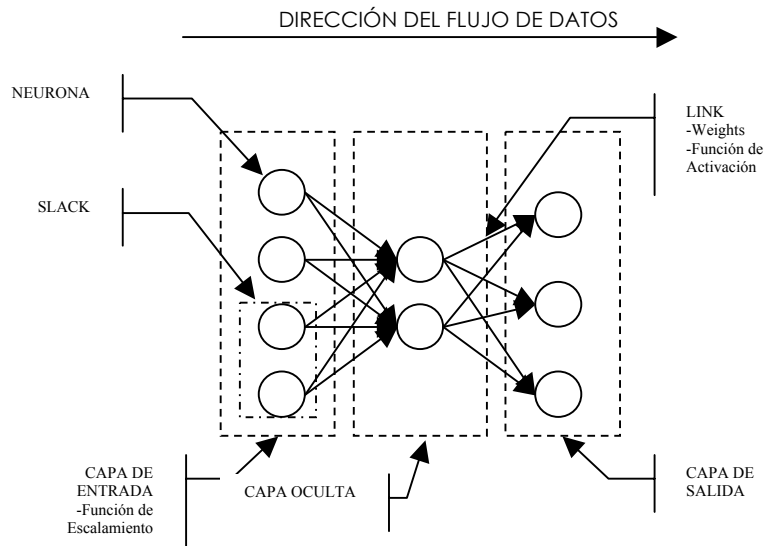


Figura 4.5. Componentes de una red neuronal backpropagation

Cada dato de entrada corresponde a una característica en particular, la cual es introducida como un valor numérico. Esta característica representa atributos que bien pueden ser categorías o valores continuos. Los datos de salida de la red representan la solución del problema, como un resultado binomial (SI ó NO), asignándole valores numéricos de 1 ó 0 respectivamente. Luego de un post-procesamiento, en el cual se realizan ponderaciones en base a criterios, por ejemplo en los casos en que el resultado sea un 0.85 o un 0.4, se puede obtener una salida interpretada de manera razonable.

Mediante diversos parámetros, como los *pesos*, el *momentum* y la *tasa de aprendizaje*, se expresan las fuerzas relativas (o valor matemático) que son transferidos por las muchas conexiones de una capa a otra. En otras palabras, estos parámetros expresan la importancia relativa de cada valor de entrada a los elementos de la red y en última instancia las salidas de la red. Es así como a través de estos ajustes que la red realiza su aprendizaje.

Es así que, durante el procesamiento de la información, diversos elementos de la red se encuentran ejecutando operaciones de forma simultánea. Este proceso es entonces paralelo, de forma idéntica a la forma como el cerebro trabaja, diferente al procesamiento en serie de la computación convencional.

4.4.7. Diseño de la red neuronal artificial

El diseño de la red se considera el punto medular en las redes neuronales. Existe una gran variedad de soluciones dentro de un mismo diseño, que está en función de las variables escogidas, de la elección de la arquitectura (la cual incluye la disposición y número de neuronas), el tiempo de entrenamiento, la determinación de los valores de momentum, la tasa de aprendizaje y de los pesos iniciales.

Asimismo, muchos problemas responden de forma diferente a diferentes tipos de redes y arquitecturas, y no se ha determinado aún un método definitivo para determinar cual de ellas funciona mejor. Las redes más avanzadas pueden obtener una diferencia significativa en la habilidad predictiva en algunos casos. Sin embargo, es posible obtener una solución aceptable empleando una red simple de tres capas (Ward System Group, 1996).

El diseño de una arquitectura óptima (o *quasi* óptima) para una red neuronal puede verse como un problema de búsqueda en el espacio de arquitecturas donde cada punto representa una arquitectura distinta. Dado un criterio de decisión para definir la mejor arquitectura, como por ejemplo, el error de entrenamiento más bajo o la complejidad

más baja, el nivel de desempeño de todas estas arquitecturas formará una superficie discreta en el espacio. Obtener el diseño de arquitectura más cercano al óptimo es equivalente a encontrar el punto más alto sobre esta superficie (Figura 4.6).

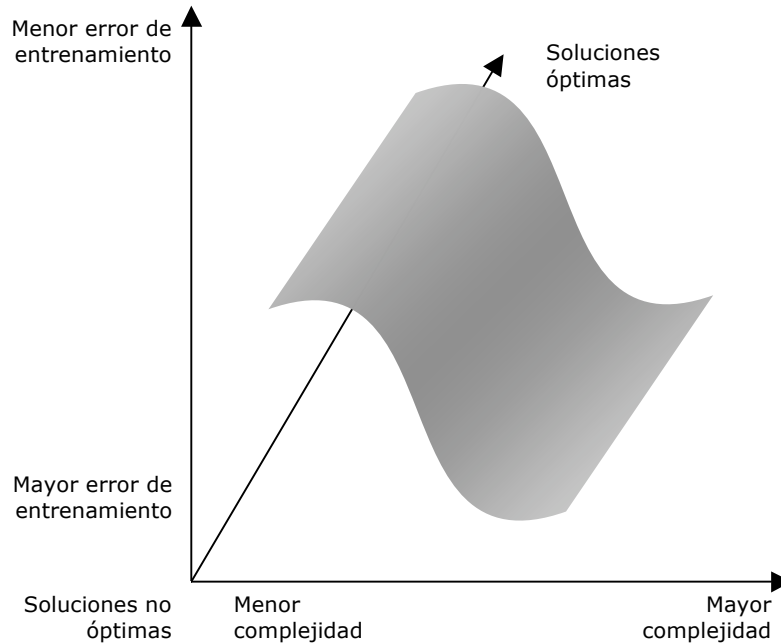


Figura 4.6. Búsqueda del mejor diseño de arquitectura en un espacio 3D

La definición del modelo en este estudio se determinó realizando ensayos con diferentes tipos de modelos, modificando en ellos los parámetros de forma sistemática, con auxilio de la bibliografía relacionada para casos similares, con el fin de obtener aquellos modelos que muestren el mejor ajuste de los resultados predichos con relación a los resultados esperados en función de las variables empleadas en el estudio (Figura 4.7).

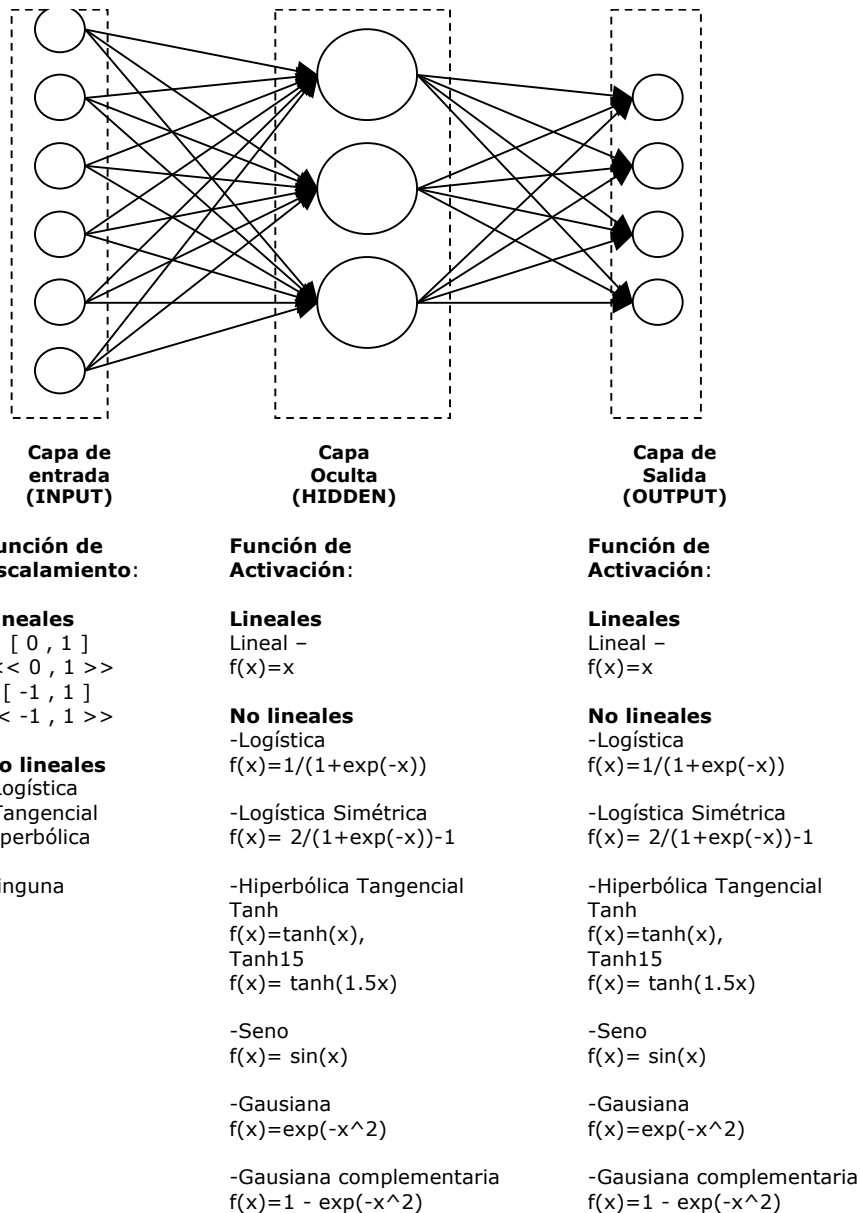


Figura 4.7. Funciones en una red backpropagation

4.4.8. Diseño de la arquitectura de la red

Existen tres grupos de arquitecturas prediseñadas para redes del tipo *backpropagation* incluidas en NeuroShell 2© que proporcionan diferentes niveles de complejidad en su estructura. Las mismas pueden ser modificadas en todos sus parámetros, sin embargo,

no pueden ser diseñadas nuevas arquitecturas. Estas redes de tipo *backpropagation* se componen por tres grupos esenciales:

- a. **Redes Estándar.** Redes lineales en las que cada capa se encuentra conectada solo con la capa previa.
 1. Lineal de tres capas (con una capa oculta)
 2. Lineal de cuatro capas (con dos capas ocultas)
 3. Lineal de cinco capas (con tres capas ocultas)

- b. **Redes recurrentes (Jordan-Elman nets).** Son redes retroalimentadas en que los datos vuelven a circular en capas anteriores.
 1. Recurrente con retroalimentación en la capa de entrada.
 2. Recurrente con retroalimentación en la capa oculta.
 3. Recurrente con retroalimentación en la capa de salida.

- c. **Redes múltiples (Ward nets).** Múltiples capas ocultas con diferentes funciones de activación, que pueden saltar o no capas ocultas.
 1. Red *backpropagation* con dos grupos de neuronas (slacks) teniendo diferentes funciones de activación seleccionadas.
 2. Red *backpropagation* con tres grupos de neuronas (slacks) teniendo diferentes funciones de activación seleccionadas.
 3. Red *backpropagation* con dos grupos de neuronas (slacks) teniendo diferentes funciones de activación seleccionadas y salto (jump) de conexión seleccionado.

4.4.9. Funciones de activación (FA)

Las funciones de activación (FA) sólo pueden ser aplicadas y modificadas en las redes *backpropagation*. Las capas ocultas producen salidas basadas en la suma de los pesos modificados que pasan por ellas. La forma en que tales capas producen sus salidas se realiza por medio de la aplicación de una FA a la suma de los pesos valuados. La FA, también llamada función *squashing* (que se traduce como aplastamiento), corrige la suma del valor de salida, para luego ser expedida hacia la siguiente capa (Ward Systems Group, 1996).

Aunque la función logística (*logistic*) es la más popular, hay otras funciones las cuales pueden ser usadas. Algunos parámetros en el programa de Neuroshell 2© reducen la necesidad de utilizar otras funciones de activación. Sin embargo, algunos problemas responderán mejor a unas funciones que a otras

Dado que el uso de las funciones de activación (FA) es un parámetro que maneja una gran variabilidad de opciones, es que se determinó con base a estudios similares, y a las recomendaciones proporcionadas en el programa, emplear tres FA combinadas en las capas ocultas y la capa de salida. Estas funciones fueron:

a. FUNCIÓN LOGÍSTICA (Logistic)

Está expresada como:

$$f(x) = 1 / (1 + \exp(-x))$$

La función logística, también llamada SIGMOIDEA LOGÍSTICA, se recomienda cuando las salidas son categóricas, en rangos de 0 a 1 (Figura 4.8).

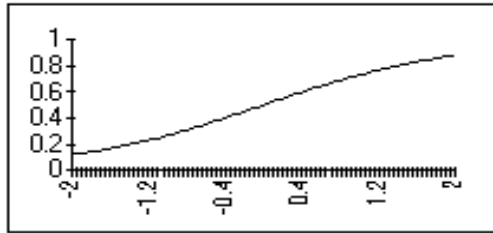


Figura 4.8. Función logística.

b. FUNCIÓN GAUSIANA (Gaussian)

Esta expresada cómo:

$$f(x) = \exp(-x^2)$$

Esta función es única, porque a diferencia de las otras, no es una función creciente. (Es la clásica curva en forma de campana). Esta función produce salidas en [0,1].

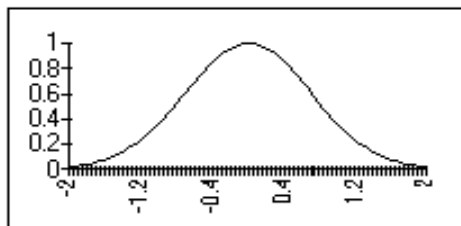


Figura 4.9. Función gaussiana.

El programa recomienda que si se decide utilizar una red de tres capas, se utilice la FA gaussiana en la capa oculta. Éste función es usualmente más poderosa en redes de tres capas (Figura 4.9).

c. FUNCIÓN GAUSIANA COMPLEMENTARIA (Gaussian compl.)

Esta expresada cómo:

$$f(x) = 1 - \exp(-x^2)$$

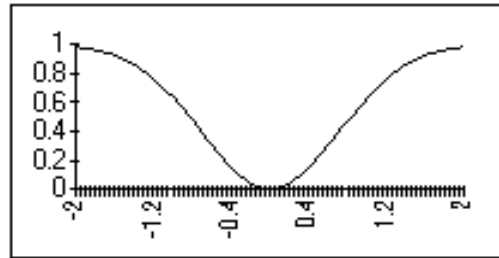


Figura 4.10. Función gaussiana complementaria.

Muy utilizada en las redes múltiples, esta función de activación resalta las características de los datos que se encuentran a los extremos (Figura 4.10).

d. FUNCIÓN TANGENTE HIPERBÓLICA (Tanh)

Está expresada como:

$$f(x) = \tanh(x)$$

La función tangente hiperbólica es de mejor uso para salidas de valores continuos, sin embargo, se recomienda usar cuando se utiliza la función lineal en la capa de salida. Si es utilizada en la primera capa oculta (hidden layer), se deberá ajustar la escala de las entradas en el rango de $[-1,1]$ en lugar del rango $[0,1]$. Se han experimentado con buenos resultados cuando se utilizó la función tangente hiperbólica en la capa escondida de una red de tres capas, y usar la función logística o la función lineal sobre la capa de salida (Figura 4.11).

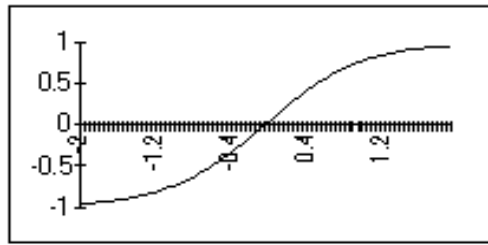


Figura 4.11. Función tangente hiperbólica.

4.4.10. Modo de actualización de pesos (MAP)

Un factor cuya influencia es significativa en el desempeño de las redes neuronales es la manera en que son aplicados los pesos programados en los *links* (enlaces) de la red. Este parámetro guarda relación con otros tres factores de la red que son (Ward Systems Group, 1996):

a. Tasa de aprendizaje (LR) ó Learning Rate

Cada vez que un patrón es presentado a la red, los pesos que son dirigidos al nodo de salida son modificados ligeramente durante el aprendizaje en la dirección requerida para producir un error más pequeño para la próxima vez que los datos sean nuevamente presentados a la red. La cantidad del peso modificado está en relación con el error que corrige la tasa de aprendizaje. Es decir, si la tasa de aprendizaje es de 0.5, el cambio en el error será solo de la mitad del error detectado. Si la tasa de aprendizaje es alta, la modificación en el peso será alta, y el aprendizaje será más rápido. Sin embargo, puede ocurrir oscilación y la no convergencia del modelo si la tasa de aprendizaje es muy alta.

b. Momento (Mo) ó Momentum

Tasas de aprendizaje con valores altos (mayor a 0.9) provocan oscilaciones por lo que el modelo converge a una solución que no es la óptima. Una forma de lograr el aprendizaje sin la oscilación antes descrita es la utilización del parámetro denominado Momento o *momentum*, que consiste en aplicar una función en dicha capa para crear un efecto de atenuamiento. El factor momentum permite determinar la proporción del último cambio de peso que es agregado al cambio de un nuevo peso.

c. Peso inicial (InW) ó Initial Weight

Dado que las neuronas transmiten valores desde una capa de la red a la siguiente capa en las redes *backpropagation*, los valores son modificados por un valor-peso en el enlace que representa la conexión entre las neuronas.

Cuando la red es diseñada, los pesos comienzan como números aleatorios que disminuyen dentro de un rango especificado. Como cada patrón transita a través de la red, el peso debe ser elevado para reforzar positivamente una conexión. Para reforzar negativamente o inhibir una conexión, el peso es disminuido.

La combinación de estos tres factores permite el aprendizaje de la red, pero no existe un método definido para lograr que dicha combinación sea la óptima. Dado lo anterior, el ajuste de los pesos de la red en el programa NeuroShell 2© puede lograrse utilizando alguno de los tres modos que se incluyen en el diseño de la red, que son:

Modo Vanilla

Solo se aplica la tasa de aprendizaje a los links, pero el momentum no es aplicado. Este algoritmo permite velocidad y razonable exactitud, pero por lo general provoca una oscilación en el aprendizaje, oscilación que es evitada por el Momentum aplicando un efecto suavizante, en especial en tasas de aprendizaje mayores al 0.9.

Modo Momentum

Aplica el efecto suavizante a la oscilación de la tasa de aprendizaje, lo que permite en otras palabras mantener a la red en una dirección, y es recomendado su utilización cuando son datos con alto nivel de *noise* (ruido).

Modo TurboProp

Tiene la ventaja que no es sensitivo a la tasa de aprendizaje (LR) y al momentum (Mo). Esto significa que mientras los dos modos anteriores requieren que el usuario seleccione una LR y un Mo el cual no cambia durante el proceso, TurboProp selecciona un modo de actualización de pesos que es independiente para cada uno de ellos, a diferencia de los modos anteriores, los cuales utilizan los mismos pesos en todo el proceso. El intervalo de calibración tiene el mismo tamaño de una época, por lo que los pesos son actualizados cada vez que el entrenamiento termina una época completa.

Los parámetros anteriormente descritos son modificables en redes *backpropagation*, y su efecto no es de gran importancia para el desarrollo de la red. Los parámetros referidos en el punto anterior implican variaciones importantes en el desarrollo de la red.

4.4.11. Neuronas en la capa oculta

El número de neuronas en la capa oculta usualmente se define automáticamente, pero puede ser modificado. Usando la calibración, la cual limita el sobre-aprendizaje y previene la memorización, el número de neuronas ocultas no es un factor crítico. Si no se emplea la utilidad de *calibración*, el usar pocas neuronas es mejor para generalizar las observaciones.

El número de neuronas que está dado por una red de tres capas se calcula mediante la fórmula siguiente:

$$NNO = \frac{VIn + VOut}{2} + \sqrt{ObTrn}$$

Donde:

NNO = Número de neuronas ocultas

VIn = Variables de entrada

VOut = Variables de Salida

ObTrn = No. de observaciones usadas para el entrenamiento de la red

Para un mayor número de capas, se divide NNO entre el número de capas ocultas. En las redes *backpropagation*, el número de neuronas ocultas determina como el problema puede ser aprendido. Si se usan demasiadas, la red tenderá a memorizar el problema, aún utilizando la utilidad de *calibración*. Si se usan pocas, la red no generaliza lo suficiente.

La definición del número de neuronas ocultas es un proceso de prueba y error, sin embargo, la fórmula antes mencionada suele obtener resultados satisfactorios en la mayoría de las redes.

4.4.12. Calibración

La calibración optimiza la aplicación de la red durante el entrenamiento. Encuentra la red óptima para los datos en un set de prueba. La calibración trabaja de forma diferente dependiendo del tipo de red que se utiliza. Para redes *backpropagation*, la calibración utiliza los juegos de entrenamiento y prueba para calcular el punto óptimo en que la red alcanza a generalizar nuevos datos sin sobreentrenarse.

La calibración se realiza mediante el cálculo del cuadrado medio del error (mean squared error) entre los datos que son introducidos a la red y los datos que la red calcula. Se crea un conjunto de datos completamente separado del juego de entrenamiento y del conjunto de datos de prueba, el cual evalúa la precisión de las predicciones o entrenando la red. Por lo general se utiliza un 10% del tamaño del conjunto de datos del entrenamiento.

4.4.13. Secuencia para la entrada de los datos

El programa NeuroShell 2© permite dos alternativas en el modo en que la red lee los datos para ser ingresados en la red neuronal durante la etapa de entrenamiento (*training*). El primero es en forma *secuencial*, donde el programa lee en orden desde el primer hasta el último dato, y el segundo es en forma *aleatorio*, donde se leen los datos sin seguir un orden preestablecido. Las redes recurrentes (Jordan-Elman) donde la lectura de los datos se debe realizar sólo en forma secuencial, mientras que en todas las redes *backpropagation* es posible seleccionar ambas modalidades.

Las pruebas realizadas en las redes demostraron que la forma como se ingresan los datos no tiene influencia significativa en su desempeño, pues el error mínimo promedio

del entrenamiento no tiene modificaciones importantes, en comparación con otros parámetros de la red.

4.4.14. Selección de variables de entrada

Tanto en los sistemas lineales como no lineales resulta siempre difícil determinar la precisa contribución de una variable de entrada en la funcionalidad de un sistema. Uno de los métodos más recurridos es tomando la totalidad de las variables y evaluar el efecto que tienen en conjunto, variando algunos valores de cada una. El inconveniente de estos métodos es que se pueden obtener diferentes efectos y omitir otros, además que el número de combinaciones puede llegar a ser abrumador.

Asimismo, en otros estudios se ha intentado emplear un algoritmo genético para seleccionar combinaciones de datos de entrada, y posteriormente, probar en la red cada una de ellas hasta encontrar la combinación que produzca los mejores resultados. Aquí el inconveniente es que en las redes de *backpropagation* los pesos iniciales (*Initial Weight* o InW) son seleccionados de forma aleatoria, por lo que dichos pesos no siempre son los mismos, provocando variaciones en los resultados. Este inconveniente solo se subsana al determinar en el modelo un peso específico, lo cual limita al modelo al no poder variar este parámetro. Otro problema es que tales combinaciones solo muestran la interacción entre esos mismos datos de entrada, pero no es posible averiguar el nivel de contribución de cada uno de ellos al modelo. Finalmente, el tiempo que toma para probar un gran número de redes neuronales puede ser demasiado.

El programa NeuroShell 2© proporciona herramientas para evaluar la contribución de cada dato de entrada en la red, ya sea mediante un análisis de pesos en el

entrenamiento de la red *backpropagation* (método supervisado), o mediante un análisis genético adaptativo (método no supervisado).

Dado lo anterior, se estableció emplear un método *no supervisado* denominado Redes Neuronales de Regresión General (*General Regression Neural Networks*) o GRNN el cual utiliza un algoritmo genético para la selección de las variables de entrada, a fin de emplear en dicho modelo aquellas que tuvieran la mayor contribución a la red. Este método es una arquitectura que se encuentra disponible en el programa NeuroShell 2©. Este tipo de redes son conocidas por su capacidad para entrenar rápidamente conjuntos de datos, que produce salidas en valores continuos y permite entradas multidimensionales (Ward Systems Group, 1996).

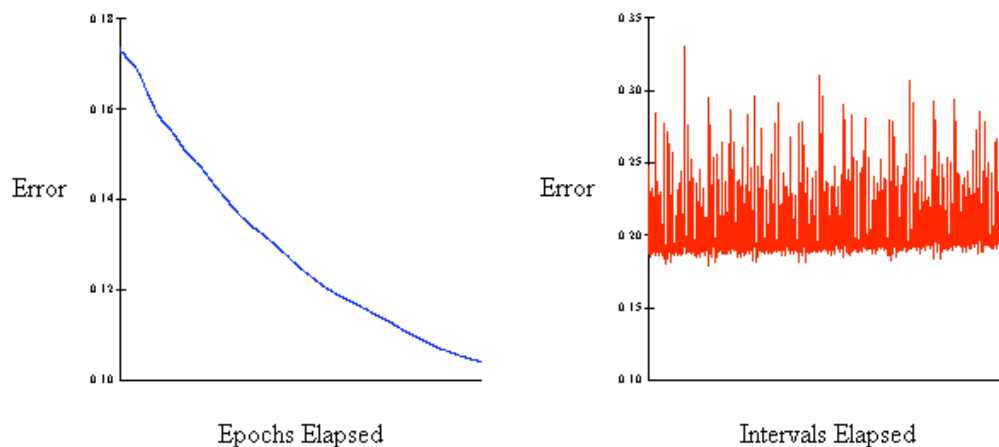
4.5. Criterios para la evaluación de redes

Los criterios para la evaluación de los modelos de redes neuronales fueron los recomendados en los estudios de Vega-García *et al.* (1996), en el estudio de Fizelew y García (2000) y principalmente en las sugerencias contenidas en el mismo programa NeuroShell 2©. Aunque si bien tales parámetros simplifican en gran medida el análisis del desempeño de las redes neuronales, el criterio más eficiente es la verificación de las observaciones correctas después de aplicada la red.

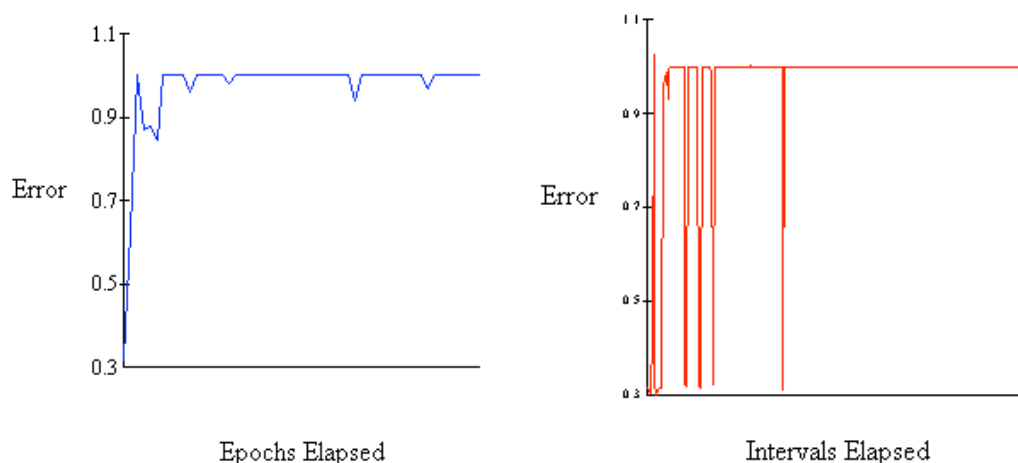
4.5.1. Error mínimo promedio (EMP)

El EMP constituye el más bajo nivel para el error que la red calcula para los datos en el set de entrenamiento. Las mejores redes son aquellas que logran una gráfica EMP versus Épocas (Gráfica 4.1) cuando desciende mas cercana al valor cero. Conforme la

red va aprendiendo, va minimizando el error durante el entrenamiento, hasta que se alcanza un punto en que la red ya no aprenderá más (Fizelew y García, 2000).



Gráfica 4.1. Error Promedio vs. Épocas/Intervalos evaluados del conjunto de datos de entrenamiento (izquierda) y del conjunto de datos de prueba (derecha) en NeuroShell 2© durante un aprendizaje normal.



Gráfica 4.2. Gráficas de Error Promedio vs. Épocas/Intervalos evaluados del conjunto de datos de entrenamiento (izquierda) y del conjunto de datos de prueba (derecha) en NeuroShell 2© donde no hay aprendizaje.

Cuando el entrenamiento es defectuoso, se observa en la gráfica que el EMP no disminuye, ni en la etapa de entrenamiento ni en la etapa de prueba (Gráfica 4.2). Asimismo, la gráfica se muestra errática e irregular, debido a que la red no logra alcanzar los valores que tiene como datos de salida .

4.5.2. Coeficiente de determinación múltiple (R^2)

Es el coeficiente de determinación múltiple utilizado como indicador estadístico y que es aplicado en el análisis de regresión múltiple, compara la precisión de un modelo resultante con un modelo base. Un ajuste perfecto sería aquel en que la R^2 es 1, un buen ajuste es aquel cuando dicho valor es cercano a 1, y un pobre ajuste cuando el valor es cercano a 0. Este valor está representado por:

$$R^2 = 1 - \frac{SSE}{SSyy}$$

Donde:

$$SSE = \sum (y - \hat{y})^2, \quad SSyy = \sum (y - \bar{y})^2,$$

" y " es el valor original, el valor " \hat{y} " son los valores predichos y " \bar{y} " es el valor promedio.

Este parámetro si bien es útil, no es el decisivo para producir en la red buenos resultados, especialmente en redes de clasificación. El buen funcionamiento de ésta es en función del número de clasificaciones correctas.

4.5.3. Coeficiente de correlación lineal de Pearsons (CCr)

Este valor representa que tan fuerte es la relación entre los valores de los datos de salida (predichos) en comparación con los valores reales, siendo un coeficiente que oscila entre -1 a 1. Un valor de r cercano a 1 determina una fuerte relación lineal positiva, en tanto que un valor de r cercano a -1 denota una fuerte relación lineal negativa. Cuando r es cercano a 0, significa que no hay relación lineal. Se pueden observar estos resultados graficando los valores de las salidas de actuales y predichas.

La fórmula utilizada para determinar el coeficiente de correlación r es:

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}}$$

Donde:

$$SS_{xy} = \sum XY - \frac{\sum X \sum Y}{n} \quad SS_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} \quad SS_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$$

“n” es igual al número de patrones, donde “x” se refiere al conjunto de datos de salida iniciales, en tanto que “y” se refiere a los datos de salida predichos.

El procedimiento de selección del mejor desempeño de la red, se realiza cuando el error promedio mínimo del conjunto de datos de prueba alcanza el valor más bajo. Aún cuando el conjunto de datos de entrenamiento conforme pasa el tiempo llega a reducirse hasta 0.05, el conjunto de datos de prueba no disminuye más. Esto es porque no es posible alcanzar mayor nivel de aprendizaje que el que se consigue con el entrenamiento. Aún si no se limitara el número de generaciones, no se consigue disminuir el EPM del conjunto de datos de prueba. Para evitar el sobreentrenamiento, el programa NeuroShell 2© utiliza la *calibración*, con la cual se guarda el mejor conjunto de datos de entrenamiento, con el objeto de que en la aplicación de la red, sea éste último el utilizado para la evaluación de los datos en esta arquitectura, con los pesos y parámetros seleccionados.

Posteriormente, durante la etapa de aplicación de la red, se obtienen los siguientes valores que evalúan el desempeño general de la red, mediante una comparación de los datos de salida iniciales con los datos de salida calculados en el procedimiento. Dichos parámetros son:

- Coeficiente de correlación (R^2)
- Coeficiente de correlación de Pearson (r^2)
- Error medio cuadrado (*Mean squared error*)
- Error medio absoluto (*Mean absolute error*)
- Error mínimo absoluto (*Minimum absolute error*)
- Error máximo absoluto (*Maximum absolute error*)
- Coeficiente de correlación "r" (*Correlation coefficient r*)
- Porcentaje dentro del 5%, 10%, 20%, 30% y mayor al 30% (*Percent within 5%, 10%, 20% and 30% and over 30%*)

Si se obtiene una R^2 razonablemente alta, la red se considera candidata para ser evaluada. De ser bajo el valor, se modifican los parámetros hasta agotar las posibilidades.

4.6. Procesamiento de imágenes del sensor AVHRR

En el año 1999 se inició el análisis con imágenes AVHRR en México. Diariamente la CONABIO recibe del Instituto de Geografía de la Universidad Nacional Autónoma de México (UNAM) imágenes nocturnas y diurnas del sensor AVHRR (*Advanced Very High Resolution Radiometer*) del satélite NOAA-14 (*National Oceanic and Atmospheric Administration*). La imagen nocturna es registrada entre las 3:00 y 5:00 hrs. El Instituto de Geografía realiza la calibración y georreferencia de las imágenes, y las envía a la CONABIO dos veces en el día, el primer envío a las 10:00 de la mañana y el segundo a las 11:00, con el objeto de reducir los errores que puedan ocurrir en la transferencia de la información. La imagen diurna es registrada entre las 15:00 y la 18:00 hrs. Al igual que las imágenes nocturnas el Instituto de geografía envía dos veces durante la tarde.

Las imágenes presentan un desplazamiento principalmente en los extremos de la República, dependiendo del paso del satélite y la extensión del país captada (CONABIO, 2003).

Tabla 4.2. Resoluciones espaciales y espectrales del sensor AVHRR.

Resolución espacial	1.1. por 1.1 km al nadir	
	Canal	Resolución
Resolución espectral en micrómetros	1	0.58 - 0.68 visible
	2	0.725 - 1.10 infrarrojo cercano
	3	3.55 - 3.93 infrarrojo térmico
	4	10.3 - 11.3 infrarrojo térmico
	5	11.5 - 12.5 infrarrojo térmico
Resolución temporal	12 horas	
Sensibilidad canal 3	-50°C a 50°C	

Saturación del píxel de NOAA-14 AVHRR BANDA 3

La saturación del píxel se considera ocurre a los 50°C; sin embargo, se necesita tan sólo el 0.1% del área que cubre el píxel con temperaturas entre 250°C y 500°C. Para que el píxel sea saturado el punto de calor tiene más de 500°C. Aún cubriendo un área de solo 0.01% del píxel, es suficiente para saturarlo. Por tal motivo no se puede hablar de número de incendios, ni de superficie afectada (Tabla 4.2.).

Extracción de información

Con base a la información obtenida en Conabio, 2003; se considera que un píxel es candidato a ser punto de calor cuando se complen las siguientes consideraciones:

a) Imágenes nocturnas: En el canal 3 de la imagen (que registra la temperatura de brillantez) presenta una temperatura mayor a 293°K = 25°C y la diferencia entre canal 3 y canal 4 es mayor a ocho grados centígrados.

Posteriormente se realiza una discriminación por la influencia de nubes: cuando en el canal 5 la temperatura es menor a $272^{\circ}\text{K} = -1^{\circ}\text{C}$ se considera que es nube.

b) Imágenes diurnas: El canal 3 mayor a 38°C , la diferencia entre el canal 3 y 4 es mayor a 8°C y canal 2 menor a 12 %.

Condiciones:

Condición 1: La suma del canal 1 y canal 2 mayor a 60%

Condición 2: Temperatura en el canal 5 menor de $272^{\circ}\text{K} = -1^{\circ}\text{C}$

Condición 3: La suma de canal 1 y canal 2 mayor a 40%

Condición 4: Temperatura en el canal 5 menor a $280^{\circ}\text{K} = 7^{\circ}\text{C}$

- Será una nube si cumple la condición 1 ó 2 ó cumpla la condición 3 y 4.
- Se aplican mascararas de cuerpos de agua, luces estables y los límites del país, para eliminar posibles confusiones.
- Posteriormente se realiza una sobreposición de los resultados con información adicional como es el tipo de vegetación, nombre de Municipios, Estados y Áreas Naturales Protegidas.
- Una vez obtenida la imagen de incendios se procede a convertirlos a un mapa.

La información sobre la detección de puntos de calor con imágenes nocturnas e imágenes diurnas es proporcionada diariamente como una tabla con los datos del municipio ubicación, extensión aproximada, y otras características. Asimismo, de las vistas diurna y nocturna de la imagen sin procesar y un mapa dinámico para su consulta en línea.

5. MATERIALES Y EQUIPO

5.1. Caracterización de la zona de estudio

Datos obtenidos del Gobierno del Estado de México, 2003, mencionan que el Estado de México se localiza en la parte central de la República Mexicana, entre los 18° 21' y 20° 17' de latitud Norte y los 98° 35' y 100° 36' de longitud Oeste. Está constituido por 125 municipios los cuales se encuentran asentados en una extensión de 22,499.95 Km² (el 1.15 % del total nacional). Políticamente la entidad limita al Norte con los estados de Querétaro e Hidalgo; al Este con Puebla y Tlaxcala; al Sur con Morelos y Guerrero y al Oeste con Michoacán y una pequeña porción de Guerrero. También limita con el Distrito Federal, rodeándolo hacia el Norte, Oriente y Occidente. (Figura 5.1)



Figura 5.1. Ubicación de la zona de estudio

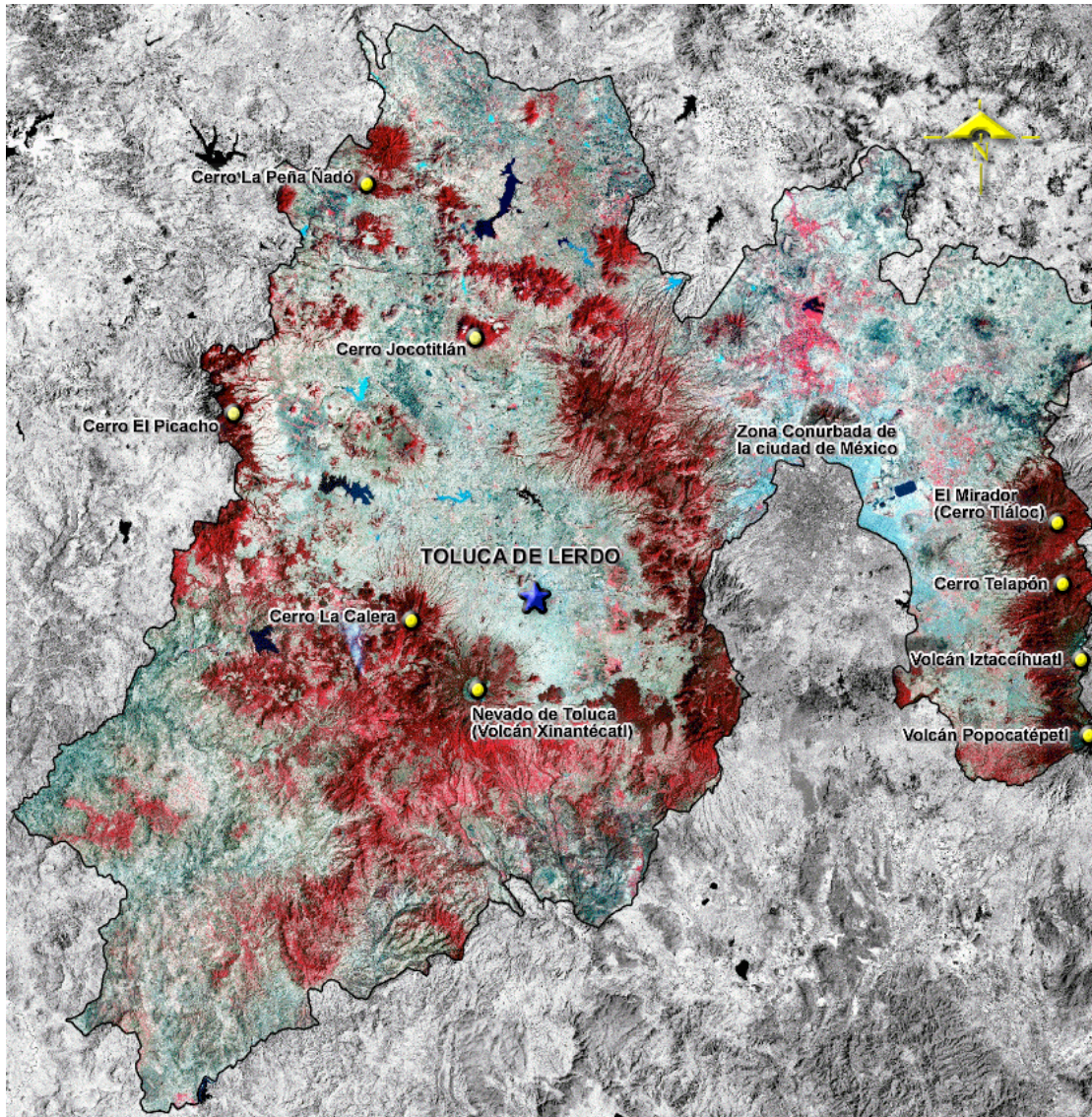


Figura 5.2. Imagen satelital en falso color del Estado de México
 (Fuente: INEGI, 2007)

Clima y Relieve

El relieve de la entidad se caracteriza por presentar grandes y pequeños valles agrícolas, sierras y volcanes aislados que proporcionan la diversificación de altitudes, tipos de roca, yacimientos minerales, suelos, climas, vegetación, flora, fauna y una gran diversidad de actividades económicas que generan una gama de regiones y paisajes característicos del territorio estatal (Figura 5.2).

El clima predominante es el cálido húmedo en las zonas bajas del suroeste; semisecos y secos en la región norte y noroeste, los templados se localizan en valles y planicies y los fríos en las zonas montañosas (Gobierno del Estado de México, 2003).

Hidrología

La precipitación media anual es del orden de 1,000 mm/año, lo que implica un volumen medio anual precipitado de 22,467 millones de m³. De este volumen se evapotranspira un promedio de 17,773 millones de m³, por lo que finalmente la disposición es de 4,694 millones de m³, de éstos 3,707 millones de m³, forman los escurrimientos superficiales y 967 millones de m³, presentan la infiltración a los acuíferos locales.

Dentro de la entidad se originan 3 de las 37 regiones hidrológicas del país, la del río Pánuco que cubre la porción Norte del Estado, con una superficie de 902,628 ha, la cual se caracteriza por la gran disponibilidad de acuíferos, abarcando las zonas de Cuautitlán, Teotihuacan, Texcoco y Chalco, las cuales forman parte de la Cuenca del Valle de México.

La región del Río Lerma con un área de 539, 545 ha, en la parte centro-oeste, cuenca donde existen una gran cantidad de bordos que permiten captar escurrimientos de afluentes al colector principal, y además contribuyen al control de inundaciones en las partes bajas de los Valles de Ixtlahuaca y Atlacomulco.

Finalmente, la región del río Balsas en la porción Sur con 957,154 ha de superficie, donde existen importantes acuíferos en calizas cretácicas.

Respecto a la recarga de los acuíferos estudiados, ésta es básicamente por infiltración vertical, producto de la precipitación y en menor proporción por flujo horizontal subterráneo, siendo la calidad del agua aceptable para el consumo humano. De los 159 escurrimientos superficiales existentes en la entidad, es decir ríos y arroyos, el 43.3% se concentra en la Cuenca Hidrológica Moctezuma, la cual presenta una extensión de 13,791 ha aproximadamente.

Recursos naturales

En cuanto a riqueza de recursos naturales, el Estado de México cuenta con 609,000 ha. arboladas, 560,000 Ha. de bosques característicos de clima templado frío y 49,000 ha. de matorral, chaparral y selva baja caducifolia. La mayor concentración y diversidad forestal se encuentra en el interior de la cuenca del Río Balsas y en los principales sistemas montañosos como las sierras de Monte Alto y Monte Bajo, de las Cruces, del Ajusco, de Río Frío y Sierra Nevada; en menor proporción en la Sierra de San Andrés, Mesa de los Pinos y la Sierra de Carimangacho al norte del Estado (Gobierno del Estado de México, 2003).

Con respecto al recurso suelo, por su origen y formación en la entidad predominan principalmente los suelos Feozem en 534,635 ha, (el 23.7% del total estatal), los andosoles en 459,489 ha, (el 20.4%) y regosoles abarcando una superficie de 264,388 ha, (el 11.7%).

Uso del suelo

De acuerdo a INEGI, 2007, los suelos Feozem se caracterizan por presentarse en lugares con clima templado siendo muy aptos para la agricultura. También se

encuentran en zonas con acumulación de nutrientes y áreas de poca pendiente. Los principales usos del suelo en el Estado de México son el agrícola (47.43%), forestal (33.0%) y pecuario (16.9%), lo que en conjunto suma el 85.9% del total de la superficie de la entidad. (Tabla 5.1 y Figura 5.3). De las 609,000 ha. arboladas, 198,000 son consideradas como aprovechables con incrementos anuales de 950,000 m³ que representan el potencial de aprovechamiento anual de éstos. El 42% corresponde a “pino” (*Pinus*), el 27% a “oyamel” (*Abies*) y el 22% a “encino” (*Cupressus*) (Figura 5.4).

Dentro de la entidad se localizan diez Parques Nacionales, los cuales abarcan una extensión en territorio mexiquense de 98,692.00 ha; 24 Parques Estatales abarcando una superficie de 282,059 ha.; cinco Parques Municipales en 193.72 ha; y once Reservas Ecológicas en una superficie de 100,866.71 ha, así como diez Parques sin decreto, ocupando una extensión de 679 ha. La ubicación geográfica de la entidad (sobre el Eje Neovolcánico Transversal), propicia la existencia de una gran diversidad de fauna, entre la que destacan los mamíferos, aves, reptiles y anfibios, contando con 111 especies distribuidas en 4 órdenes y 45 géneros, predominando los anfibios y reptiles. Respecto a las especies endémicas, se han detectado cinco especies de anfibios, una de reptiles y una de mamíferos.

Del 14.1% restante, el 4.7% se encuentra destinado al uso urbano e industrial, 1.4% a cuerpos de agua y 7.8% para otros usos. Considerando la regionalización por cuencas hidrográficas, más de la mitad de la superficie de la entidad puede ser utilizada para labores agrícolas, sin embargo, las condiciones físicas de la zona (topografía, clima, suelo, etc.), determinan diferentes alternativas de uso con distribución irregular. Con respecto al uso forestal, los bosques del Estado de México, no sólo ocupan grandes superficies, sino que presentan condiciones favorables tales que permiten planear

esquemas de explotación racional a través de la industrialización o comercialización del recurso maderable y no maderable. Del total de la superficie forestal, se advierte que existen grandes extensiones de bosques donde se encuentran especies forestales tales como pino, oyamel, cedro, encino, eucalipto, etc., así como diferentes tipos de selvas.

Tabla 5.1. Principales especies y utilidad por uso de suelo

Concepto	Nombre científico	Nombre local	Utilidad
Agricultura			
47.43 % de la superficie estatal	<i>Zea mays</i>	Maíz	Comestible
	<i>Phaseolus vulgaris</i>	Frijol	Comestible
	<i>Hordeum sativum</i>	Cebada	Comestible
	<i>Avena sativa</i>	Avena	Comestible, Forraje
	<i>Solanum tuberosum</i>	Papa	Comestible
Pastizal			
14.47 % de la superficie estatal	<i>Bouteloua gracilis</i>	Navajita	Forraje
	<i>Aristida divaricata</i>	Zacate	Forraje
	<i>Buchloe dactyloides</i>	Zacate chino	Forraje
	<i>Muhlenbergia rigida</i>	Zacatón	Forraje
Bosque			
27.81 % de la superficie estatal	<i>Abies religiosa</i>	Oyamel	Comercial, Industrial
	<i>Pinus montezumae</i>	Ocote blanco	Comercial, Industrial
	<i>Pinus teocote</i>	Pino chino	Comercial, Industrial
	<i>Quercus rugosa</i>	Encino quebracho	Comercial, Industrial
	<i>Quercus laurina</i>	Encino laurelillo	Comercial, Industrial
Selva			
5.67 % de la superficie estatal	<i>Acacia cymbispina</i>	Huizache	Forraje, Doméstico
	<i>Ipomoea wolcottiana</i>	Cazahuate	Ornamental, Medicinal
	<i>Bursera copallifera</i>	Copal	Medicinal, Comercial
	<i>Eysenhardtia polystachya</i>	Vara dulce	Forraje, Medicinal
	<i>Haematoxylon brasiletto</i>	Palo brasil	Artesanal, Industrial
Matorral			
0.90 % de la superficie estatal	<i>Opuntia streptacantha</i>	Nopal	Comestible, Comercial
	<i>Mimosa biuncifera</i>	Uña de gato	Medicinal, Leña
	<i>Jatropha dioica</i>	Sangre de drago	Medicinal
	<i>Acacia farnesiana</i>	Huizache	Medicinal, Artesanal
	<i>Opuntia tomentosa</i>	Nopal	Comestible

(Fuente: INEGI, 2007)

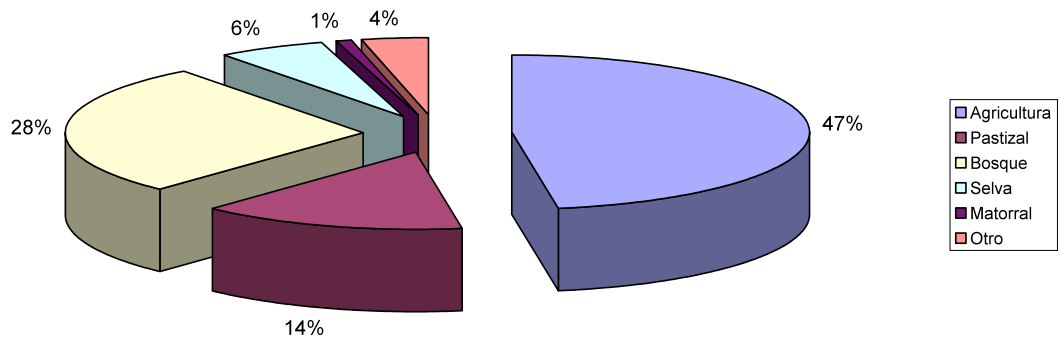


Figura 5.3. Cobertura por tipo de uso de suelo
(Fuente: INEGI, 2007)

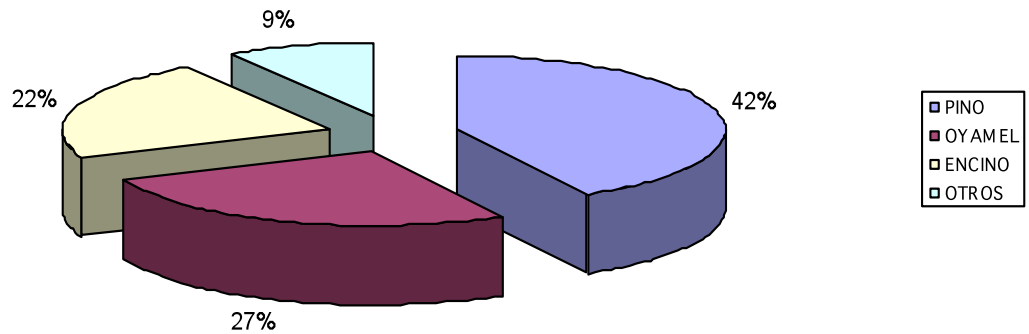


Figura 5.4. Aprovechamiento potencial por tipo de bosque
(Fuente: Gobierno del Estado de México, 2003)

5.2. Obtención del material cartográfico

5.2.1. Información obtenida de la CONABIO

La Comisión Nacional para el Uso de la Biodiversidad (CONABIO) es una entidad gubernamental que se encarga de la evaluación y el manejo de aspectos relacionados con los recursos naturales en México. La CONABIO utiliza una gran cantidad de información geográfica, principalmente imágenes de satélite del sensor AVHRR que son

proporcionadas por el Instituto de Geografía de la UNAM, a las cuales se les aplican técnicas de análisis de percepción remota para obtener información sobre aspectos relacionados con características de la corteza terrestre relacionadas con la biodiversidad. Mucha de esta información es luego procesada en Sistemas de Información Geográficos para poner a disposición del público de forma gratuita diversos materiales sobre recursos naturales en el país y en otras partes del continente.

Uno de los programas que CONABIO realiza desde 1999 es el programa para la detección de puntos de calor, los *shapefiles* (archivos en formato vector manejados por la compañía *Environmental Systems Research Institute Inc.*) correspondientes a las variables que fueron planteadas para el estudio se presentan en la Tabla 5.2:

Tabla 5.2. Información digital obtenida de CONABIO.

TÍTULO	AÑO	ESCALA	PROYECCIÓN
Cabeceras Municipales	2000	No especificado	CÓNICA
Precipitación Media Anual	1990	1 : 4,000,000	CÓNICA
Regímenes de humedad del suelo	1992	1 : 4,000,000	CÓNICA
Temperatura Media Anual	1990	1 : 4,000,000	CÓNICA
Mapa base del Estado de México	1999	1 : 250,000	CÓNICA

Asimismo se obtuvieron de la misma página los metadatos correspondientes a cada imagen. Las imágenes se encuentran disponibles en coordenadas geográficas UTM (Universal Transversal Mercator) y en proyección Cónica de Lambert, siendo utilizadas las primeras para este estudio empleando un Datum de NAD 27-14n. Cada mapa una vez procesado corresponde a una variable de entrada, para la red neuronal, siendo doce las empleadas en el estudio. Se obtuvieron los *shapefiles* de los PUNTOS DE CALOR (PC) en el territorio nacional detectados por el sensor AVHRR correspondientes a los condensados anuales de los años 2000, 2001, 2002, 2003, 2004 2005 y 2006

(este último año en el período comprendido entre los meses de de enero a julio). Estos PC son los que constituyeron las variables de salida (también denominados como OUTPUT) para la red neuronal propuesta (Tabla 5.3).

Tabla 5.3. Puntos de calor detectados en el período 2000 a 2006 en el Estado de México.

AÑO	PUNTOS DE CALOR DETECTADOS
2000	338
2001	220
2002	183
2003	359
2004	162
2005	376
2006	301
TOTAL	1,939

Nota: Se incluyen los PC diurnos y nocturnos detectados en este período, solamente en las áreas con vegetación susceptible a incendios forestales .

Es importante mencionar que la información obtenida no especifica de esos puntos de calor cuales corresponden a incendios forestales, ya que la metodología aplicada para determinar cuales pixeles se consideran puntos de calor solo determina la temperatura del mismo.

5.2.2. Material obtenido de la PROFEPA

Fue proporcionada por la Delegación de la Procuraduría Federal de Protección al Ambiente (PROFEPA) en el Estado de México, información cartográfica en formato *shapefile* recabada por esta Institución sobre aspectos humanos relacionados directa o indirectamente con riesgo o peligro de incendio forestal, o por resultar de interés en el Estado de México, del cual ya no fue necesario adaptarlo debido a que dicha información se encuentra georeferenciada al mismo Datum NAD 27 y al mismo sistema

de coordenadas UTM en el que se está manejando la información de éste estudio (Tabla 5.4).

Tabla 5.4. Información digital obtenida de PROFEPA-EDOMEX.

TITULO	ESCALA	PROYECCIÓN
Líneas de conducción eléctrica	1:250,000	UTM
Áreas naturales protegidas	1:250,000	UTM
Aprovechamientos forestales	1:250,000	UTM
Calidad ecológica	1:250,000	UTM
Modelos digitales de elevación	1:250,000	UTM
Tipo de vegetación INEGI	1:250,000	UTM

Finalmente, se obtuvieron las cartas topográficas escala 1:250,000 en formato digital elaborado por Instituto Nacional de Estadística, Geografía e Informática y sus toponímicos, mismas que fueron empleadas para generar el mapa de curvas de nivel del Estado de México.

5.3. Software y equipo utilizado

El procesamiento y la elaboración del material cartográfico se realizó en el programa Arcview© versión 3.2a con los módulos instalados de Image Analisys©, 3D Analisys© y Spatial Analyst©; asimismo, el programa Idrisi© fue utilizado para la conversión de formatos *shapefile* a formato vector, posteriormente a formato raster y de raster al formato ASCII. Para el manejo y procesamiento de las bases de datos fue utilizada la hoja de cálculo Microsoft Excel©. El diseño y aplicación de las redes neuronales se realizó en el programa NeuroShell 2© versión 3.0.

El material cartográfico, así como los datos para las redes neuronales fueron procesados en su totalidad en una computadora PC personal con un procesador

Pentium© 4 HT de 3.1 GHz y 512 MB en RAM en el sistema operativo Windows© XP Professional. Para la realización de comprobaciones en campo, fue utilizado también un receptor GPS de la marca Garmin©, modelo *E-Trex Legend* con el Datum NAD 27 y sistema de coordenadas UTM, procurándose un mínimo de 5 metros de precisión en la lectura, con al menos 8 satélites registrados en paralelo. En la mayoría de los casos por tratarse de áreas abiertas se conseguía una visión despejada del cielo, a excepción de las zonas arboladas o montañosas, donde la precisión disminuye en relación a la densidad de cobertura de copa.

6. METODOLOGÍA

6.1. Preprocesamiento del material cartográfico

Una vez obtenido el material cartográfico se procedió a generar las bases de datos espaciales en formato *shapefile* de cada una de las trece variables consideradas para este estudio. Se procuró incluir variables de entrada relacionadas con actividades humanas, así como variables relacionadas a características del terreno y del ambiente, todas empíricamente relacionadas directa o indirectamente al riesgo o peligro de incendio forestal, dado que estudios anteriores muestran que dichos tipos variables son útiles para predecir incendios forestales (Vega-García *et al.*, 1999). Las trece variables seleccionadas son:

- a. Distancia a la carretera más cercana
- b. Zonas de campamento y recreación
- c. Areas naturales protegidas o bajo algún tipo de protección
- d. Áreas bajo aprovechamiento forestal
- e. Líneas de conducción eléctrica
- f. Vegetación susceptible a incendios forestales
- g. Pendientes categorizadas
- h. Altitudes categorizadas
- i. Regímenes de humedad del suelo
- j. Temperatura media
- k. Precipitación media
- l. Calidad ecológica

La información digital enumerada anteriormente se definió en base a diferentes los siguientes criterios:

6.1.1. Límite del Estado de México (EDO_MEX)

Utilizando el conjunto MAPA BASE DEL ESTADO DE MEXICO, del cual se seleccionó el MAPA DE LIMITE ESTATAL, el cual sería posteriormente utilizado para la delimitación del área del Estado de México en los mapas utilizados en éste estudio, mediante el empleo de la utilidad Wizard Geoprocessing del programa ArcView©. Con base en este mapa, se realizaron recortes de las diferentes características del estudio, con el fin de obtener concordancia en cada uno de los píxeles que forman la imagen.

6.1.2. Distancia a cabeceras municipales (CAB_MUN)

En un principio se pensó utilizar el mapa de POBLACIONES2000 proveído por la CONABIO, el cual incluye la totalidad de asentamientos humanos en el territorio nacional, seleccionando aquellas que se encontrarían dentro del área que comprende el Estado de México, para posteriormente determinar un *buffer* (creación de un área o superficie a partir de un punto, una línea o un polígono) de cinco kilómetros de cada población. Sin embargo, se verificó que existía un gran número de poblaciones distribuidas dentro del territorio, por lo que el *buffer* de todas estas poblaciones se empalma en distancias menores a cinco kilómetros. Para subsanar esto, se determinó seleccionar la población que se considere como las de mayor concentración en habitantes, además de que tuviese importancia como centro económico y turístico y que se encontraran distribuidas en el Estado. De tal manera, se decidió seleccionar a partir del mapa de MUNICIPIOS del mismo conjunto MAPA BASE DEL ESTADO DE MEXICO las cabeceras municipales, ampliándose el *buffer* a diez kilómetros. Esto en base a que

se considerarían como los principales lugares donde a partir de los cuales se realizan diversas actividades antropogénicas como turismo, comercio y otros. Se asume que a menor distancia del sitio poblado, mayor es la probabilidad de presentar un punto de calor.

6.1.3. Distancia a la carretera más cercana (DIST_CARR)

A partir del mapa de CARRETERAS, el cual se encuentra contenido dentro del conjunto de MAPA BASE DEL ESTADO DE MEXICO, se incluyeron vías principales y secundarias aplicándose a los mismos un *buffer* de diez kilómetros. No se incluyeron terracerías ni veredas debido a que, como el caso del mapa de poblaciones, hay una alta densidad de éstos últimos, por lo que no era factible el uso de *buffers* a tan corta distancia. Se consideró que a menor distancia de una vía principal de comunicación, mayor es la probabilidad de la presencia de un punto de calor.

6.1.4. Mapa de zonas de campamento (ZC)

Para determinar las zonas con presencia de campistas o paseantes se determinó utilizar en primera instancia información escrita (revistas y mapas de turismo y actividades de esparcimiento), sobre las ubicaciones de los sitios que son visitados para realizar actividades al aire libre.

Posteriormente, ésta información fue transformada a formato espacial y combinada con los mapas de MUNICIPIOS y TIPOS DE VEGETACION, definiéndose *buffers* de un kilómetro de distancia a partir de la ubicación de sitios de interés turístico, como lo son paisajes escénicos, grutas, formaciones rocosas y otras similares. De esta forma se

generó un mapa con polígonos que limitan las zonas que acostumbran ser frecuentadas por paseantes y campistas. En este caso, es más probable que dentro de una zona de campamento y esparcimiento, ocurra un punto de calor.

6.1.5. Áreas naturales protegidas (ANP)

Usando información proporcionada por la Delegación Estatal de la Procuraduría Federal del Ambiente (PROFEPA) se complementó el mapa de ÁREAS NATURALES PROTEGIDAS de la CONABIO el cual sólo incluía las áreas federales. Con esta información se agregaron aquellas áreas estatales en sus diversas modalidades de protección (parques estatales, zonas de protección ecológica, etc.). Se consideró que es más probable que un punto de calor ocurra fuera de un área bajo algún tipo de protección.

6.1.6. Aprovechamientos forestales (APROV)

A partir de la información proporcionada por la PROFEPA en el Estado de México se obtuvo la relación de los aprovechamientos forestales autorizados en el Estado. Se seleccionaron aquellos cuya autorización estaba vigente desde el año 1999 a la fecha, eliminándose aquellos que dejaron de tener actividades antes de ese año y que por tanto ya no estaban laborando en el período de estudio. Posteriormente se representaron mediante puntos en el mapa, y se les aplicó un *buffer* de dos kilómetros a cada característica, considerándose esta distancia como adecuada en base a que es el máximo aproximado que se desplazan las personas dentro del área autorizada para el aprovechamiento. Se consideró que es más probable que un punto de calor ocurra dentro de un área donde se están desarrollando aprovechamientos forestales.

6.1.7. Líneas de conducción eléctrica (LIN_COND)

Se generó del conjunto MAPA BASE DEL ESTADO DE MEXICO aplicando un *buffer* de quinientos metros en el mapa LINEAS DE CONDUCCION del recorrido que siguen las torres que soportan las líneas de conducción eléctrica. Se seleccionó esta variable a fin de determinar si existía alguna relación de los incendios al efectuarse labores de limpieza en el interior del área considerada como de derecho de vía de esa estructura. Es más probable que un punto de calor ocurra al encontrarse dentro de un área donde existan líneas de conducción eléctrica.

6.1.8. Vegetación susceptible a incendios (US_SUELO)

Del mapa de TIPO DE VEGETACIÓN modificado por INEGI proporcionado por PROFEPA Estado de México, se seleccionaron las áreas que por su tipo son más susceptibles a ser afectadas por incendios forestales. No se consideraron las áreas urbanas, terrenos de uso completamente agrícola ni los cuerpos de agua. Los nueve tipos de vegetación seleccionados y su régimen de fuego fueron:

Tabla 6.1 Tipo de vegetación y su régimen de fuego.

TIPO DE VEGETACIÓN	REGIMEN DE FUEGO
Bosque templado (coníferas y encinos)	Dependiente del fuego
Bosque mesófilo de montaña	Sensible al fuego
Selva subhúmeda	Sensible al fuego
Pastizal cultivado	Dependiente del fuego
Pastizal natural (semidesértico)	Dependiente del fuego
Matorrales de zonas áridas	Sensible ó dependiente al fuego (según tipo)
Pradera de alta montaña	Dependiente ó independiente del fuego (según tipo)
Plantaciones forestales	No aplica (por no ser vegetación natural)
Vegetación halófila y gipsófila	Dependiente del fuego

Dado que sólo la superficie comprendida en estos tipos de vegetación se considera cómo el área de interés para el estudio, éstas se sometieron al proceso de entrenamiento y prueba dentro de la red neuronal artificial sin tomar en cuenta la totalidad de las celdas de la matriz de 40,000 observaciones. Se determinó que tal conjunto de píxeles (8,938 observaciones) conformarían el conjunto de celdas de *Vegetación Susceptible a Incendios Forestales* o datos-objetivo, mismas que una vez finalizado el proceso de la RNA estas serían reintegradas a la matriz original a fin de conformar una nueva imagen de 200 filas por 200 columnas. Las categorías de vegetación fueron acomodadas en orden, de acuerdo a su mayor o menor susceptibilidad a incendiarse, de acuerdo a si tipo, cantidad de combustible producido, etc. Young (1996) menciona que los incendios forestales son mucho más comunes en los bosques de coníferas que en las comunidades de árboles caducifolios de hojas anchas. Datos de INEGI (2007) mencionan que los pastizales poseen la mayor superficie afectada por los incendios forestales (2,750.30 ha.) que los matorrales (284.25 ha.). Asimismo, solo fueron afectadas en 2007 tres y media hectáreas de renuevo.

6.1.9. Pendientes categorizadas (PEND_CAT)

Dado que no se contaba con un mapa de pendientes de la totalidad del área de estudio, se realizó una unión de los modelos digitales de elevaciones (MDE) de seis cartas de escala 1:250,000 con el módulo CONCAT (concatenar) del programa Idrisi© para posteriormente delimitar el límite del Estado de México. A partir de éste MDE fue posible generar un mapa de pendientes con el módulo SLOPE (pendiente) del programa Idrisi©. Posteriormente se realizó una categorización en tres grupos, considerándose tres grupos: en el primero se agruparon las pendientes del 0 al 15 por

ciento, en el siguiente grupo, las pendientes entre el 16 al 30 por ciento, y finalmente el tercer grupo con las pendientes mayores al 30 por ciento. Se consideró que la presencia de puntos de calor es más probable cuando hay un mayor grado de pendiente en el terreno, ya que es una de las condiciones más frecuentes para que se desarrolle un probable incendio forestal.

6.1.10. Altitudes categorizadas (ALT_CAT)

Usando como base el MDE del Estado de México se generó un mapa de CURVAS DE NIVEL a partir del cual se seleccionaron cuatro grupos bajo el siguiente criterio: el primer grupo conformado por altitudes menores a los 1,000 msnm, el segundo grupo con altitudes dentro del intervalo de 1,001 a 2,000 msnm, el tercer grupo con las altitudes en el rango de de 2,001 a 3,000 msnm, y finalmente un cuarto grupo con altitudes superiores de los 3,000 msnm. Se consideró que en altitudes más bajas es más probable la presencia de puntos de calor.

6.1.11. Régimen de humedad del suelo (REHUSUEL)

A partir del mapa de REGIMEN DE HUMEDAD DEL SUELO obtenido de CONABIO, se realizó la delimitación del Estado de México, considerándose la clasificación de éste régimen en cinco grupos de acuerdo la cantidad de humedad reportada en el suelo, los cuales son Arídico, Xérico, Ústico, Údico y Acuico. Se determinó que los suelos con menor cantidad de humedad provee condiciones más factibles para la presencia de puntos de calor.

6.1.12. Temperatura media (TEMP_MED)

A partir del mapa de TEMPERATURA MEDIA ANUAL obtenido de CONABIO se realizó la delimitación del Estado de México, con lo cual se determinaron para éste territorio cinco categorías de temperatura que son: fría, semifría, templada, semicálida y cálida. Se consideró que los puntos de calor son más frecuentes cuando los ambientes se encuentran sometidos a temperaturas altas, que por lo general guarda relación con el contenido de humedad de los materiales combustibles.

6.1.13. Precipitación media (PREC_MED)

Asimismo, del mapa de PRECIPITACIÓN MEDIA ANUAL a nivel nacional se delimitó el Estado de México para obtenerse así cinco categorías de milímetros (mm/año) de agua en cinco categorías, las cuales fueron: de 400 a 600 mm, de 601 a 800 mm, de 801 a 1,200 mm, de 1,201 a 1,500 mm y de 1,501 mm en adelante. Se consideró que en áreas donde la precipitación es menor, existe una menor incidencia de los puntos de calor.

6.1.14. Calidad ecológica (CAL_ECO)

A partir de información proporcionada por la PROFEPA en el Estado de México, se utilizó un mapa denominado CALIDAD ECOLOGICA, en el cual se determinaron zonas de acuerdo al nivel de impacto producido por el hombre (mínima tala clandestina, de poca a nula presencia de paseantes, mayor diversidad de especies en los estratos arbóreo, arbustivo y herbáceo y a su estado fitosanitario). Asimismo evalúa la efectividad de la aplicación de sistemas de gestión y conservación al suelo y al ambiente y donde las propias comunidades han realizado actividades de sustentabilidad agrícola. Este factor permite determinar así cuales zonas que son más apreciadas por

su estado de conservación y se encuentra categorizada. Se conformó con base a las siguientes cinco categorías: muy baja, baja, media, alta, zona urbana y cuerpos de agua. Para este caso, se determinó que áreas donde la calidad ecológica es muy baja tienen mayor presencia de puntos de calor.

6.2. Adaptación de los datos para la RNA

Dado que los mapas generados se hallaban en formato vectorial, se hizo necesario exportarlo al programa Idrisi© en donde se procedió a transformarlos al formato raster, a fin de obtener de cada imagen una matriz de celdas de 200 filas por 200 columnas, cada celda equivalente a 1.1 km por lado, que corresponde a la resolución espacial del sensor AHVVR.

El formato raster permitió mediante una transformación en código ASCII reconocer los mapas en una hoja de cálculo de EXCEL®, distribuyendo en una columna el valor de cada una de las celdas y permitiendo con ello, su procesamiento en el programa NeuroShell 2®, el cual es un paquete que procesa en redes neuronales artificiales los datos distribuidos en variables de entrada con sus respectivas variables de salida.

Todo este proceso se repitió para cada uno de las variables utilizadas, a fin de que cada celda que conforma la matriz tuviese la información de todas las variables de entrada, para posteriormente tales datos ser introducidos en la red neuronal (Figura 6.1).

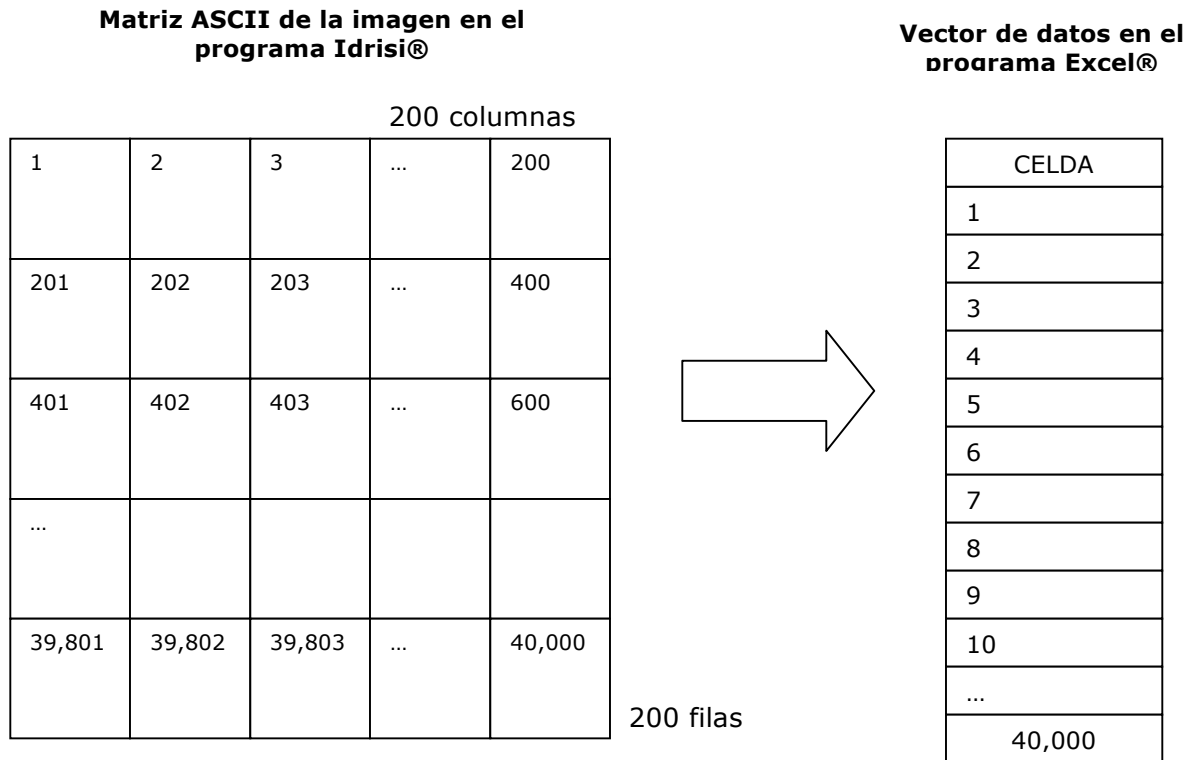


Figura 6.1. Transformación a código ASCII para la alimentación de la red.

Todas las variables fueron categorizadas entre 0 y 1. Se determinó que el valor 1 es el que corresponde a aquella categoría que de acuerdo a los criterios de aplicación se considera con mayor probabilidad de estar expuesta a presentar algún punto de calor. Los valores de las variables de entrada fueron escalados proporcionalmente mediante la fórmula:

$$Ve = (1/n) * c$$

Donde:

Ve = Valor escalado entre 0 y 1.

n = Número de categorías de la variable de entrada.

c = Categoría seleccionada en el orden establecido según clasificación.

En la Figura 6.2 se describe el procedimiento seguido para el análisis.

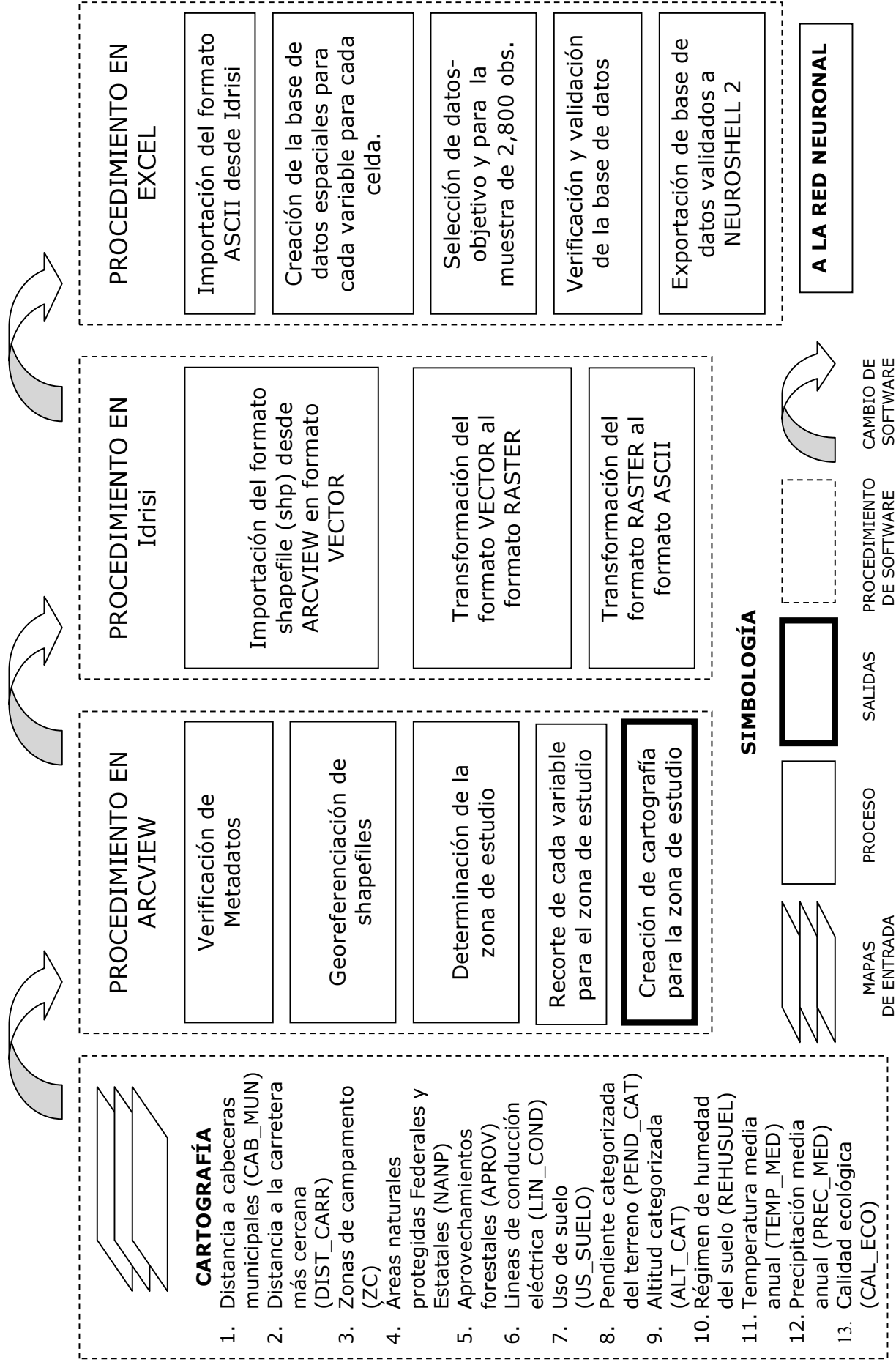
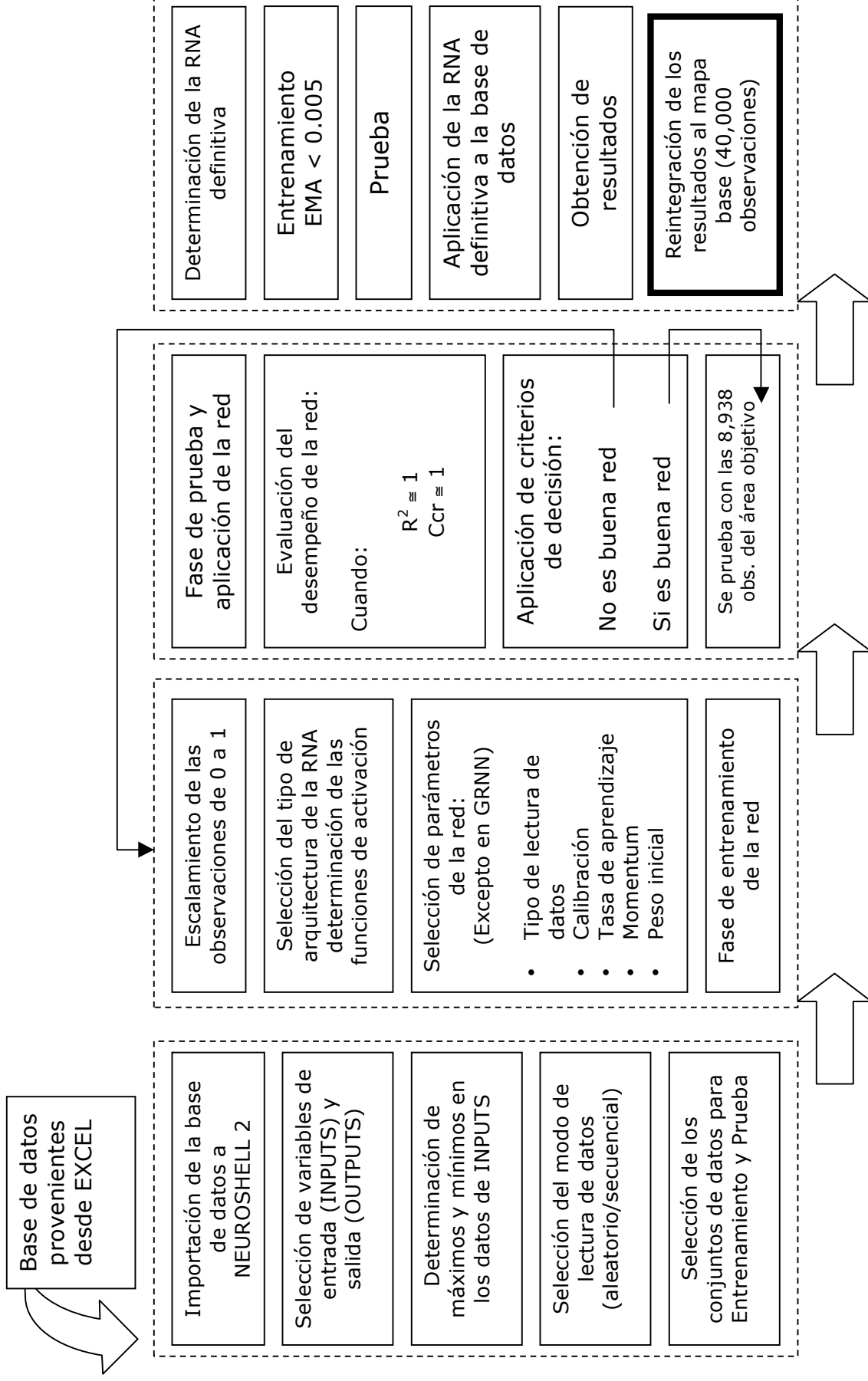


Figura 6.2. Diagrama de flujo para el procesamiento de la base de datos de la red



(Continuación). Diagrama de flujo para el procesamiento de la base de datos de la red

Se determinó que el valor 1 es el que corresponde a aquella categoría que de acuerdo a los criterios de aplicación se considera con mayor probabilidad de estar expuesta a presentar algún punto de calor.

6.3. Selección de las variables de entrada y salida

Las trece variables de entrada se dividieron artificialmente en dos grupos tomando en consideración las características de cada variable. Los grupos se denominaron FACTORES HUMANOS y FACTORES GEOGRÁFICOS. De esta forma, los *factores humanos* considerados fueron:

- a. Distancia a cabeceras municipales (CAB_MUN).
- b. Distancia a la carretera más cercana (DIST_CARR).
- c. Zonas de campamento (ZC).
- d. Áreas naturales protegidas Federales y Estatales (ANP).
- e. Aprovechamientos forestales (APROV).
- f. Líneas de conducción eléctrica (LIN_COND).

Por otra parte, se consideraron como *factores geográficos*:

- a. Uso de suelo (US_SUELO).
- b. Pendiente categorizada del terreno (PEND_CAT).
- c. Altitud categorizada (ALT_CAT).
- d. Régimen de humedad del suelo (REHUSUEL).
- e. Temperatura media anual (TEMP_MED).
- f. Precipitación media anual (PREC_MED).
- g. Calidad ecológica (CAL_ECO).

La descripción de los factores humanos y de los factores geográficos qué valores les fueron asignados en éste estudio, se presentan a detalle las Tablas 6.1 y 6.2 respectivamente.

Tabla 6.2. Variables seleccionadas como factores humanos.

TIPO	CLAVE	NOMBRE	CATEGORIZACIÓN	CRITERIO
FACTORES HUMANOS	CAB_MUN	Distancia a Cabeceras Municipales	Distancia a 1 Kilómetro	Es MÁS probable que a menor distancia de un centro de población importante, Sí ocurra un PC
			Distancia a 2 Kilómetros	
			Distancia a 3 Kilómetros	
			Distancia a 4 Kilómetros	
			Distancia a 5 Kilómetros	
	DIST_CARR	Distancia a la carretera más cercana	Distancia a 1 Kilómetro	Es MÁS probable que a menor distancia de carretera, Sí ocurra un PC
			Distancia a 2 Kilómetros	
			Distancia a 3 Kilómetros	
			Distancia a 4 Kilómetros	
			Distancia a 5 Kilómetros	
			Distancia a 6 Kilómetros	
			Distancia a 7 Kilómetros	
			Distancia a 8 Kilómetros	
			Distancia a 9 Kilómetros	
			Distancia a 10 Kilómetros	
	ZC	Zonas de campamento	Dentro del área	Es MAS probable que SI ocurra un PC dentro de una ZC
			Fuera de área	
	NANP	Áreas Naturales Protegidas	Fuera de área	Es MÁS probable que Sí ocurra un PC fuera de una ANP o bajo algún tipo de protección.
			Dentro del área	
APROV	Aprovechamientos forestales	Dentro del área	Es MAS probable que SI ocurra un PC dentro de un área de APROV	
		Fuera de área		
LIN_COND	Líneas de conducción	Dentro del área	Es MAS probable que SI ocurra un PC dentro de un área de LN-COND	
		Fuera de área		

Tabla 6.3. Variables seleccionadas como factores geográficos.

TIPO	CLAVE	NOMBRE	REATEGORIZACION	COMENTARIOS
FACTORES GEOGRÁFICOS	US_SUELO	Vegetación susceptible a incendios forestales	Bosque templado (coníferas y encinos)	Es MÁS probable que Sí ocurra un PC tomando en cuenta la carga de combustible que cierto tipo de vegetación puede producir.
			Bosque mesófilo de montaña	
			Selva subhúmeda	
			Pastizal cultivado	
			Pastizal natural (semidesértico)	
			Matorrales de zonas áridas	
			Pradera de alta montaña	
			Plantaciones forestales	
			Vegetación halófila y gipsófila	
PEND_CAT	Pendientes categorizadas	Alta	Es MÁS probable que Sí ocurra un PC cuando la pendiente es mayor	
		Media		
		Baja		
ALT_CAT	Altitudes categorizadas	Menor de 1,000 m.s.n.m.	Es MÁS probable que Sí ocurra un PC cuando son altitudes bajas	
		De 1,001 a 2,000 m.s.n.m.		
		De 2,001 a 3,000 m.s.n.m.		
		Mayor de 3,000 m.s.n.m.		
REHUSUEL	Régimen de humedad del suelo	Árido	Es MÁS probable que Sí ocurra un PC cuando existe menor cantidad de humedad en el suelo	
		Xérico		
		Ústico		
		Údico		
		Ácuico		
TEMP_MED	Tempertatura media	Cálida	Es MÁS probable que Sí ocurra un PC cuando existe mayor temperatura en el ambiente	
		Semicálida		
		Templada		
		Semifría		
		Fría		
PREC_MED	Precipitación Media	De 400 a 600 mm	Es MÁS probable que Sí ocurra un PC cuando existe menor precipitación en la zona	
		De 600 a 800 mm		
		De 800 a 1200 mm		
		De 1200 a 1500 mm		
		De 1500 a 2000 mm		
CAL_ECO	Calidad ecológica	Muy Baja	Es MÁS probable que Sí ocurra un PC si la calidad ecológica es muy baja	
		Baja		
		Media		
		Alta		

Por otra parte, las variables de salida que se conforman por los puntos de calor obtenidos de CONABIO, fueron agrupados en dos partes: los datos del 2000 al 2005 se emplearon para llevar a cabo el entrenamiento de la red, y los datos de 2006 conformaron el conjunto para la aplicación de la mejor red seleccionada.

6.4. Definición de cuatro modelos para entrenar la RNA

Con el fin de poder evaluar la totalidad de las variables propuestas en el estudio, se determinaron cuatro combinaciones o modelos de variables de entrada:

- **MODELO 1**

El primer arreglo (MODELO 1) se conformó con las variables seleccionadas mediante el procedimiento GRNN, el cual determinó cuáles son las variables que más contribuyen a obtener una R^2 más cercana a 1.

- **MODELO 2**

El segundo arreglo (MODELO 2) se generó con la totalidad de las trece variables de entrada, incluyendo tanto los factores humanos (US_SUELO, DIST_POBL, DIST_CARR, ZC, ANP, APROV, LIN_COND) como los geográficos (PEND_CAT, ALT_CAT, REHUSUEL, TEMP_MED, PREC_MED, CAL_ECO).

- **MODELO 3**

El tercer arreglo (MODELO 3) incluyó solamente las siete variables de entrada consideradas como factores humanos (US_SUELO, DIST_POBL, DIST_CARR, ZC, ANP, APROV, LIN_COND).

- MODELO 4

Finalmente, el cuarto arreglo (MODELO 4) está conformado por las variables de entrada consideradas como factores geográficos (PEND_CAT, ALT_CAT, REHUSUEL, TEMP_MED, PREC_MED, CAL_ECO).

En Vega-García *et al.* (1996) se plantearon modelos similares para probarse en la red, realizando análisis exploratorios en los cuales las variables fueron introducidas a la red y para ser evaluadas usando como criterio principal obtener los más bajos valores en el cuadrado medio del error (*Root Mean Square Error* o RMSE) en el conjunto de datos de prueba (*test*).

En la Figura 6.3 se presenta el mapa con los puntos de calor diurnos y nocturnos detectados en los años 2000, 2001, 2002, 2003, 2004, 2005 y 2006 dentro de las áreas con vegetación susceptible a incendios forestales los cuales fueron usados como datos de salida.

Los mapas correspondientes a las variables de entrada fueron generados en el programa ArcView© con una plantilla de referencia UTM y su correspondiente simbología. Éstos se presentan en las Figuras 6.4 a 6.16.

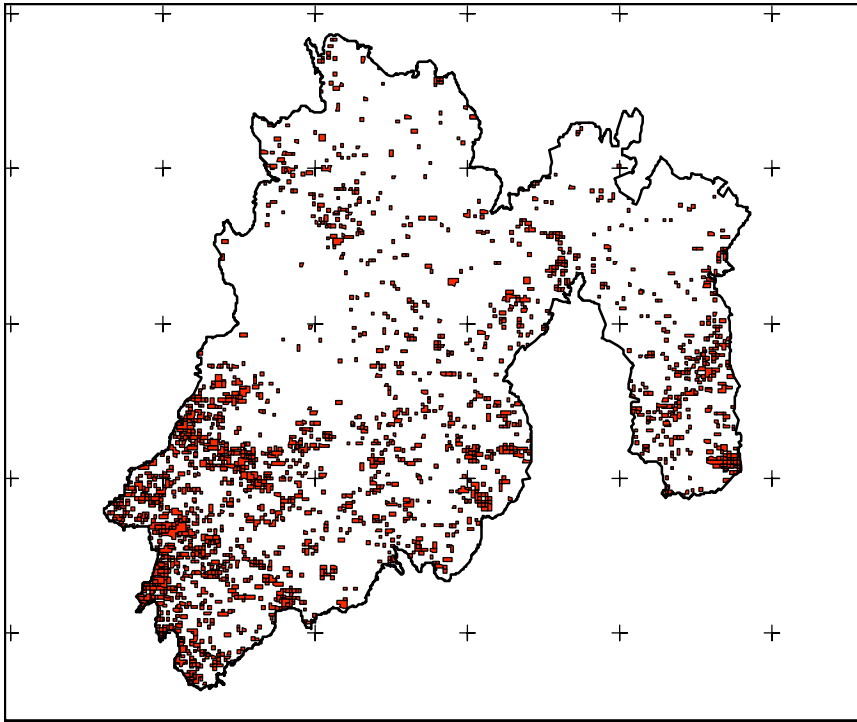


Figura 6.3. Puntos de calor diurnos y nocturnos del 2000 al 2006 en vegetación susceptible

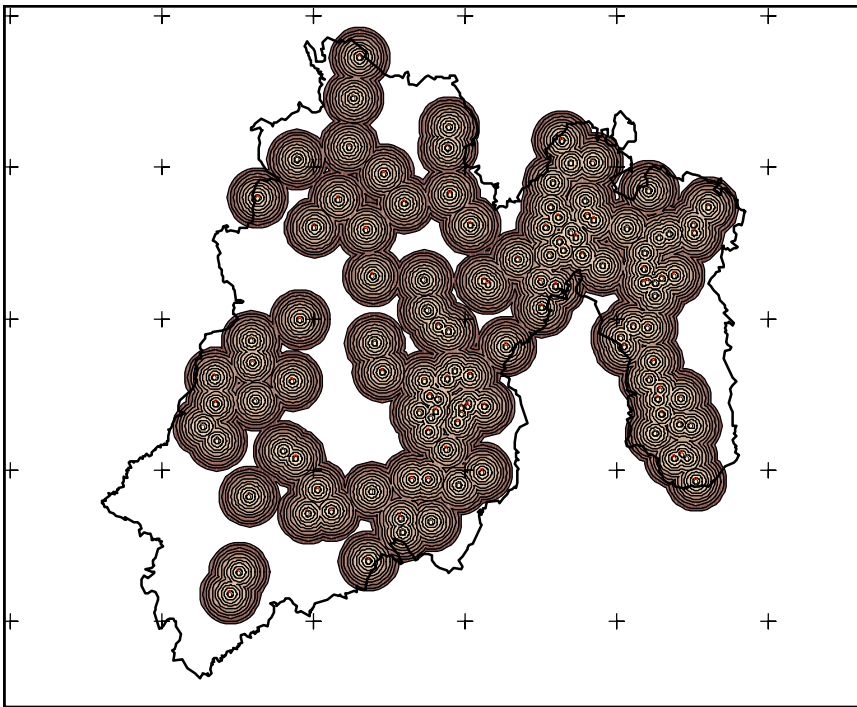


Figura 6.4. Mapa de cabeceras municipales (CAB_MUN)

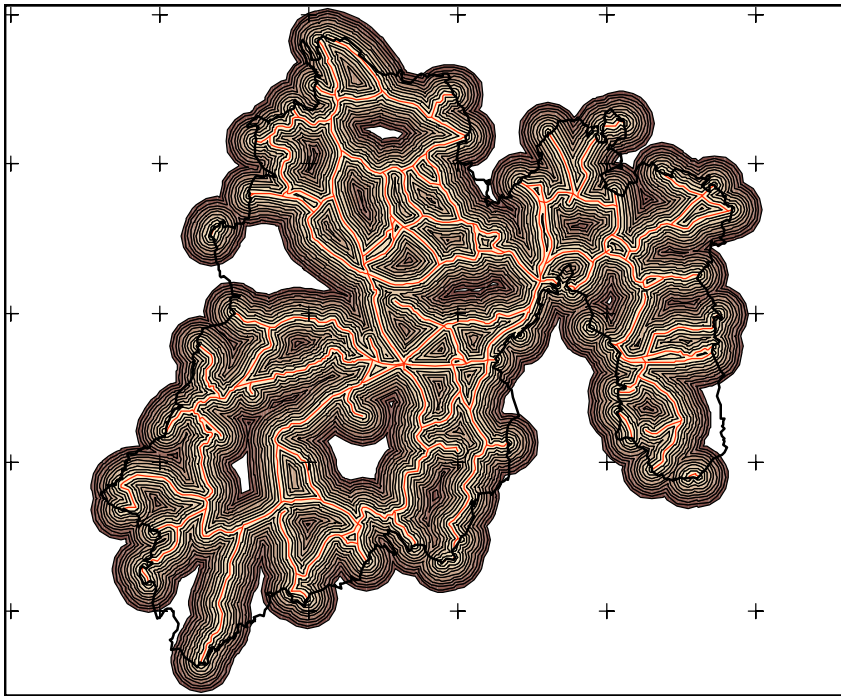


Figura 6.5. Mapa de distancia a la carretera principal más cercana (DIST_CARR)

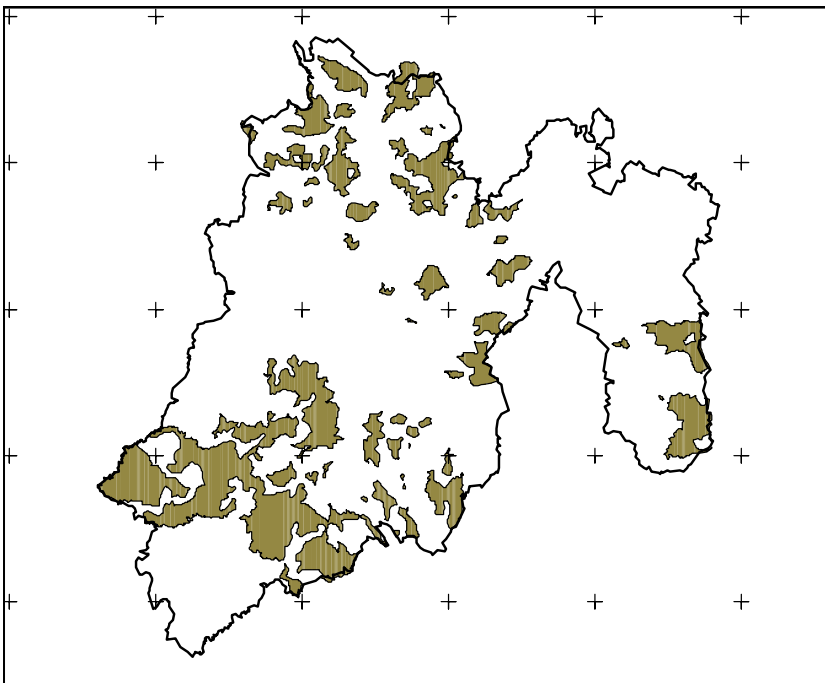


Figura 6.6. Mapa de zonas de campamento (ZC)

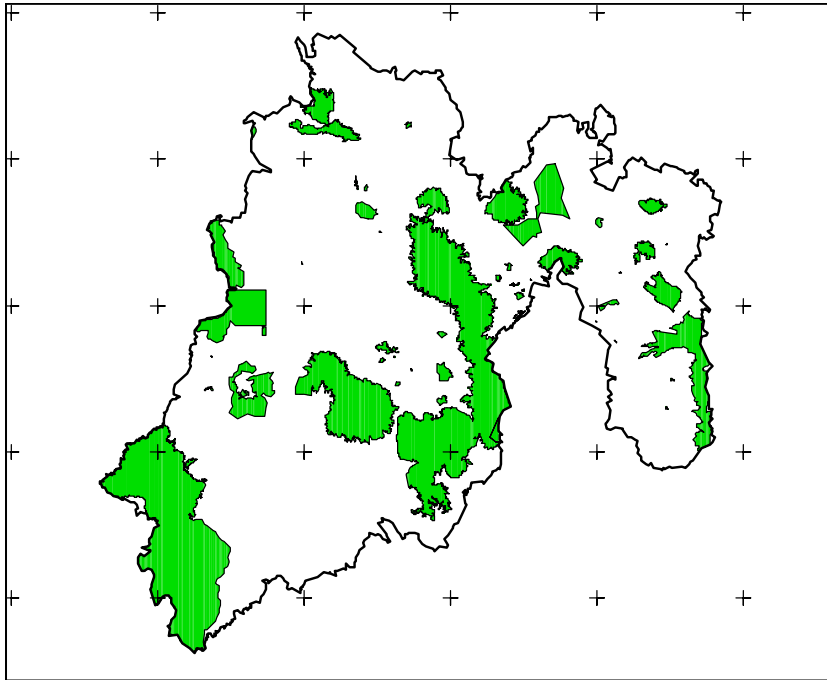


Figura 6.7. Mapa de áreas naturales protegidas (ANP)

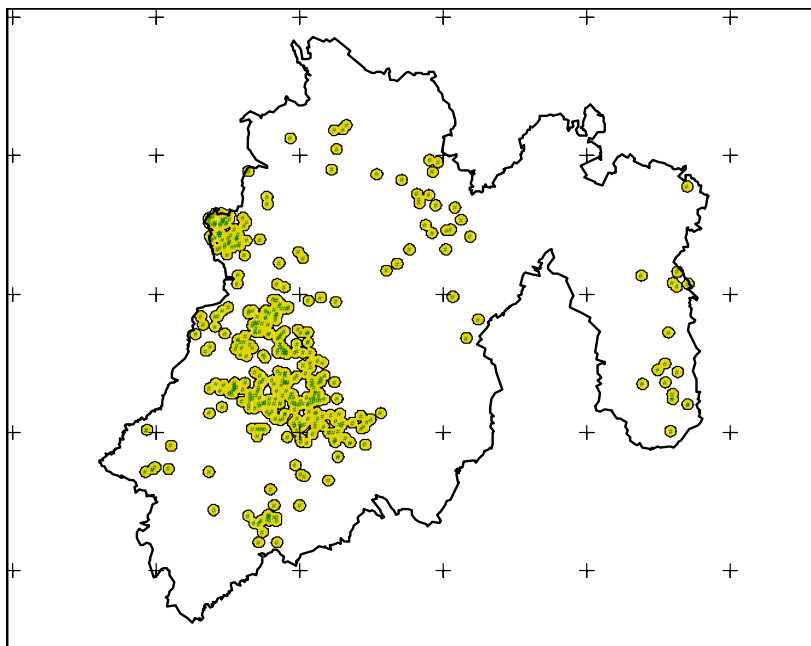


Figura 6.8. Mapa de aprovechamientos forestales autorizados (APROV)

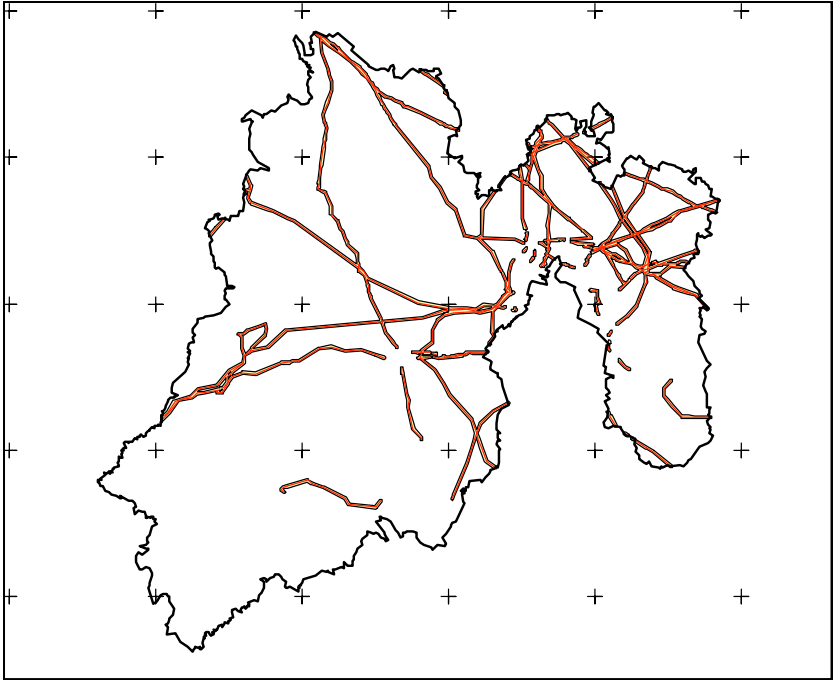


Figura 6.9. Mapa de líneas de conducción eléctrica (LIN_COND)

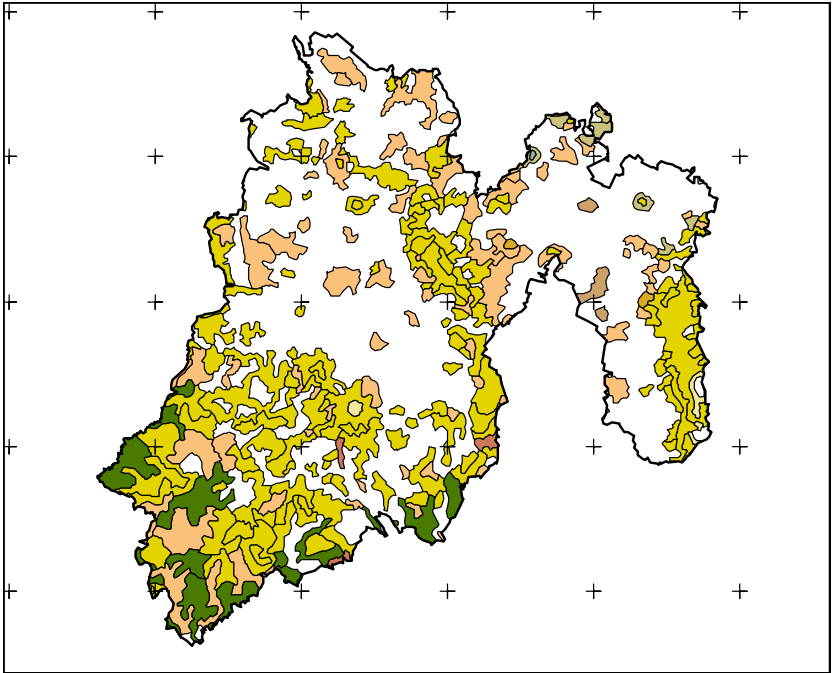


Figura 6.10. Mapa de vegetación susceptible a incendios forestales (US_SUELO)

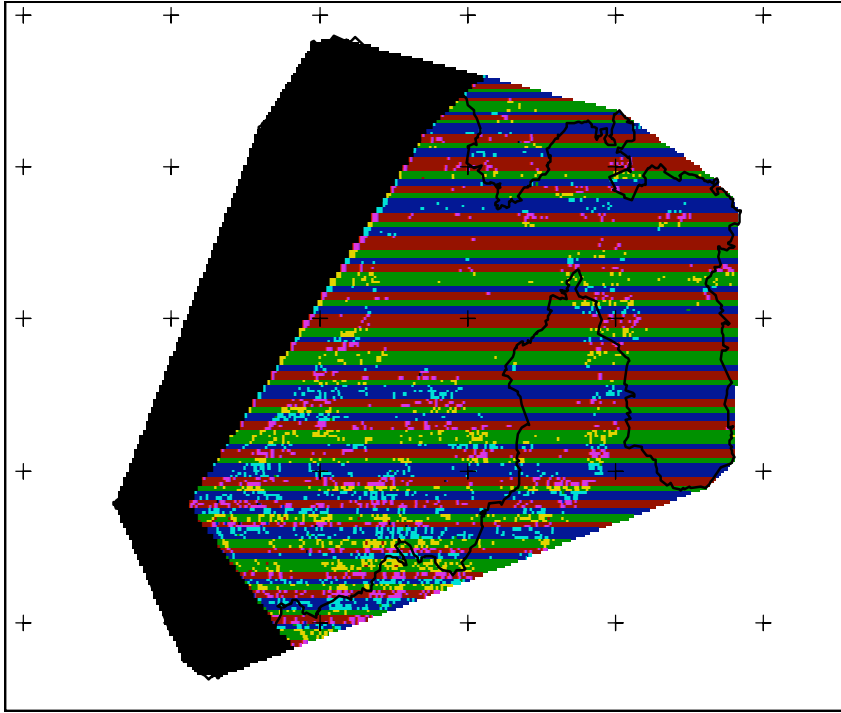


Figura 6.11. Mapa de pendientes categorizadas (PEND_CAT)

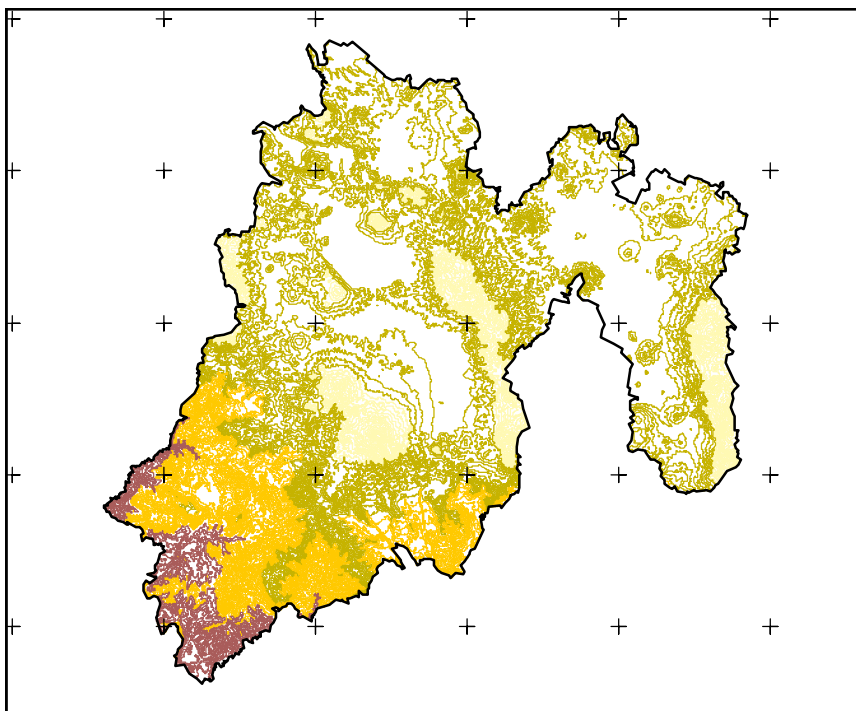


Figura 6.12. Mapa de altitudes categorizadas (ALT_CAT)

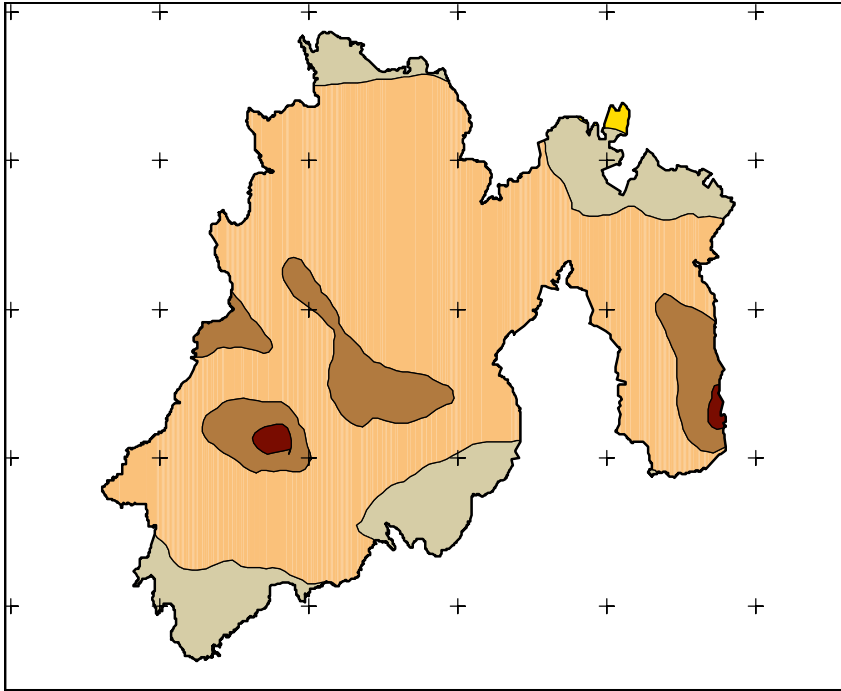


Figura 6.13. Mapa del régimen de humedad en el suelo (REHUSEUL)

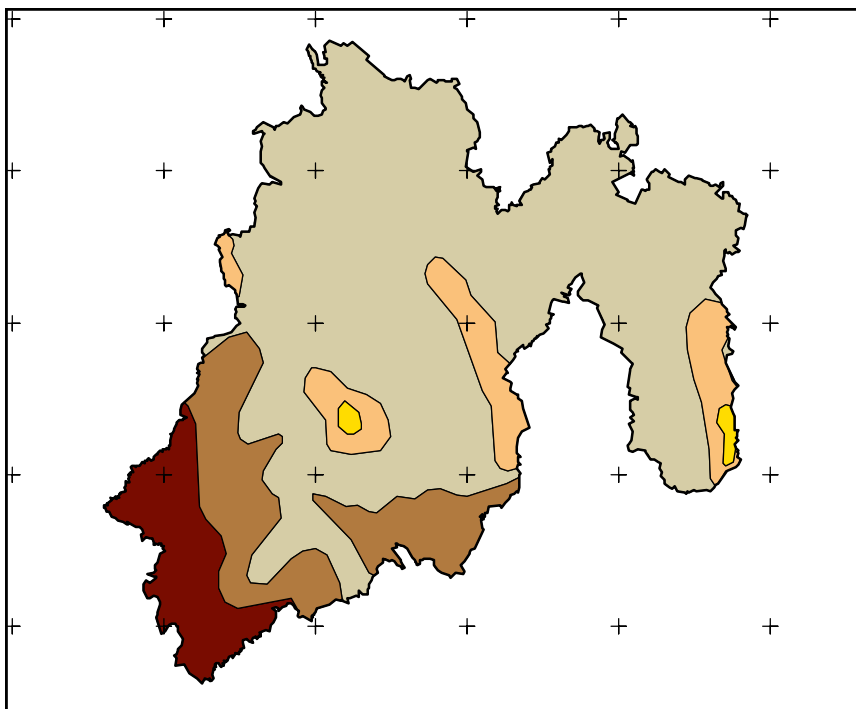


Figura 6.14. Mapa de temperatura media anual (TEM_MED)

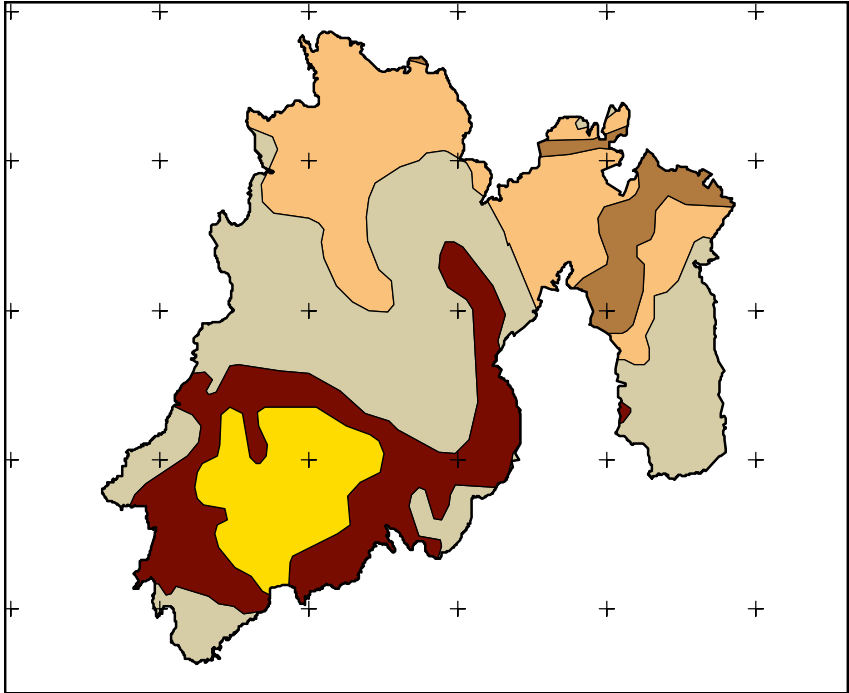


Figura 6.15. Mapa de precipitación media anual (PREC_MED)

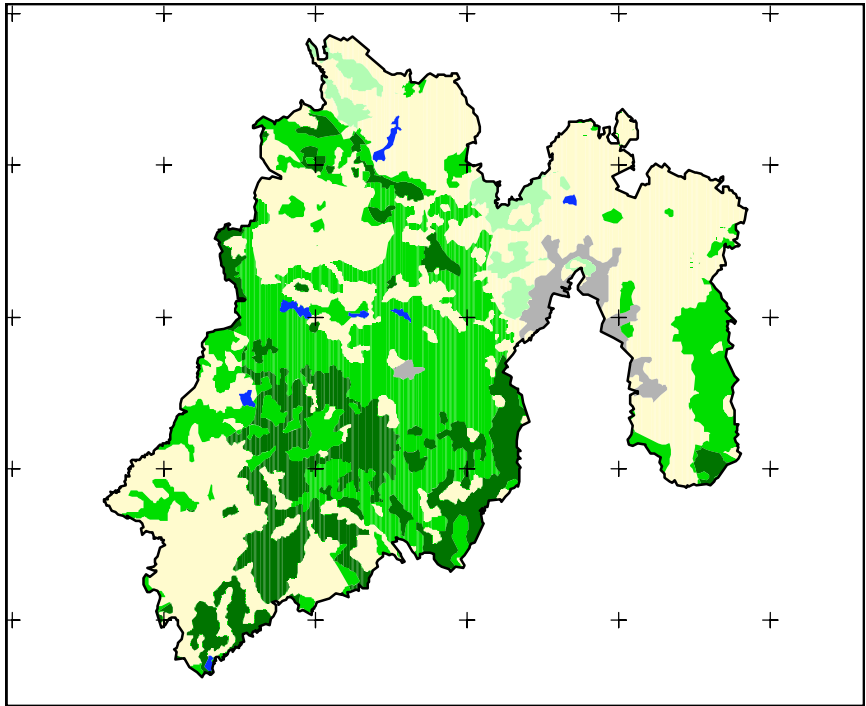


Figura 6.16. Mapa de calidad ecológica (CAL_ECO)

6.5. Alimentación de la base de datos

Fueron seleccionadas de la base de datos en formato de columna aquellas celdas que representan áreas de uso de suelo que son susceptibles a presentar puntos de calor en base a su cobertura vegetal. Esto permitió reducir el número de celdas que se introdujeron a la RNA de un total de 40,000 celdas que componen la imagen a 21,728 que forman el límite estatal del Estado de México, y posteriormente a 8,938 celdas, que representan sólo la vegetación susceptible a incendios. Una vez obtenidos los resultados de la RNA, estos datos se reintegraron a la base de datos, y posteriormente, se reconvirtieron en imágenes raster para su representación espacial. Esto es posible, debido a que cada celda está identificada en un número progresivo del 1 al 40,000. Cada celda en la base de datos le corresponde un valor por cada variable de entrada. Espacialmente, dichas características corresponden a lo existente en campo, lo cual es corroborado una vez que los resultados son obtenidos.

El procedimiento *backpropagation* requiere que para cada observación existirá un valor de salida al cual se desea llegar, mediante los cálculos que la red precise ajustando los pesos de cada neurona con el fin de obtener un menor valor en el Error Cuadrado Medio (ECM). Cada observación se asocia un valor correspondiente a la presencia o ausencia de un punto de calor.

Se emplearon como datos de salida los valores de presencia o no de puntos de calor de los años 2000 al 2005 representados en una escala del 0 al 4, el cual representa las veces en que un punto de calor se ha presentado en una determinada celda en este período de tiempo (Tabla 6.3).

Tabla 6.4. Valores de salida que representan incidencia de puntos de calor.

NO DE INCIDENCIAS DE PUNTOS DE CALOR EN LA CELDA	VALOR	SALIDA
Ninguna ocurrencia en los 6 años	0	0
Una sola ocurrencia en los 6 años	1	0 - 0.25
Dos ocurrencias en los 6 años	2	0.25 - 0.50
Tres ocurrencias en los 6 años	3	0.50 - 0.75
Cuatro ocurrencias en los 6 años	4	0.75 - 1

De acuerdo a lo anterior, los valores resultantes de la red representan mayor o menor probabilidad de que se presenten uno o varios eventos que involucran la presencia de puntos de calor. Los resultados una vez obtenidos en la salida de la red son categorizados, de manera que magnitudes de un intervalo de entre 0 a 1 correspondan al número de veces en que se han presentado puntos de calor (Figura 6.17) Ninguna celda presentó más de cuatro puntos de calor en los seis años evaluados.

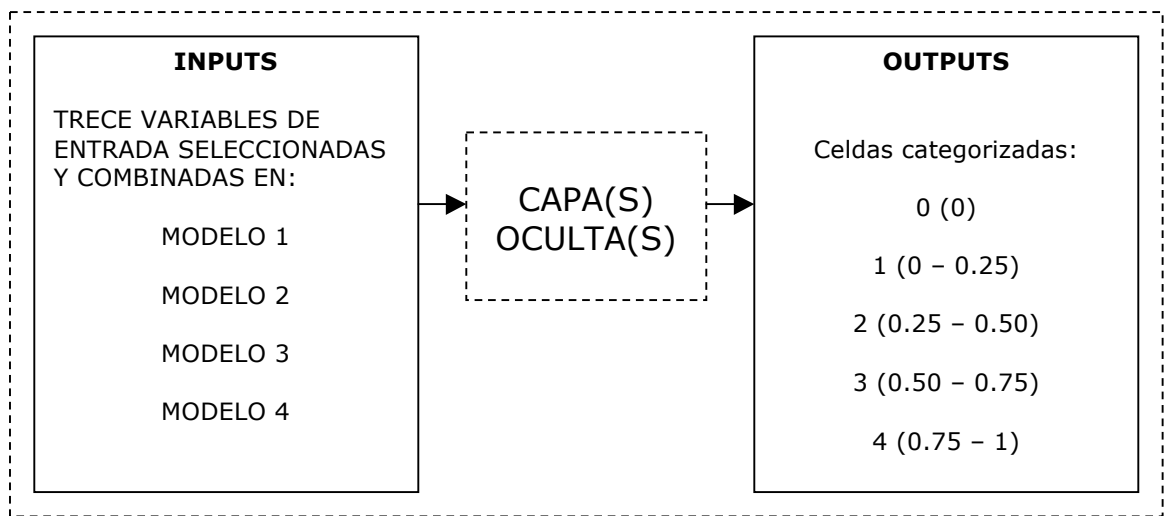


Figura 6.17. Datos que intervienen en la alimentación de la red

6.6. Definición de la muestra para el entrenamiento.

Una vez que se obtuvieron las 8,938 celdas correspondientes a las zonas con vegetación susceptible a presentar puntos de calor en el Estado de México, se observó que existía un gran número de observaciones de no presencia de punto de calor, en comparación con aquellas celdas que presentaban por lo menos la presencia de un solo punto de calor. Al efectuarse las pruebas con estos datos, se observó niveles en el error cuadrático medio superiores al 0.5 que en ocasiones podían oscilar hasta el 0.8, en función de la arquitectura y de las funciones de activación empleadas. Asimismo, se requería de un mayor número de épocas para alcanzar el nivel mínimo del error. Dado lo anterior, se pensó que tal inequivalencia entre las observaciones influiría en el entrenamiento de las redes, provocando que el error no disminuyera y haciendo que la red se volviese ineficiente para procesar todos los datos del estudio.

Con base a la metodología empleada en el estudio realizado en *Vega-García et al. (1996)*, se determinó emplear una muestra con el mismo número de observaciones de celdas con presencia de puntos de calor y el mismo número de celdas con ausencia de estos puntos. Para dicha selección, se utilizaron las observaciones con presencia de por lo menos un punto de calor correspondientes al período 2000-2005, tomando un número equivalente seleccionado al azar en la base de datos del mismo año, obteniéndose así una muestra de 2,800 observaciones. Esta muestra fue utilizada para determinar los modelos más adecuados para el procesamiento de las 8,938 celdas (Figura 6.18).

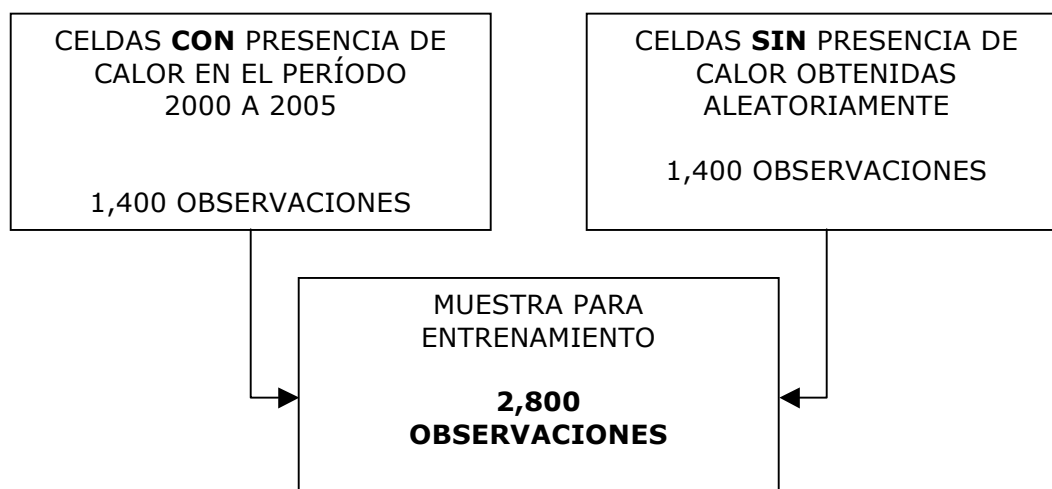


Figura 6.18. Conformación de la muestra de entrenamiento.

Se probaron combinaciones de parámetros en una misma red, con el fin de conocer su funcionamiento y comparar su desempeño ante diferentes funciones de activación (FA), número de neuronas en la(s) capa(s) oculta(s), tasas de aprendizaje (LR), momento (Mo), y pesos iniciales (InW). Otro factor evaluado fue la secuencia de entrada de datos (SeD), y se aplicaron diversos niveles de calibración (Ca), con el fin de evitar el sobreentrenamiento de la red. Para su identificación en la base de datos del proceso en la RNA, se designó con letra minúscula para cada tipo de arquitectura *backpropagation* aplicada (Tabla 6.4).

Tabla 6.5. Identificación por tipo de arquitectura *backpropagation*.

ID	TIPO DE ARQUITECTURA	CARACTERÍSTICAS
" a "	Lineal	Tres capas
" b "	Lineal	Cuatro capas
" c "	Lineal	Cinco capas
" d "	Recurrente (Jordan-Elman)	Con retroalimentación en la capa de entrada
" e "	Recurrente (Jordan-Elman)	Con retroalimentación en la capa oculta
" f "	Recurrente (Jordan-Elman)	Con retroalimentación en la capa de salida
" g "	Múltiple (Ward)	Dos capas con diferente función de activación
" h "	Múltiple (Ward)	Tres capas con diferente función de activación
" i "	Múltiple (Ward)	Dos capas con jumper activado

7. RESULTADOS

7.1. Determinación de las variables para el MODELO 1 con GRNN

Se determinaron para el MODELO 1 las mejores variables mediante el procedimiento GRNN (*General Regression Neural Network*) de las trece variables propuestas. Para ello se utilizaron con las 2,800 observaciones del conjunto de datos de la muestra como datos de entrenamiento aplicándose en ellas cinco funciones de activación: (0,1), (-1,1), Logística, Tanh y None. De la misma manera que se determina en las redes *backpropagation*, la aplicación de diferentes funciones de activación en la capa de entrada producen diferentes niveles desempeño de la red, mismo que es evaluado con R^2 , produciendo los resultados presentados en la Tabla 7.1:

Tabla 7.1. Resultados del proceso GRNN para la determinación del MODELO 1.

F.A. [0,1]			F. A. [-1,1]		
VARIABLE	FACTOR DE ATENUAMIENTO	TIPO DE VARIABLE	VARIABLE	FACTOR DE ATENUAMIENTO	TIPO DE VARIABLE
DIST_CARR	3	H	DIST_CARR	2.98824	H
ALT_CAT	2.87059	G	ALT_CAT	2.90588	G
APROV	2.14118	H	APROV	2.14118	H
PREC_MED	1.87059	G	PREC_MED	2	G
REHUSUEL	1.34118	G	REHUSUEL	1.4	G
TEMP_MED	1.09412	G	TEMP_MED	0.88235	G
US_SUELO	1.02353	G	NCAL_ECOL	0.83529	G
NCAL_ECOL	0.84706	G	PEND_CAT	0.69412	G
PEND_CAT	0.76471	G	ANP	0.62353	H
ANP	0.64706	H	CAB_MUN	0.16471	H
ZC	0.17647	H	ZC	0.11765	H
CAB_MUN	0.11765	H	US_SUELO	0.10588	G
LIN_COND	0	H	LIN_COND	0.04706	H

H = Variable humana G = Variable geográfica

Continuación.

F. A. LOGISTIC		
VARIABLE	FACTOR DE ATENUAMIENTO	TIPO DE VARIABLE
ALT_CAT	2.82353	G
REHUSUEL	2.8	G
US_SUELO09	2.70588	G
NCAL_ECOL	2.51765	G
DIST_CARR	2.04706	H
LIN_COND	1.8	H
PREC_MED	1.76471	G
ANP	1.49412	H
APROV	1.49412	H
PEND_CAT	1.34118	G
TEMP_MED	0.83529	G
ZC	0.65882	H
CAB_MUN	0.27059	H

F. A. NONE		
VARIABLE	FACTOR DE ATENUAMIENTO	TIPO DE VARIABLE
PREC_MED	2.98824	G
ALT_CAT	2.90588	G
REHUSUEL	2.77647	G
APROV	2.76471	H
LIN_COND	1.78824	H
TEMP_MED	1.55294	G
ANP	1.44706	H
DIST_CARR	1.09412	H
NCAL_ECOL	0.98824	G
US_SUELO	0.82353	G
PEND_CAT	0.57647	G
CAB_MUN	0.49412	H
ZC	0.43529	H

F. A. TANH		
VARIABLE	FACTOR DE ATENUAMIENTO	TIPO DE VARIABLE
ALT_CAT	2.88235	G
DIST_CARR	2.58824	H
PREC_MED	2.55294	G
REHUSUEL	2.23529	G
US_SUELO	1.30588	G
PEND_CAT	1.11765	G
LIN_COND	1.05882	H
NCAL_ECOL	0.94118	G
ANP	0.69412	H
CAB_MUN	0.37647	H
ZC	0.07059	H
APROV	0.03529	H
TEMP_MED	0.02353	G

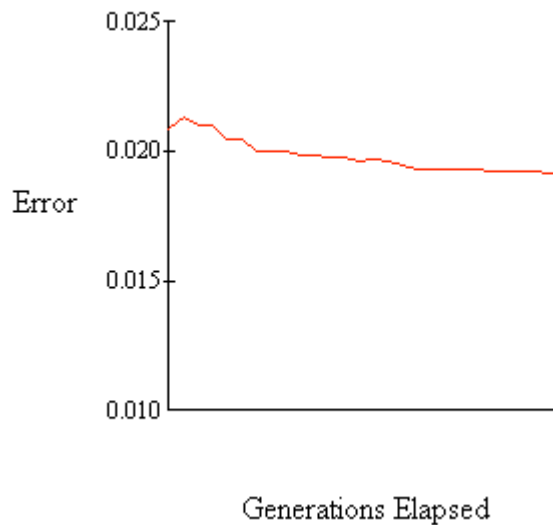
CUADRO COMPARATIVO

F.A.	R2	r2	Ccr
[1,0]	0.3749	0.3884	0.6232
[-1,1]	0.3902	0.4015	0.6337
TANH	0.3114	0.3230	0.5684
LOGISTIC	0.3253	0.3413	0.5842
NONE	0.3166	0.3258	0.5708

H = Variable humana G = Variable geográfica

Se determinó que las funciones de activación [-1,1] y [0,1] son las que presentan el mayor nivel de R^2 en comparación con las restantes, pues con ellas se obtuvieron 0.3902 y 0.3749 respectivamente, siendo la primera la que se utilizó para determinar las cinco variables del MODELO 1: DIST_CARR, ALT_CAT, APROV, PREC_MED Y REHUSUEL, bajo el criterio de que fueron estas las que presentaron individualmente un

factor de atenuamiento mayor a 1, lo que significa que son las variables que se encuentran más correlacionadas con los valores de salida. Este modelo representó la mejor combinación de variables geográficas con las variables humanas del estudio. El error mínimo promedio alcanzado fue menor a 0.02135 como se muestra en la Gráfica 7.1.



Gráfica 7.1. Error mínimo promedio con FA [-1,1], obtenido en el procedimiento GRNN para MODELO 1 en la muestra de 2,800 observaciones.

7.2. Aplicación de la muestra

Una vez determinados los cuatro modelos del estudio, se realizó la evaluación de la muestra de 2,800 observaciones (las 1,400 observaciones con incendio y 1,400 observaciones sin incendio escogidas de forma aleatoria), realizando modificaciones en los parámetros que de acuerdo a los ejercicios de prueba y error resultaron ser las que permitieron obtener una R^2 más alta, como fueron variables en el modelo, arquitectura empleada, funciones de escalamiento y de activación y modo de actualización de los

pesos, lo que permitió perfeccionar la arquitectura dada para posteriormente realizar su aplicación en la red final de 8,938 observaciones.

Se determinaron como constantes en todas las redes el modo de actualización de pesos denominada TURBOPROP, así como una FA LOGISTICA en todas las capas de salida y evaluadas en un máximo de 1,000 épocas de entrenamiento. La determinación de esta función en la capa de salida, fue porque en todas las pruebas realizadas las redes *backpropagation* con esta capa obtuvieron mejores resultados de R^2 que aquellas que emplearon una función distinta (TANH, GAUSSIANA O SENO). Asimismo se consideró que este número de épocas eran suficientes para alcanzar el error mínimo promedio, ya que la mayoría de las redes lo lograban entre las 400 y 500 épocas de entrenamiento, y ninguna redujo su error mínimo promedio después de las 1,000 épocas.

Asimismo, tanto para los datos de muestra como para los datos-objetivo, la modalidad TURBOPROP ofreció un aprendizaje más rápido y más preciso que las dos modalidades de actualización de pesos VANILLA y MOMENTUM.

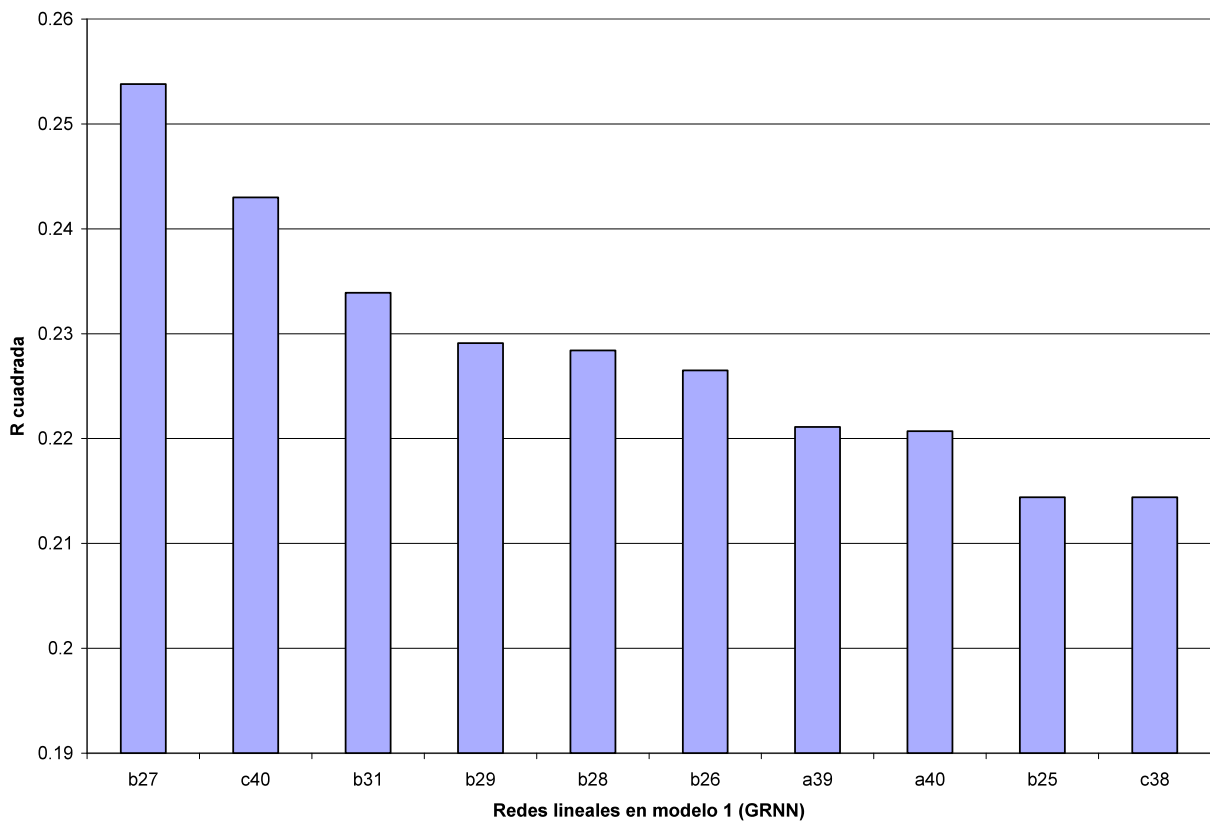
Se presentan en las siguientes gráficas las diez mejores redes para cada modelo en cada arquitectura evaluada.

7.2.1. Redes lineales

MODELO 1 (Variables seleccionadas del modelo GRNN)

El modelo que consideró las variables DIST_CARR, ALT_CAT, APROV, PREC_MED Y REHUSUEL no presentó un comportamiento óptimo, ya que sólo se obtuvieron valores

de R^2 que oscilaron entre 0.1836 a 0.2538 con esta arquitectura. La mejor red (“b27”) se obtuvo al emplear una arquitectura de cuatro capas, con FA función de escalamiento de [-1,1] y funciones GUASIANA y TANH en las capas ocultas. Se observó que las funciones de escalamiento <<0,1>> y [0,1] presentan resultados idénticos bajo las mismas condiciones. Por lo anterior, fue posible determinar que por lo menos para esta arquitectura, basta el empleo de una sola de ellas para evaluar redes con similares FA (Gráfica 7.2).

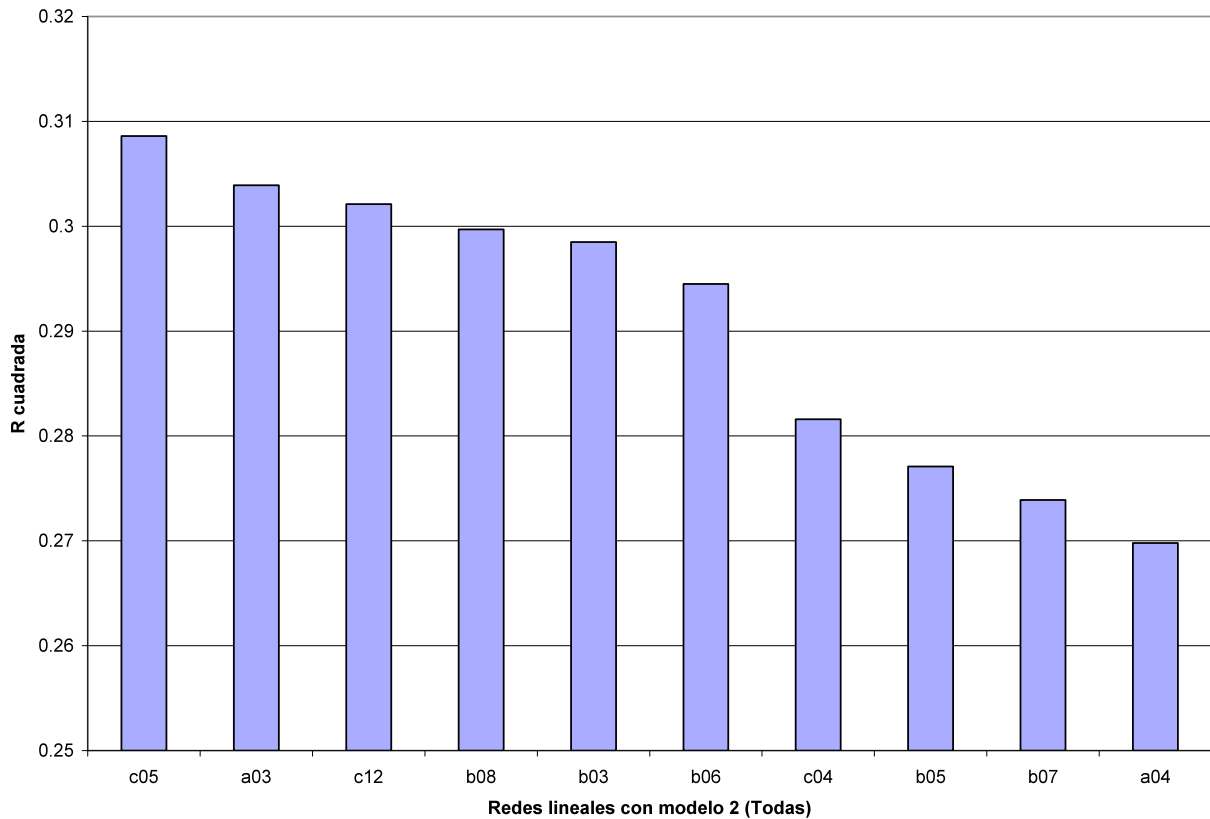


Gráfica 7.2. Valores de R^2 obtenidos para MODELO 1 en arquitecturas de redes lineales.

MODELO 2 (Todas las variables seleccionadas)

El modelo c05 diseñado con las trece variables propuestas para el estudio tampoco ofreció resultados aceptables en las arquitecturas de redes lineales. Se esperaba que

los resultados fueran similares a los obtenidos por Buendía (2000), quien obtuvo una R^2 cercana al 0.9 empleando una red lineal de cuatro capas con FA logística. Sin embargo, solo se obtuvieron valores de R^2 que oscilaron enormemente de entre 0.1701 al 0.3086, siendo la mejor red una de cinco capas, con una función de escalamiento (FE) de tipo LOGISTICA y una combinación de funciones de activación (FA) en capas ocultas de GAUSIANA, TANH y LOGISTICA (Gráfica 7.3).

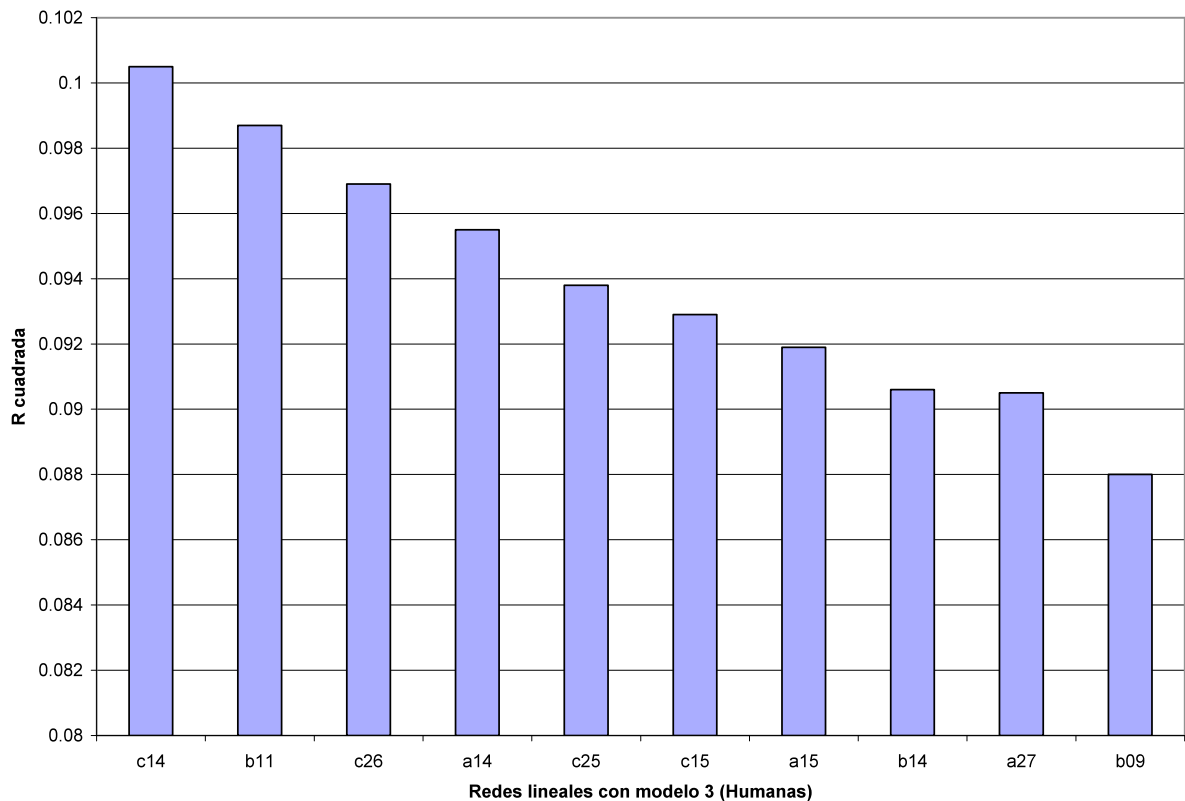


Gráfica 7.3. Valores de R^2 obtenida para MODELO 2 en arquitecturas de redes lineales.

Resultados similares se obtuvieron al emplear FE de $[-1,1]$ y TANH, sin embargo, no hay en este modelo una clara diferenciación de una arquitectura sobresaliente.

MODELO 3 (Variables humanas)

Al igual que los resultados obtenidos en los MODELOS 1 y 2, se obtuvieron valores de R^2 que no superaron el valor de 0.1005, mismos que fueron bajos e inconsistentes. En esta combinación, la red “c14” resultó mejor, con la utilización de arquitecturas de cinco capas, usando una FE de [0,1], una combinación de FA de tipo TANH, LOGISTICA y LOGISTICA y una salida LOGISTICA (Gráfica 7.4).

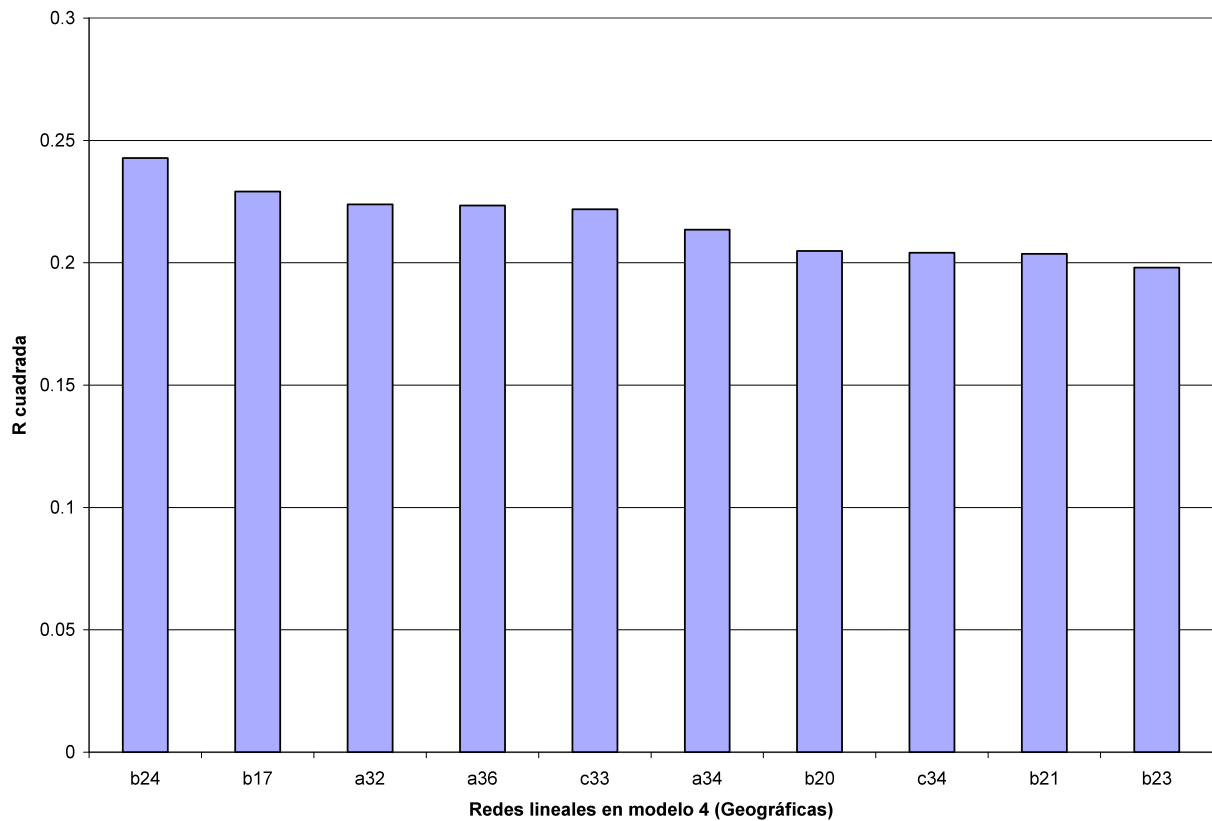


Gráfica 7.4. Valores de R^2 obtenidos para MODELO 3 en arquitecturas de redes lineales.

MODELO 4 (Variables geográficas)

Resultados similares al modelo anterior se obtuvieron al evaluar el MODELO 4 en redes lineales, ya que fueron obtenidos empleando una arquitectura de cuatro capas con una FE TANH en la capa de entrada y una combinación de FA GAUSIANA, TANH y LOGISTICA en las capas ocultas. Los resultados no fueron muy variables, pues no

superaron la R^2 de 0.2428. En general para este modelo y esta arquitectura, se observó menor variabilidad, en comparación con el MODELO 3 (Gráfica 7.5).

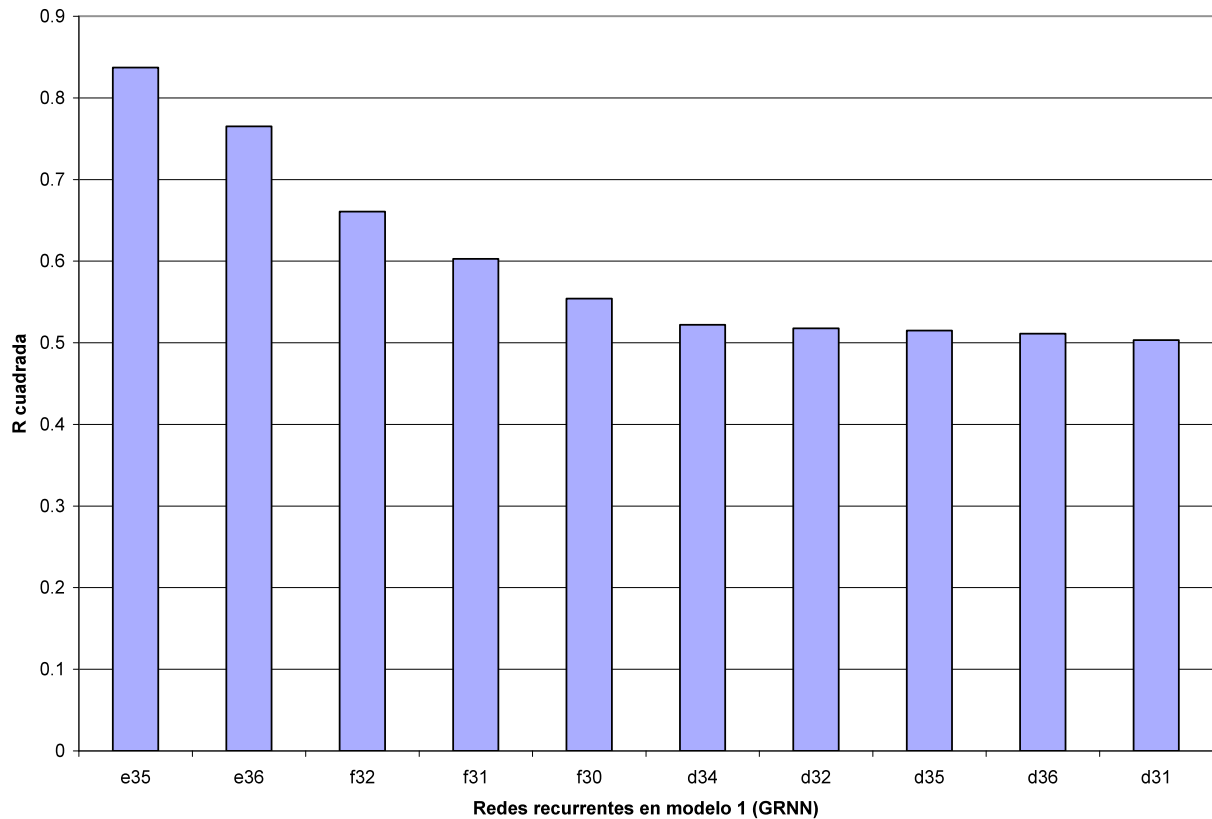


Gráfica 7.5. Valores de R^2 obtenidos para MODELO 4 en arquitecturas de redes lineales.

7.2.2. Redes recurrentes (Jordan-Elman Nets)

MODELO 1 (Variables seleccionadas del modelo GRNN)

El segundo mejor resultado para las observaciones de la muestra fueron obtenidos con la arquitectura de redes recurrentes Jordan-Elman en este modelo, tal como se muestra en la figura 7.6, al emplearse una red con capa oculta retroalimentada de FA LOGISTICA y con una FE de [-1,1] en la capa de entrada. Esta red proporcionó un valor de R^2 de 0.8518 para éste modelo, así como una r^2 de 0.8372 y una Ccr de 0.9159.

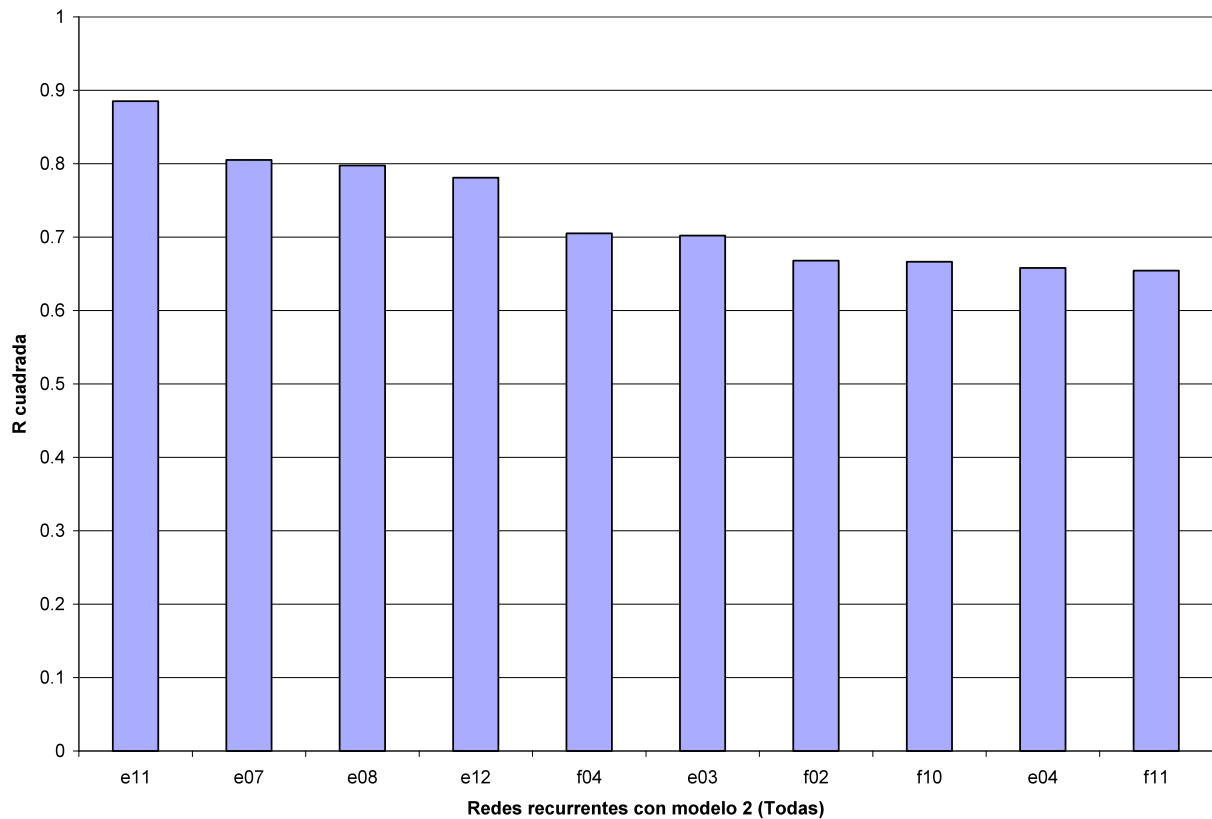


Gráfica 7.6. Valores de R^2 obtenidos para MODELO 1 en arquitecturas de redes recurrentes.

En general, las redes que emplearon FE de $[-1,1]$ y TANH tuvieron los mejores resultados. Por otra parte, valores similares de R^2 fueron obtenidos empleando la misma arquitectura con diferentes capas retroalimentadas (entrada, oculta y salida). Solo hubo diferencia significativa en este modelo al aplicar una FE TANH en la que se obtuvo 0.0946 (Gráfica 7.6).

MODELO 2 (Todas las variables)

El modelo que incluyó las trece variables presentó en esta arquitectura el mejor comportamiento de red en la muestra al obtenerse el valor superior de R^2 de todas que fue de 0.8852, una r^2 de 0.8862 y una Ccr de 0.9414. Este modelo y ésta arquitectura permitieron como valor mínimo una R^2 de 0.3240 (Gráfica 7.7)



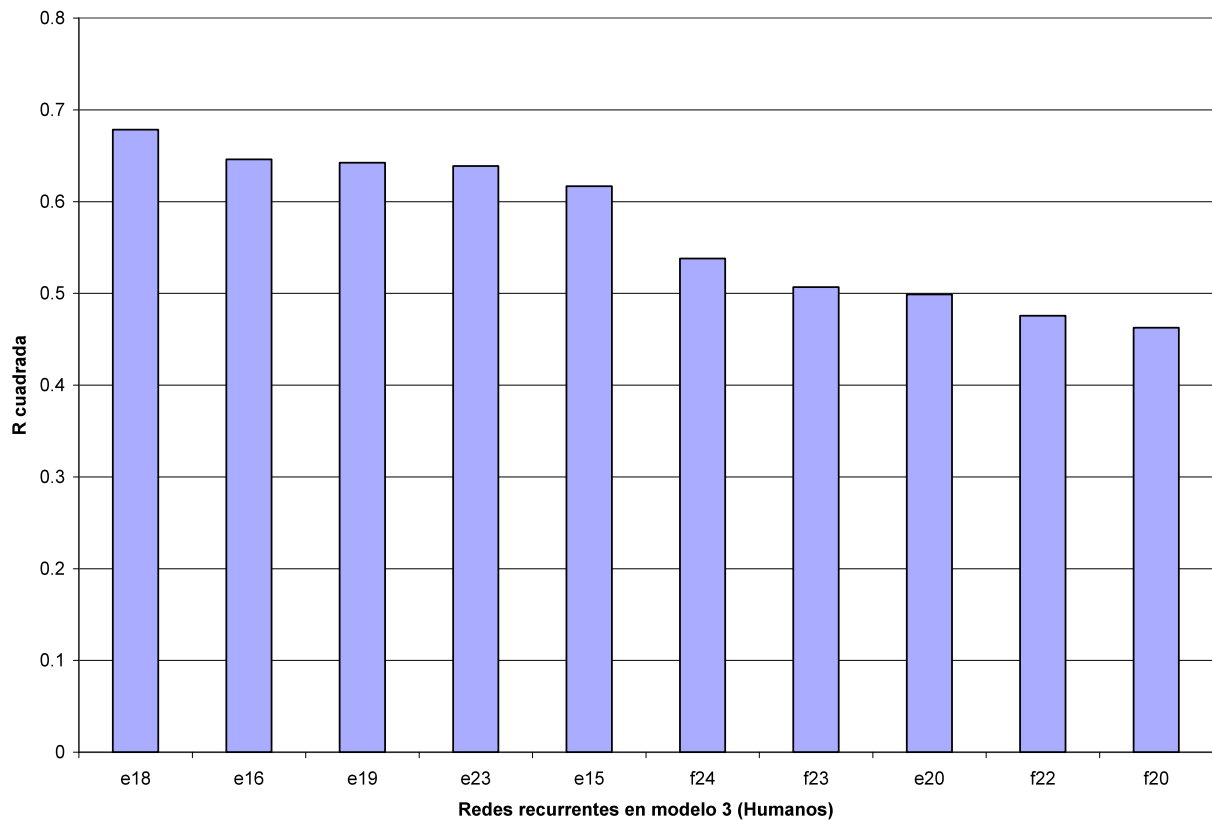
Gráfica 7.7. Valores de R^2 obtenidos para MODELO 2 en arquitecturas de redes recurrentes.

En la gráfica es posible observar que existe variabilidad de la R^2 en las diez mejores redes de este modelo, siendo las arquitecturas con retroalimentación en la capa de entrada y oculta las de mejor comportamiento. Las capas ocultas con FA GAUSIANA y TANH, presentaron los mejores resultados, siendo el factor principal la FE empleada, que en lo general correspondió a $[-1,1]$ y TANH.

MODELO 3 (Variables humanas)

Al igual que con el MODELO 3, se obtuvieron buenos resultados al utilizar una arquitectura con una FE de $[-1,1]$ y una FA GAUSIANA en la capa oculta. Los más bajos

valores de R^2 para este modelo fueron obtenidos al usarse una FE LOGISTICA tanto en redes retroalimentadas en las capas de entrada, oculta y de salida (Gráfica 7.8).



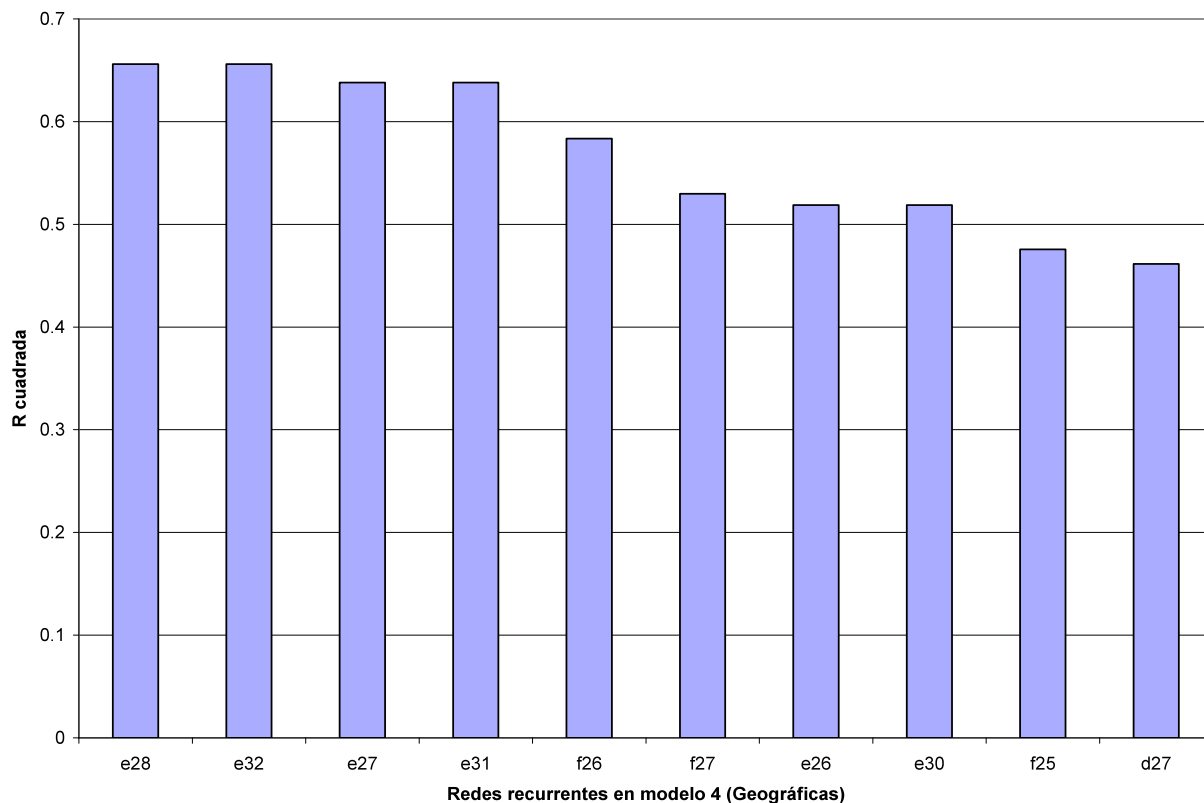
Gráfica 7.8. Valores de R^2 obtenidos para MODELO 3 en arquitecturas de redes recurrentes.

La gráfica muestra gran variabilidad en esta combinación, siendo solo la red anteriormente descrita como la tercera mejor red en general pero la mejor red para el MODELO 3.

MODELO 4 (Variables geográficas)

La utilización del MODELO 4 en esta arquitectura permitió la obtención de una R^2 no mayor de 0.6560 con una FE de TANH y una FA de tipo GAUSIANA en la capa oculta. Como se observa en la Figura 7.9 los valores de R^2 oscilaron hasta 0.4614 como

mínimo, pero no alcanzaron el desempeño de los MODELOS 1 y 2, aunque es importante ver que tuvieron un buen comportamiento que a pesar de ser variables de un solo tipo (Figura 7.9).



Gráfica 7.9. Valores de R^2 obtenidos para MODELO 4 en arquitecturas de redes recurrentes.

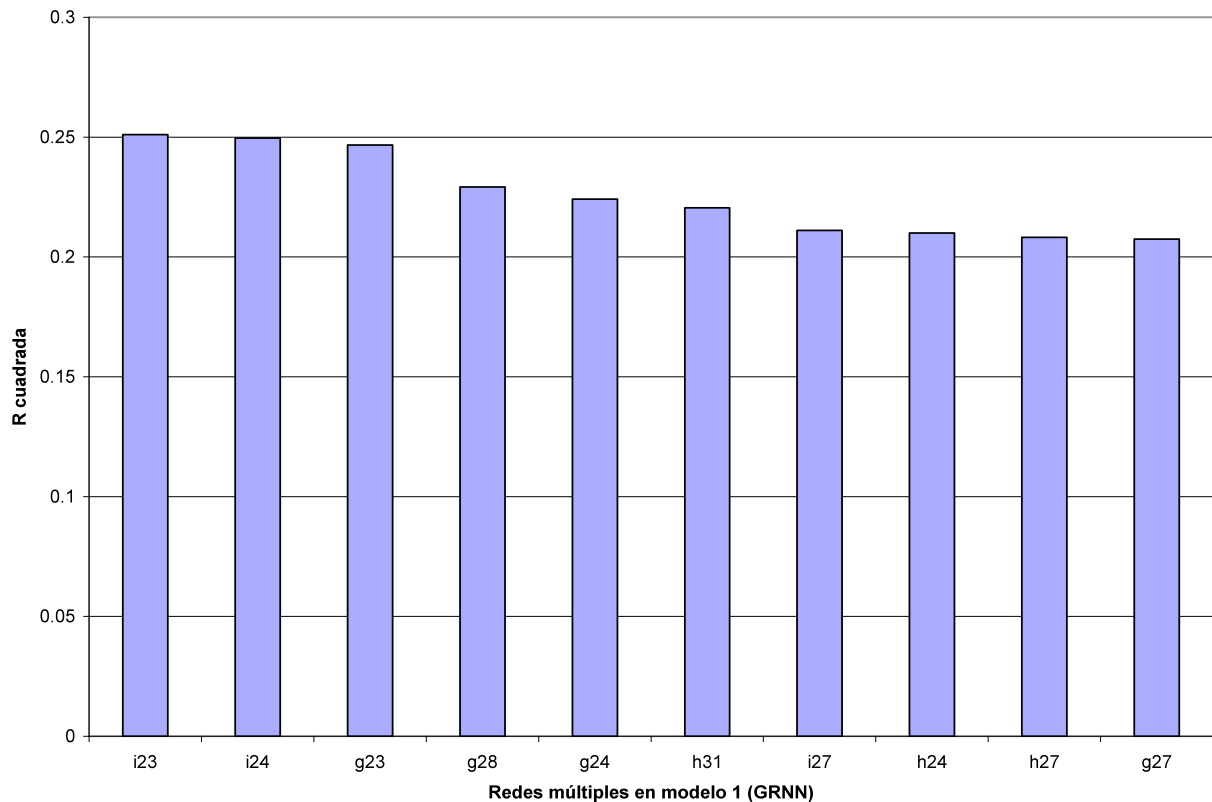
Esta tipo de red podría constituir una buena opción en estudios que impliquen solamente factores similares de tipo geográfico.

7.2.3. Redes múltiples (Ward nets)

MODELO 1 (Variables seleccionadas del modelo GRNN)

Las redes múltiples no ofrecieron mejores resultados que las redes lineales y recurrentes. En general, para el MODELO 1 se obtuvieron valores de R^2 que oscilaron de 0.2075 al 0.2511, y que como mínimo alcanzaron solo 0.1139. El mejor resultado se

logró aplicando una red con dos capas ocultas, FA GAUSIANA y TANH respectivamente, así como una FE de [0,1] en la capa de entrada (Gráfica 7.10).



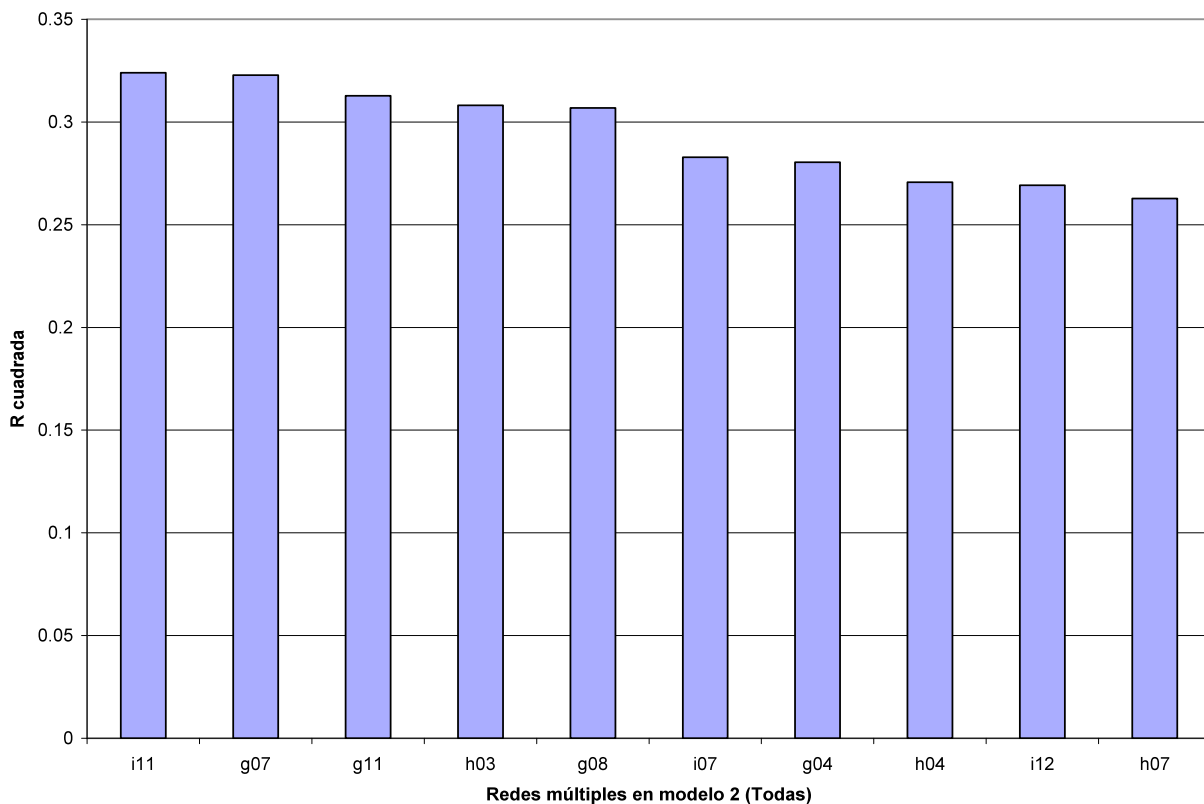
Gráfica 7.10. Valores de R^2 obtenidos para MODELO 1 en arquitecturas de redes múltiples.

Dichas redes no aportaron resultados aceptables. Las variables mixtas en redes múltiples no ofrecen resultados tan buenos por lo menos para la red como en el caso de las redes recurrentes (tipo Jordan-Elman).

MODELO 2 (Todas las variables)

Comportamiento similar al MODELO 1 se observó al aplicarse las trece variables en esta arquitectura. Los valores de R^2 para este modelo oscilaron entre 0.3243 a 0.1693. La mejor FE en la capa de entrada fue la [-1,1] y en menor proporción la FE de tipo

TANH. Para este modelo, no se observó que el tipo de FA en las capas ocultas tuviera trascendencia en el valor de R^2 (Gráfica 7.11).

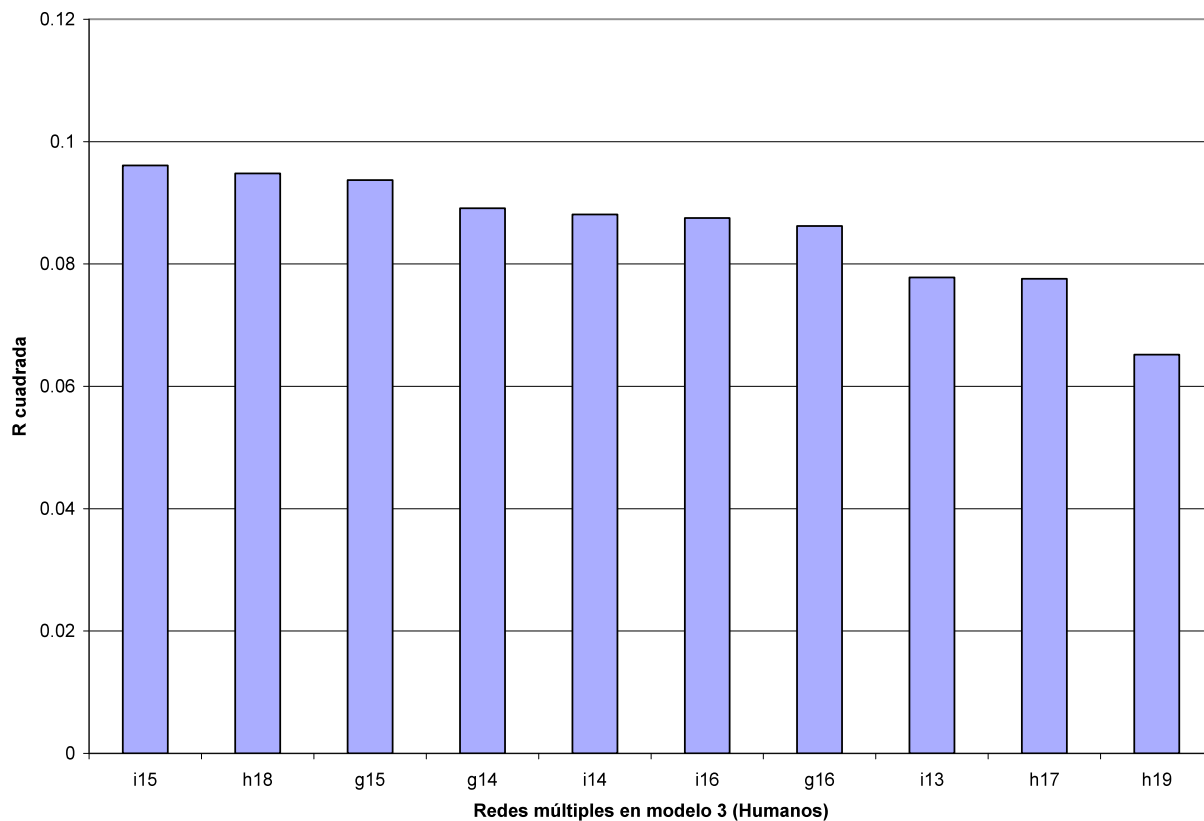


Gráfica 7.11. Valores de R^2 obtenidos para MODELO 2 en arquitecturas de redes múltiples.

MODELO 3 (Variables humanas)

Este modelo presentó los más bajos valores obtenidos para las observaciones de este estudio. Los valores de R^2 no superaron el valor de 0.1. Los valores más altos de R^2 fueron obtenidos utilizando las redes múltiples con diferentes FA (GAUSSIANA y TANH) y salto de conexión seleccionado (Gráfica 7.12). En general, los niveles más bajos de R^2 fueron obtenidos para este modelo con la arquitectura de tres grupos de neuronas con diferentes FA seleccionadas.

A pesar de ser esta red la recomendada por el mismo programa Neuroshell 2© como la mejor opción para llevar a cabo labores de predicción, las redes múltiples Ward no aportaron buenos resultados para el conjunto de datos seleccionados para muestra.

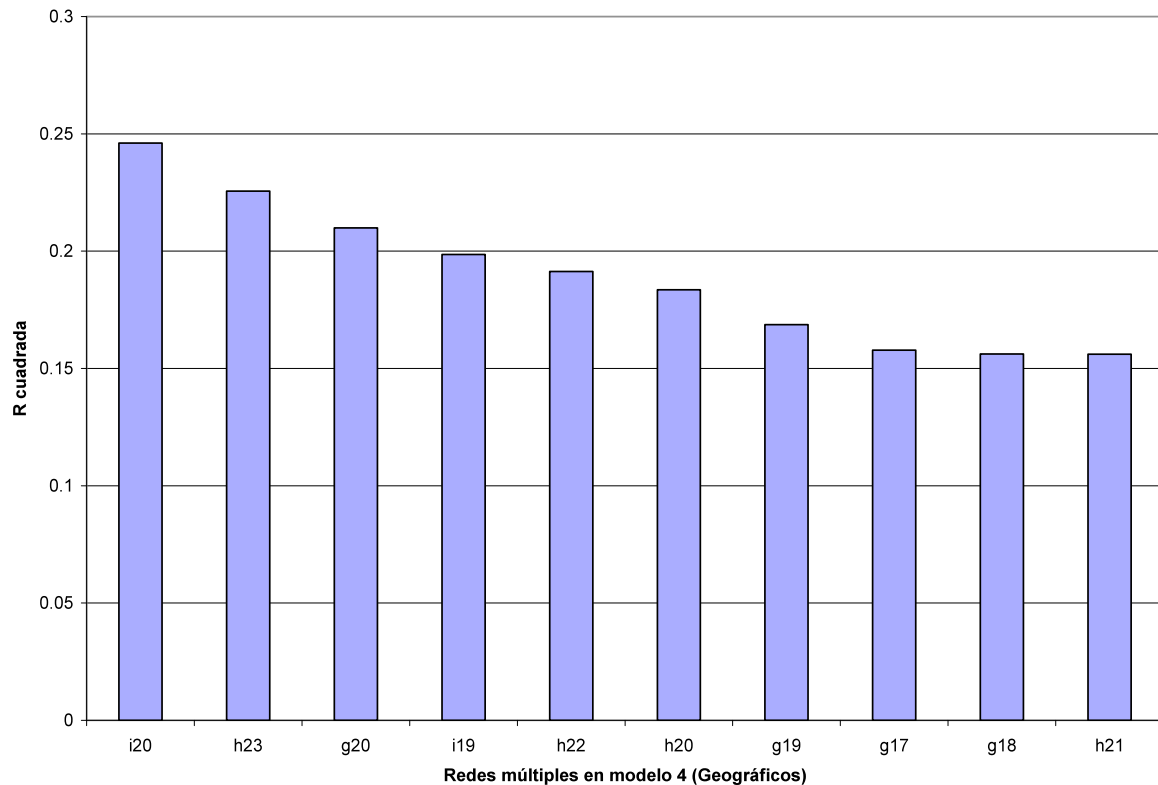


Gráfica 7.12. Valores de R^2 obtenidos para MODELO 3 en arquitecturas de redes múltiples.

MODELO 4 (Variables geográficas)

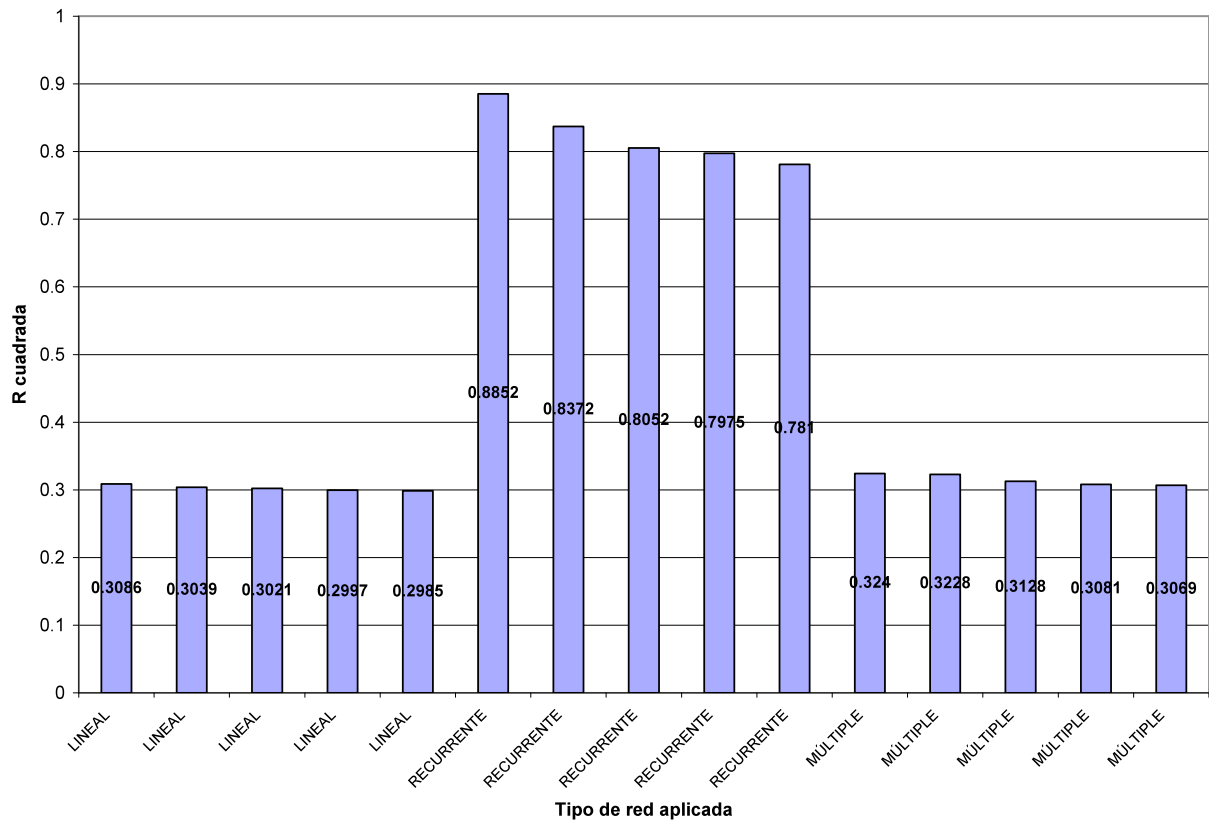
Valores de R^2 que oscilaron entre 0.1216 a 0.2461 fueron obtenidos mediante la utilización de FE TANH y [-1,1], así como combinaciones de FA GAUSIANA, LOGÍSTICA y TANH en la capa oculta. Al igual que en los modelos anteriores, esta arquitectura de red no presentó mejoría en los datos de la muestra (Gráfica 7.13).

De acuerdo a los datos obtenidos, es posible determinar que las redes múltiples no son las adecuadas para entrenar con las variables propuestas en el estudio.



Gráfica 7.13. Valores de R^2 obtenidos para MODELO 4 en arquitecturas de redes múltiples.

A continuación, se procedió a evaluar las mejores redes obtenidas del procedimiento anterior, que fueron escogidas mediante el criterio de que tuviesen los niveles de R^2 más altos, que en este caso fueron las redes recurrentes con retroalimentación en la capa oculta. Con el entrenamiento de cada una de estas redes, se aplicó. Al comparar los resultados, se observó que aunque tenían un nivel de R^2 alto, ninguna de las redes lograr mejorar la R^2 en la etapa de aplicación de la mejor red (“e11”) con los 8,938 datos que conforman el conjunto del celdas con vegetación susceptible a incendios forestales.

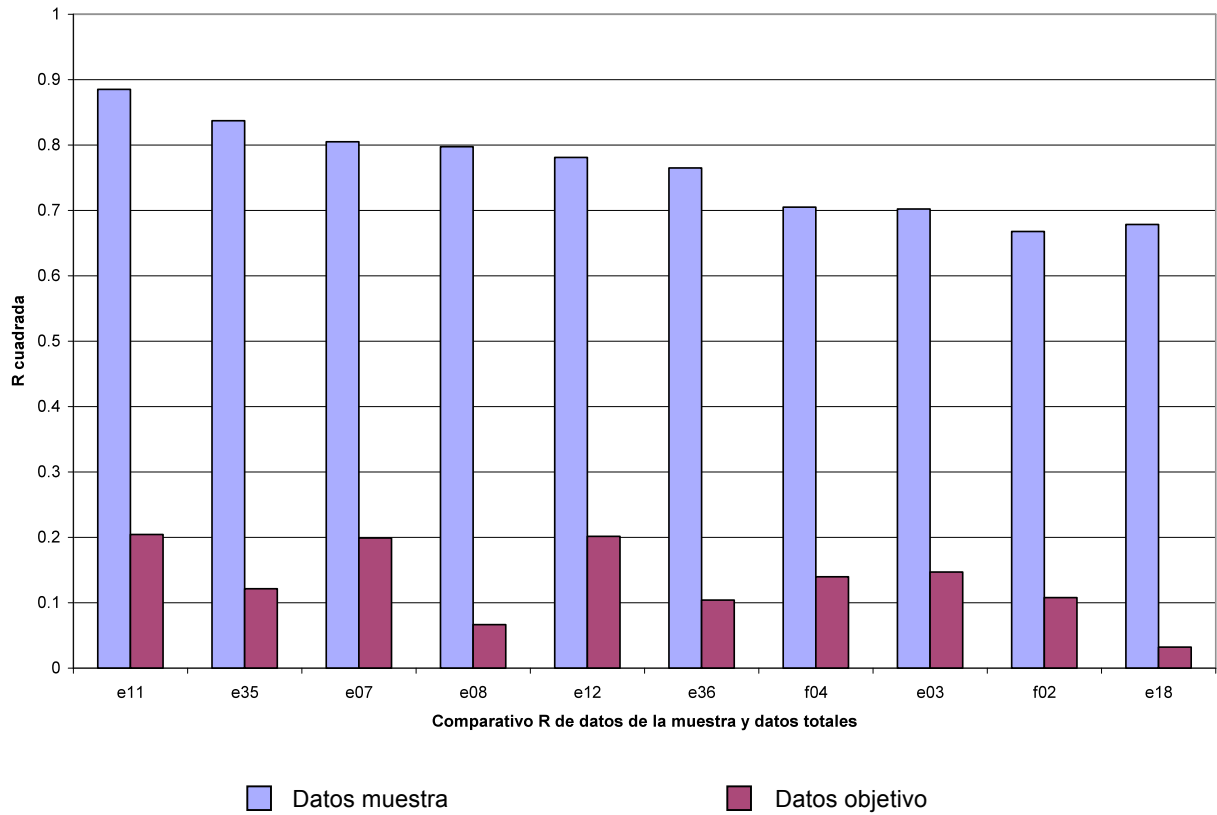


Gráfica 7.14. Mejores valores de R^2 obtenidos para cada arquitectura de red.

De lo anterior, fue posible determinar que las mejores redes que funcionaron para la muestra fueron las de arquitectura recurrente o Jordan-Elman, ya que presentaron valores de R^2 cercanos a 0.9. en tanto que las redes múltiples (Ward) y lineales no sobrepasaron el 0.4 del mismo parámetro. La Gráfica 7.14 muestra los valores de R^2 alcanzados por las cinco mejores redes de cada arquitectura.

Aplicación de la mejor red en los datos-objetivo

El análisis de los resultados de los procedimientos practicados a los datos de la muestra de 2,800 observaciones permiten determinar que las mejores redes del estudio fueron las recurrentes (Jordan-Elman) mismos que obtuvieron valores de R^2 mayores al 0.8, en tanto que las redes múltiples (Ward) y lineales no sobrepasaron el 0.4 del mismo parámetro. La Gráfica 7.15 muestra los valores de R^2 alcanzados por las cinco mejores redes de cada arquitectura.



Gráfica 7.15. Comparativa de valores de R^2 obtenidos en la muestra y en la aplicación de la red a los datos-objetivo

La Figura 7.15 muestra una gran diferencia entre los niveles de R^2 obtenidos con la muestra de 2,800 observaciones con los obtenidos al procesar los 8,938 datos que conforman el conjunto de los datos-objetivo. Al igual que en la muestra, fue posible verificar que los niveles de R^2 más altos se obtuvieron con redes recurrentes retroalimentadas en la capa oculta y en la capa de entrada (redes de arquitectura “e” y “f”), pero los mismo no superaron valores de R^2 de 0.20, habiendo una diferencia mayor de 0.6 en promedio.

Se aplicó la mejor arquitectura de red denominada “e11”, la cual corresponde a una red de tipo recurrente con retroalimentación en la capa oculta con las trece variables de entrada y como salida los puntos de calor del período 2000 a 2005 en modo

TURBOPROP, con una FE de [-1,1] y FA GAUSIANA en las capa oculta y LOGÍSTICA en la capa de salida, empleando la totalidad del conjunto de los datos-objetivo (8,938 observaciones) a fin de evaluar su comportamiento.

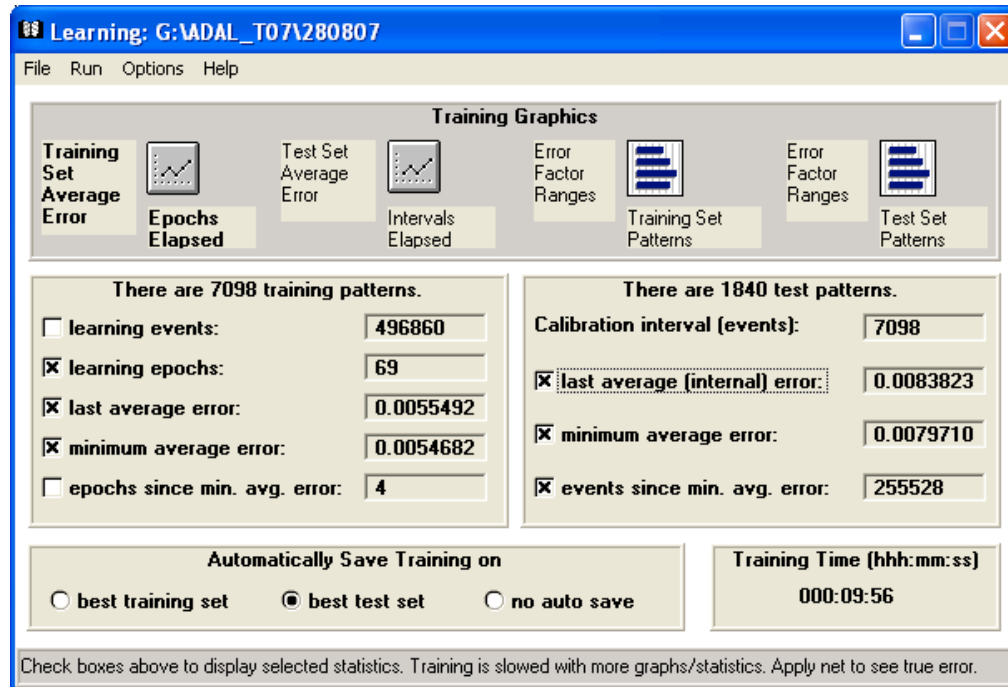
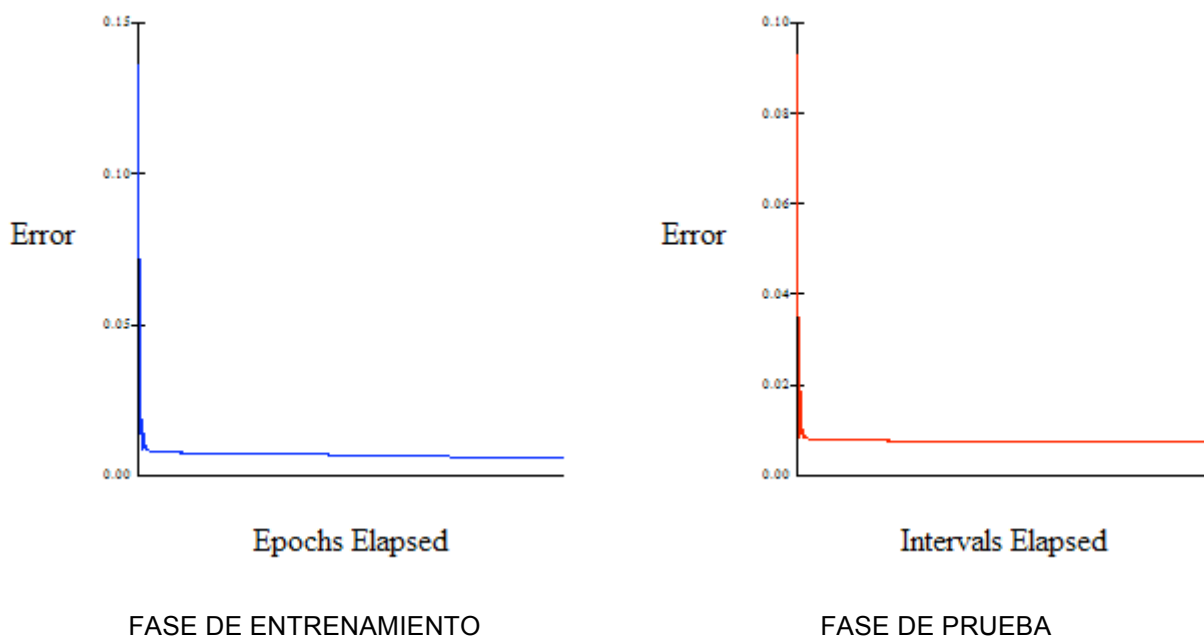


Figura 7.1. Resultados del entrenamiento de la mejor red para los datos-objetivo.

De la etapa de entrenamiento llevado a cabo en 69 épocas se alcanzó un EMP de 0.0054682 y de 0.007971 en la etapa de prueba. Al aplicarse la red al conjunto de datos solo fue posible alcanzar una R^2 de 0.2044, una r^2 de 0.2046 y un Ccr de 0.4523. El tiempo empleado para el entrenamiento de la red tomó 9 minutos con 56 segundos (Figura 7.1).

Las gráficas tanto de entrenamiento como de prueba mostraron un buen comportamiento de la red en la totalidad de las épocas, pues no presentaron oscilaciones importantes. La utilidad TURBOPROP proporcionó un entrenamiento

rápido y constante presentando poca variabilidad. Sin embargo, en la etapa de aplicación de la red, no logró obtenerse un buen nivel de R^2 (Gráfica 7.16).



Gráfica 7.16. Comportamiento de R^2 con la mejor red para los datos-objetivo.

La evaluación comparativa en la aplicación de la mejor red para los datos tanto de la muestra como de los datos-objetivo se presenta en la Tabla 7.2. Asimismo, los Factores de Contribución Relativa (FCR) por cada variable son mostrados en la Tabla 7.3:

Tabla 7.2. Valores obtenidos en la aplicación de la mejor arquitectura.

CAMPO	DATOS MUESTRA	DATOS-OBJETIVO
<i>Patterns processed</i>	2,800 obs.	8,938 obs.
<i>Output</i>	C1 (00-05)	C1 (00-05)
<i>R squared</i>	0.8852	0.2044
<i>r squared</i>	0.8862	0.2046
<i>Mean squared error</i>	0.003	0.167
<i>Mean absolute error</i>	0.033	0.251
<i>Min. absolute error</i>	0	0
<i>Max. absolute error</i>	0.524	3.281
<i>Correlation coefficient r</i>	0.9414	0.4523
<i>Percent within 5%</i>	8.857	0.145
<i>Percent within 5% to 10%</i>	9.429	0.168
<i>Percent within 10% to 20%</i>	13.343	0.347
<i>Percent within 20% to 30%</i>	9.179	0.492
<i>Percent over 30%</i>	9.143	14.511

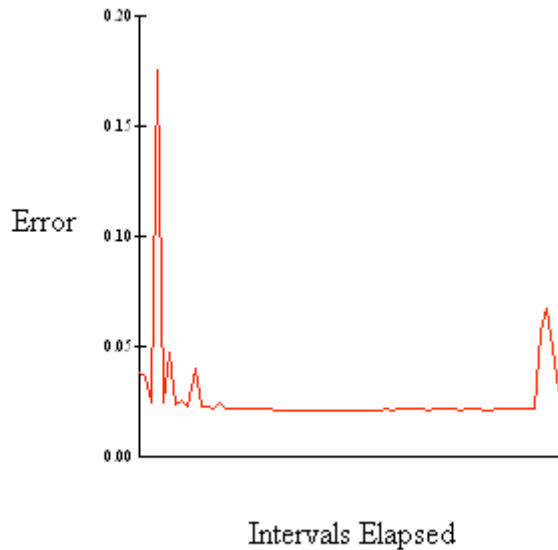
Tabla 7.3. Contribución relativa para cada variable en la mejor arquitectura.

DATOS MUESTRA		DATOS-OBJETIVO	
VARIABLE	CONTRIBUCIÓN RELATIVA	VARIABLE	CONTRIBUCIÓN RELATIVA
TEMP_MED	0.00830	TEMP_MED	0.01668
PREC_MED	0.00751	NCAL_ECOL	0.01407
NCAL_ECOL	0.00706	PREC_MED	0.01354
DIST_CARR	0.00681	REHUSUEL	0.01177
ZC	0.00670	CAB_MUN	0.01095
REHUSUEL	0.00645	ALT_CAT	0.01012
APROV	0.00635	DIST_CARR	0.00975
ALT_CAT	0.00605	ZC	0.00973
NANP	0.00516	PEND_CAT	0.00905
LIN_COND	0.00485	APROV	0.00901
US_SUELO	0.00457	US_SUELO	0.00895
PEND_CAT	0.00457	NANP	0.00875
CAB_MUN	0.00401	LIN_COND	0.00818

Los resultados obtenidos para la red propuesta no proporcionaron una R^2 aceptable. Los factores de contribución relativa de las variables son bajos, pues estos deberían de mostrar niveles más altos (de acuerdo a los parámetros establecidos en el mismo programa de Neuroshell 2®). El bajo desempeño de esta red se ve reflejado en el bajo valor de R^2 producido en la etapa de aplicación a la red, a sí como en los bajos niveles de FCR presentados en la Tabla 7.3.

7.3. Aplicación de la mejor red en los datos de 2006

Fue aplicada la misma red, ahora con el entrenamiento obtenido por la red recurrente en la capa oculta bajo los mismos parámetros utilizando como datos de salida los puntos de calor registrados en el año de 2006, obteniéndose la siguiente gráfica:



Gráfica 7.17. Etapa de prueba que muestra el nivel de EMP en relación a las épocas transcurridas en la aplicación de la red a los datos de 2006.

Tabla 7.4. Valores obtenidos en la aplicación de la red a los datos de 2006.

CAMPO	VALOR
<i>Patterns processed</i>	8938
<i>Output</i>	C1
<i>R squared</i>	0.1052
<i>r squared</i>	0.1061
<i>Mean squared error</i>	0.029
<i>Mean absolute error</i>	0.059
<i>Min. absolute error</i>	0
<i>Max. absolute error</i>	1.000
<i>Correlation coefficient r</i>	0.3257
<i>Percent within 5%</i>	0
<i>Percent within 5% to 10%</i>	0
<i>Percent within 10% to 20%</i>	0
<i>Percent within 20% to 30%</i>	0
<i>Percent over 30%</i>	3.368

En la Gráfica 7.17 se observa que existe un comportamiento errático al principio, ya que en las primeras épocas se registran valores altos de error mínimo promedio (EMP), repitiéndose casi al final de la etapa de prueba. El valor mínimo de EMP fue solo de 0.0355 en la etapa de prueba, siendo que para ser una red eficiente tendría que tener valores cercanos al 0.005.

La Tabla 7.4. muestra que posterior a la aplicación de la red, ésta no predijo correctamente los datos de PC detectados en el año 2006, ya que solo se obtuvo una R^2 de tan solo 0.1052 lo cual la hace inadecuada para éste fin. Aun así, se alcanzó un Ccr de 0.3257, el cual muestra que si existe correlación entre las variables usadas, aunque no en un nivel que pueda considerarse aceptable.

7.4. Discusión de resultados

Los valores esperados para el procesamiento de la totalidad de los datos tendrían que ser similares a los obtenidos por Buendía (2000), (una R^2 de 0.8617 en la fase de entrenamiento y una R^2 de 0.8514 en la fase de prueba) pero este nivel de desempeño sí fue alcanzado con los resultados del entrenamiento en la muestra de 2,800 observaciones.

Es de resaltar que en Vega-García *et al.* (1996), se obtuvieron valores de R^2 cercanos al 0.9 empleando también una muestra de igual número de observaciones con y sin presencia de incendios, aunque las variables empleadas además incluyeron factores de tipo temporal, además de factores geográficos y antropogénicos. Dichos resultados fueron obtenidos usando muestras similares, pero en una superficie de menor extensión. De igual forma, en Vega-García *et al.* (1999) se utilizó una metodología similar, aunque el objetivo principal de este estudio fue la comparación de dos procedimientos distintos (redes neuronales artificiales y procedimiento estadístico LOGIT). La diferencia entre los resultados del conjunto de datos de la muestra y los del conjunto de los datos-objetivo se debió a que existe un mayor número de observaciones cuyos valores de entrada fueron igual a cero, y sus correspondientes datos de salida tuvieron gran variabilidad por lo que la misma no logró generalizar adecuadamente las

salidas en la red. Comparando el estudio llevado a cabo por Buendía (2000), es de resaltar que las diferencias sustanciales fueron en el origen de los datos empleados: el tamaño de la resolución espacial de este estudio fue de 1.1 km provenientes de imágenes procesadas por el sensor AVHRR, en tanto que Buendía (2000) empleó datos procesados de imágenes Landsat, cuya resolución espacial es de 20 m. Aunque la superficie de estudio fue sustancialmente mayor, el número de observaciones entre los dos estudios fue similar (aproximadamente 8,500 observaciones en promedio).

8. CONCLUSIONES Y RECOMENDACIONES

8.1. Conclusiones

Con base a los resultados obtenidos en el estudio, es posible determinar que las redes recurrentes Jordan-Elman fueron la mejor arquitectura de red para procesar los datos de una muestra de 2,800 observaciones en todos los modelos propuestos, principalmente para los MODELOS 1 y 2, ya que se alcanzaron valores de R^2 superiores al 0.8. El mejor modelo fue el que incluyó la totalidad de las trece variables (MODELO 2), y posteriormente, el modelo que incluyó solo las seis variables seleccionadas por el procedimiento de GRNN (MODELO 1), las cuales fueron: temperatura media (TEMP_MED), zonas de campamento (ZC), calidad ecológica del sitio (CAL_ECO), régimen de humedad del suelo (REHUSUEL), aprovechamientos forestales en el área de estudio (APROV) y distancia a las carreteras (DIST_CARR).

De tal manera que para los datos de la muestra, es posible afirmar que sí existe relación de factores antropogénicos (actividades humanas) con la presencia de puntos de calor. Sin embargo, dichos factores deben ser evaluados conjuntamente con factores geográficos para incrementar su precisión, debido a que el entrenamiento de las redes con los modelos que incluyen por separado las características geográficas y las humanas independientes, no proporcionaron niveles altos de R^2 . Por otra parte, los resultados obtenidos en la evaluación de los datos-objetivo (8,938 observaciones) se consideraron no concluyentes, pues no alcanzaron valores de R^2 mayores al 0.3.

Se determinó que los principales factores que modifican el desempeño de una red *backpropagation* en orden de importancia para este estudio fueron las variables

empleadas como datos de entrada, los datos empleados como datos de salida, el tipo de arquitectura empleada, las funciones de escalamiento (FE) en la capa de entrada y de activación (FA) y el modo de actualización de pesos (MAP).

Las variables empleadas en ambos estudios fueron similares en aquellas consideradas geográficas (entre las que se pueden mencionar: altitud, tipo de vegetación, pendiente) así como la inclusión de variables consideradas humanas tal como fueron usados en los estudios realizados por Vega-García *et.al.* (1996, 1999).

TurboProp resultó ser en este estudio el mejor modo de actualización de pesos (MAP) en comparación con los métodos VANILLA y MOMENTUM, ya que su uso permitió obtener los mejores niveles de R^2 en los tres tipos de arquitectura *backpropagation* del programa NeuroShell 2©. Asimismo, esta utilidad representó una buena opción al seleccionar automáticamente los parámetros de Tasa de Aprendizaje (LR) y Momentum (Mo), y es el método que menos tiempo requiere para obtener resultados.

Hasta donde abarcó el estudio, no fue posible determinar cómo influye el factor humano en la totalidad de los datos-objetivo pero sí fue posible en la muestra, usando solamente las variables inherentes a éste (ya que se obtuvo una R^2 de 0.8097 en una red recurrente con retroalimentación en la capa oculta en el MODELO 3 que incluyó sólo los factores humanos). Sin embargo, con base a los resultados obtenidos en la muestra, el desempeño de la red aumenta al incluir también factores geográficos.

El uso de algoritmos genéticos empleados para el diseño de redes neuronales fue un buen método alternativo para la selección de variables de entrada (como se aplicó la arquitectura de GRNN para determinar las variables que conformaron el MODELO 1).

8.2. Recomendaciones

La inclusión de variables temporales tales como FECHA o TEMPORADA (como temporada de vacaciones o fines de semana) aumentarían la precisión de la RNA, ya que por si solas, las variables geográficas y humanas no aportan suficientes parámetros para diseñar redes que pronostiquen eficientemente posibles puntos de calor en la zona de estudio.

También se considera que en nivel de precisión podría incrementarse al considerar la susceptibilidad inherente que ante el fuego tiene cada tipo de vegetación (como por ejemplo, que los bosques de *Pinus* que se incendian con más frecuencia que los bosques de *Abies*).

La inclusión de mapas de combustibles como variables de entrada permitiría una mejor diferenciación en las grandes extensiones cubiertas por un solo tipo de vegetación. Estos fueron utilizados por Vega-García *et al.* (1996 y 1999) con buenos resultados, aunque sería recomendable también utilizar otras variables similares que permitan obtener estimar la misma información con fuentes distintas.

La resolución espacial del sensor AVHRR no equipara la resolución de otros tipos de sensores, como en el caso de la plataforma *Landsat*. Esta metodología puede

adaptarse a otras plataformas satelitales con el fin de evaluar otras características de la corteza terrestre, aprovechando otras resoluciones espaciales que se adapten mejor a las necesidades de cierta investigación.

Debe estudiarse más a fondo la posibilidad de integrar el empleo de los algoritmos genéticos en el diseño de redes neuronales, como se ha hecho en otros estudios similares, por ejemplo en Federico (2005); pero enfocados a evaluar características de la corteza terrestre y de los recursos forestales.

El uso de la automatización para llevar a cabo los procesos de transformación de datos espaciales, lectura, categorización, reescalamiento etc. mediante la programación optimizaría los tiempos que se requieren desde el procesamiento de las imágenes hasta la generación de resultados, ya que los métodos que se tuvieron al alcance en realización de este estudio no fueron los óptimos y consumieron gran parte del tiempo en la elaboración de las bases de datos y su adaptación para la red neuronal.

La actualización de paqueterías en redes neuronales artificiales podrían simplificar el tiempo que se emplea en el desarrollo de las mismas, pues incluyen nuevas herramientas que la versión de NeuroShell© utilizada en este estudio no se cuentan. Asimismo, nuevas paqueterías en sistemas de información geográfica como *Idrisi© Kilimanjaro* y más recientemente *Idrisi© Andes*, los cuales actualmente incluyen módulos especializados que realizan modelación de imágenes de satélite aplicando técnicas de redes neuronales.

9. LITERATURA CITADA

- Blackard, J.A., J.D. Denis. 1999. Comparative accuracies of artificial neural networks and discriminant analysis in predicting forest cover types from cartographic variables. *Computers and electronics in agriculture*. (24) 131 - 151.
- Buendía, R. E. 2000. Aplicación de redes neuronales artificiales y técnicas SIG para la predicción de coberturas forestales. Tesis de Maestría. División de Ciencias Forestales. Universidad Autónoma Chapingo, México. 74 p.
- Chuvieco E. y A. Russell. 1989. Application of remote sensing and geographics information systems to forest fire hazard mapping. *Remote Sensing Enviroment* Vol. 29, 147-159.
- Colegio de Ingenieros de Montes. 2002. Página electrónica del Colegio de Ingenieros de Montes. Ministerio del Medio Ambiente. España. <http://www.cim.es/incendiosforestales>.
- CONABIO. 2003. Página electrónica de la Comisión Nacional para el Uso y Conocimiento de la Biodiversidad. Gobierno Federal. México. <http://www.conabio.gob.mx>.
- CONAFOR. 2001. Programa Nacional Forestal 2001-2006. Comisión Nacional Forestal. Gobierno Federal. México. 256 p. http://www.conafor.gob.mx/programas_nacionales_forestales/incendios/.
- CONAFOR. 2006. Página electrónica de la Comisión Nacional Forestal. Incendios Forestales, generalidades. Gobierno Federal. México.
- CONAFOR. 2007. Reporte semanal de resultados de Incendios Forestales 2007. Nacional Forestal. Coordinación General de Coordinación y Restauración. Gerencia de Incendios Forestales. Gobierno Federal. México.
- Davis, B. y D. Dean. 2002. Identifying likely wildfire ignition points using topographic analysis, statistical, investigations and artificial intelligence. MS. Thesis. Departament of Forest Sciences, Colorado State University. 126 p.

- Federico, B. L. 2005. Entrenamiento de redes neuronales basado en algoritmos genéticos. Tesis de grado en ingeniería informática. Facultad de Ingeniería, Universidad de Buenos Aires. 239 p.
- Fiszelew, A. y R. García. 2000. Generación automática de redes neuronales con ajuste de parámetros basado en algoritmos genéticos. Facultad de Ingeniería. Universidad de Buenos Aires. 123 p.
- Giroux, E.M., D. Dean, 2002. Neural Networks an alternative to statical, modelling the semivarogram analysis portion of co-Kriging procedures. M.S. Thesis. Departament of Forest Sciences, Colorado State University. 23 p.
- Gobierno del Estado de México. 2003. Página electrónica del Gobierno del Estado de México. Diagnóstico. [http:// www. edomexico. gob. mx /se/diagedo.html](http://www.edomexico.gob.mx/se/diagedo.html).
- INEGI 2007. Página electrónica del Instituto Nacional de Estadística, Geografía e Informática. México. [http:// www. inegi. gob. mx/ entidades /espanol/edomex .html](http://www.inegi.gob.mx/entidades/espanol/edomex.html).
- ISA-UMH. 2007. Introducción a las redes neuronales. Área de Sistemas de Sistemas y Automática. Universidad Miguel Hernández de Elche. España. [http://isa.umh.es /isa/es/doct/rnia/Redes-1.PDF](http://isa.umh.es/isa/es/doct/rnia/Redes-1.PDF)
- Johnston, C. 1998. Geographic Information Systems in Ecology. Edited by Univerity of Minnessota, Blackwell Sciene. Duluth Minnessota, U.S.A. 99 p.
- Kourtz, P. 1990. Artificial Intelligence: A New Tool for Forest Management. Canadian Journal Forest Research. Vol 20. pp. 428-437.
- Lymberopoulos, N., C. Papadopoulos, L. Stefanakis, N. Pantalos, y F. Lockwood. 1996. A GIS-based forest fire management information system. EARSEL Advances in Remote Sensing 4(4): pp. 68-75.
- McCormick, R.J. 2002. On developing a meso–theoretical viewpoint of complex systems by exploring the use of artificial neural network in modeling wildfires. Presented at Forest Sat Symposium Heriot Watt University, Edinburgh. 15 p.

Microsoft. 2007. Inteligencia artificial. Artículo bibliográfico. Biblioteca de consulta Encarta.

Miller, J., K. Borne, T. Brian, Z. Huang, Y. Chi. 2005. Automated wildfire detection through artificial neural network. Goodard Space Fligth Center. Nasa technical reports server (NTRS). Document ID: 20050180456. 11 p.

Mora, F. y G. Hernandez-Cardenas. 1999. Modeling and mapping wildfire potential in Mexico based on vegetation and drought conditions using remote sensing and GIS technology. In Proceedings from the Joint Fire Science Conference and Workshop, vol. 2. Moscow, Idaho: University of Idaho. pp. 25-38.

Pew, K. L. and C.P. Larsen. 2001. GIS analysis of spatial and temporal patterns of human-caused wildfires in the temperate rain forest of Vancouver Island, Canada. Forest Ecology and Management. (140) pp. 1-18.

Pyne, S. J., P. L. Andrews, R. D. Laven. 1996. Introduction to wildland fire. Ed. Wiley. Second Edition. United States. 808 p.

PROBOSQUE, 2007. Programa de Prevención y combate de Incendios. Protectora de Bosques. Página electrónica del Gobierno del Estado de México. www.edomex.gob.mx/porta/probosque/incendiosforestales.html

Ramírez Rodríguez, J. 2001. Prevención y Combate de Incendios Forestales en el Estado de México Periodo 1999-2000. Tesis de Licenciatura. Universidad Autónoma Agraria Antonio Narro. México. 101 p.

Rodríguez Trejo, D. A. 1996. Incendios Forestales. Universidad Autónoma Chapingo. Editorial Mundi-prensa. México. 630 p.

SEMARNAT, 2002. Página electrónica de la Secretaría del Medio Ambiente, Recursos Naturales y Pesca. <http://www.semarnap.gob.mx>.

- Shlisky, A., J. Waugh, P. González, M. González, M. Manta, H. Santoso, E. Alvarado, E. Ainuddin Nuruddin, D. A. Rodríguez-Trejo, R. Swaty, D. Schmidt, M. Kaufmann, R. Myers, A. Alencar, F. Kearns, D. Zollner y W. Fulks, 2007. El fuego, los ecosistemas y la Gente: Amenazas y estrategias para la conservación global de la biodiversidad. Informe Técnico de la Iniciativa Global para el Manejo del Fuego 2007-2. The Nature Conservancy. Arlington, VA.
- Soria, J., C. Ortiz, F. Islas, V. Volke. 1998. Sensores Remotos. Principios y aplicaciones en la evaluación de los recursos naturales. Experiencias en México. CONACYT. Colegio de Postgraduados. Sociedad Mexicana de la Ciencia del Suelo A.C. Publicación Especial No. 7. México.
- Valencia, D. 2005. Pre-evaluation of wildfires in Monfraüe Regional Park using CHRIS imagery. ESA CHRIS/Proba workshop, 21-32 March ESRIN, Frascati, Italy.
- Vega-García, C., B.S. Lee, P.M. Woodard, S.J. Titus. 1996. Applying neural network technology to human-caused wildfire occurrence prediction. *AI Applications*. Vol 10, No. 3.
- Vega-García, C., P.M. Woodard, S.J Titus, W.L. Adamowicz, B.S. Lee. 1999. Dos modelos para la predicción de incendios forestales en Whitecourt Forest, Canada. *Invest. Agr.: Sist. Recur. For.* Vol. 8 (1).
- Ward Systems Group. 1996. Tutorial de referencia y ayuda para Neuroshell© Ver. 2 The Ward Systems Group Inc. Executive Park West, 5 Hillcrest, Dr. Frederick MD, U.S.A.
- Young, R. 1991. Introducción a las ciencias forestales. Editorial LIMUSA, México. pp. 411-434.