



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERIA.

**“GENERACIÓN DE DATOS FALTANTES
EN SERIES DE TIEMPO HIDROMETEOROLÓGICAS
CON REDES NEURONALES ARTIFICIALES”**

TESIS

**QUE PARA OBTENER EL TÍTULO DE
INGENIERO GEOFÍSICO**

PRESENTA:

JUAN CARLOS ALVA PACHECO

DIRECTOR DE TESIS:

DR. TOMÁS MORALES ACOLTZI

MÉXICO D.F. JUNIO DE 2012.



ÍNDICE

	Página.
Resumen	1
Abstract	1
Introducción	2
Capítulo 1. Calidad de una base de datos	3
1.1 Base de datos (BD)	3
1.2 Series de tiempo (ST) – análisis	3
1.3 Calidad de la BD	7
1.3.1 RClimDex	7
1.4 Falta de datos	9
1.4.1 Zona de estudio	9
1.4.2 ST incompletas	12
Capítulo 2. Base de datos de reanálisis	16
2.1 Reanálisis	16
Capítulo 3. Metodología. Redes neuronales artificiales	26
3.1 introducción	26
3.2 Redes neuronales biológicas (RNB)	28
3.3 Comparación de las ‘RNB’ con las ‘RNA’	30
3.4 La neurona artificial	31
3.4.1 Modelo general	31
3.4.2 Modelo estándar	32
3.4.3 Conceptos de la estructura de una neurona artificial	34
3.5 Ventajas de las RNA	37
3.6 Arquitecturas de las redes neuronales	37
3.7 Mecanismos de entrenamiento (aprendizaje)	39
3.7.1 Aprendizaje supervisado	40
3.7.2 Aprendizaje no supervisado y autoorganizado	40
3.7.3 Aprendizaje reforzado (reinforcement learning)	41
3.8 Red neuronal supervisada	41
3.8.1 El perceptrón	42
3.8.2 El perceptrón multicapa	43
3.8.2.1 Aprendizaje por BP.	45
3.8.2.2 Aceleración del BP	47
3.8.3 Capacidad de generalización de la red	48
3.8.3.1 Validación cruzada (cross-validation)	48
Capítulo 4. Experimentos numéricos	51
4.1 Modelo empleado	51
4.2 Estación de Apizaco	54
4.3 Estación de Tlaxco	66

4.4 Estación de Españita	72
Capítulo 5. Análisis de resultados	83
5.1 Análisis de temperaturas máximas	83
5.1.1 Estación de Apizaco	83
5.1.2 Estación de Tlaxco	85
5.1.3 Estación de Españita	85
5.2 Análisis de temperaturas máximas	86
5.2.1 Estación de Apizaco	86
5.2.2 Estación de Tlaxco	87
5.2.3 Estación de Españita	87
Capítulo 6. Conclusiones y trabajos futuros	88
6.1 Conclusión	88
6.2 Trabajos futuros	90
Anexos	
A1 Análisis de base de datos observados	92
B1 Altura geopotencial	99
B2 ST observados y teóricos de temperaturas máxima y mínima a escala diaria de las estaciones meteorológicas en estudio.	99
B3 Fundamentos de estadística	113
C1 Sistemas y máquinas inteligentes (historia)	115
C2 Cerebro vs. Computador	116
Glosario	
Bibliografía	

RESUMEN

La calidad de una Base de Datos (BD) es de gran importancia para su procesamiento, en especial para el procesamiento de datos Geofísicos. Al hablar de Series de Tiempo (ST) de datos diarios hidrometeorológicos, se hace referencia a las variables que rigen nuestra atmósfera, particularmente en este trabajo se analizaron las variables temperatura máxima y mínima, medidas en estaciones meteorológicas del estado de Tlaxcala. Hay ST incompletas en los registros y es de suma importancia contar con una BD completa, con el propósito de hacer un buen pronóstico numérico de tiempo-clima. Para generar los datos faltantes, se hace uso de una BD diarios de las variables atmosféricas mencionadas, llamadas datos de Reanálisis, ésta, es una BD mundial disponible en la página electrónica de la North American Regional Reanalysis (NARR, por sus siglas en inglés). La metodología empleada son las Redes Neuronales Artificiales (RNA), se hace uso de ésta debido al procesamiento de información **No Lineal**, con la cual se realizan experimentos numéricos, para poder generar los datos faltantes en las ST, objetivo principal de este trabajo. El resultado de este trabajo es una BD diarios completa y validada, para realizar, en principio, un mejor pronóstico numérico del tiempo-clima.

ABSTRACT

The quality of a database (DB) is of great importance for processing, especially for geophysical data processing. Speaking of Time Series (TS) of daily hydrometeorological data refers to the variables that govern our atmosphere, particularly in this study were analyzed on maximum and minimum temperatures measured at weather stations in the state of Tlaxcala. There incomplete TS records and is extremely important to have a complete DB with the purpose of making a good forecast weather-climate. To generate the missing data, use is made of a DB daily atmospheric variables mentioned Reanalysis data calls; this is a global DB available on the website of the North American Regional Reanalysis (NARR, for its acronym in English). The methodology employed are Artificial Neural Networks (ANN), it makes use of information processing due to Nonlinear, with which numerical experiments are performed in order to generate the missing data in the TS, the main objective of this work. The result of this work is a complete and validated daily DB, for, in principle, improved numerical weather-climate prediction (NWCP).

INTRODUCCIÓN

Los trabajos con RNA, se emplean en distintas disciplinas académicas, para nuestro tema, se hace referencia a la meteorología; actualmente se ha implementado esta herramienta como en la ingeniería civil de estructuras, trabajos como el de Juan Bojorquez Mora (2011) con su tema de “Uso de redes neuronales artificiales para estimar la respuesta sísmica de sistemas estructurales”. En el área eléctrica de instrumentación como en la médica, con el tema de “Reconocimiento de texturas en imágenes médicas” por Carolina González González (2011). “Identificación inteligente de un proceso fermentativo usando el algoritmo GMDH (GroupMethod of Data Handling) modificado” por F. Hernández y F. Herrera (2012). “Rain pattern analysis and forecast model based on GPS estimated atmospheric water vapor content”, por Andres Seco et al. (2011).

La contribución del presente trabajo de tesis, es trascender a la metodología empleada, al plantear un panorama general respecto a los modelos de RNA; en base a esto, explotar su potencial y dar resultados confiables, de esta manera expandir la capacidad del método en base a los resultados obtenidos, en este caso, la generación de datos faltantes de ST hidrometeorológicas.

La estructura del trabajo del presente estudio se divide en seis capítulos, dando a conocer en un inicio la problemática que se tiene en la BD de las estaciones meteorológicas a trabajar (datos observados), de este modo, formular la operación que se efectuará para la solución de la problemática en cuestión. Adquiriendo una nueva de BD llamados de reanálisis (datos teóricos), disponible en la red con acceso a todo público; realizando un análisis de ambas bases y cuestionar las posibles soluciones que se obtendrían con ayuda de los datos teóricos. Posteriormente en el capítulo tres, se mostrará la metodología empleada, que son las RNA, desde los fundamentos y conceptos en que se basan, llegando al modelo que se ha de utilizar. Con la teoría del método, se realizan los experimentos numéricos para la generación de los datos faltantes (capítulo cuatro), utilizando los datos observados y teóricos para tal operación; seguido del análisis de los experimentos en el capítulo cinco. Concluyendo el trabajo y dando las posibles sugerencias a trabajos futuros, mostrados en el capítulo seis. Se anexan algunos análisis de la BD observados, como subtemas que contribuyen a las RNA.

CAPÍTULO 1

CALIDAD DE UNA BASE DE DATOS

1.1 BASE DE DATOS (BD).

Un conjunto de observaciones, de variables, representando la evolución de un sistema, a través de un conjunto organizado de Series de Tiempo (ST), define una BD.

Una ST, es la secuencia o colección de mediciones de cierto suceso o experimento, registrados secuencialmente en el tiempo, a intervalos de tiempo iguales. Comúnmente la notación de una ST es de la forma:

$$[y(t_1), y(t_2), y(t_3), \dots, y(t_n)] = y(t_i), i = 1, 2, 3, \dots, n \quad (1)$$

donde $y(t_i)$ es el valor tomado del proceso en el instante t_i .

Ejemplos de ST usadas en algunas áreas son:

- Economía: Precios de un artículo, tasas de desempleo, tasa de inflación, precio del dólar, precios de acciones, ingreso nacional bruto, etc.
- Meteorología: Precipitación, temperaturas máxima y mínima, velocidad del viento, energía solar, etc.
- Geofísica: Series sismológicas, gravimétricas, electromagnéticas.
- Química: Viscosidad y temperatura de un proceso.

En el presente trabajo, el interés se vuelve hacia la **meteorología**, en la que, las variables utilizadas pueden describir un sistema dinámico, por ejemplo, una onda del este, un huracán, un tornado, un frente frío, según la estación del año y la zona de estudio.

1.2 SERIES DE TIEMPO (ST) – ANÁLISIS

Por muy simple que parezca, el paso más importante en el análisis de ST, consiste en graficarlas. Ejemplos mostrados en la figura 1, son las ST de **precipitación**, **temperaturas máximas** y **mínimas**; al analizar los gráficos, podemos hacer varias observaciones sobre el comportamiento de dichas variables, el cual, en el área de meteorología es de suma importancia.

La temperatura es considerada la esencia del tiempo-clima; siendo tal vez, la variable más estudiada y sobre la que mejor se conocen sus repercusiones en el desarrollo de la agricultura. El comportamiento de la **temperatura máxima**, también es importante, sobre todo, porque es una variable de análisis constante en el tema de **Cambio Climático (CC)**, aunque en las demás variables se puede manifestar la firma del **CC** (Morales-Acoltzi, 2012).

Una variable no menos importante, es la **precipitación**, ya que tiene una gran influencia sobre los cultivos, al depender éstos de la cantidad, distribución y frecuencia de ésta durante el año; y no sólo es primordial para los cultivos, también para el abastecimiento y rendimiento de las presas, cuyo nivel fluctúa desde los desbordes hasta el vacío de éstas.

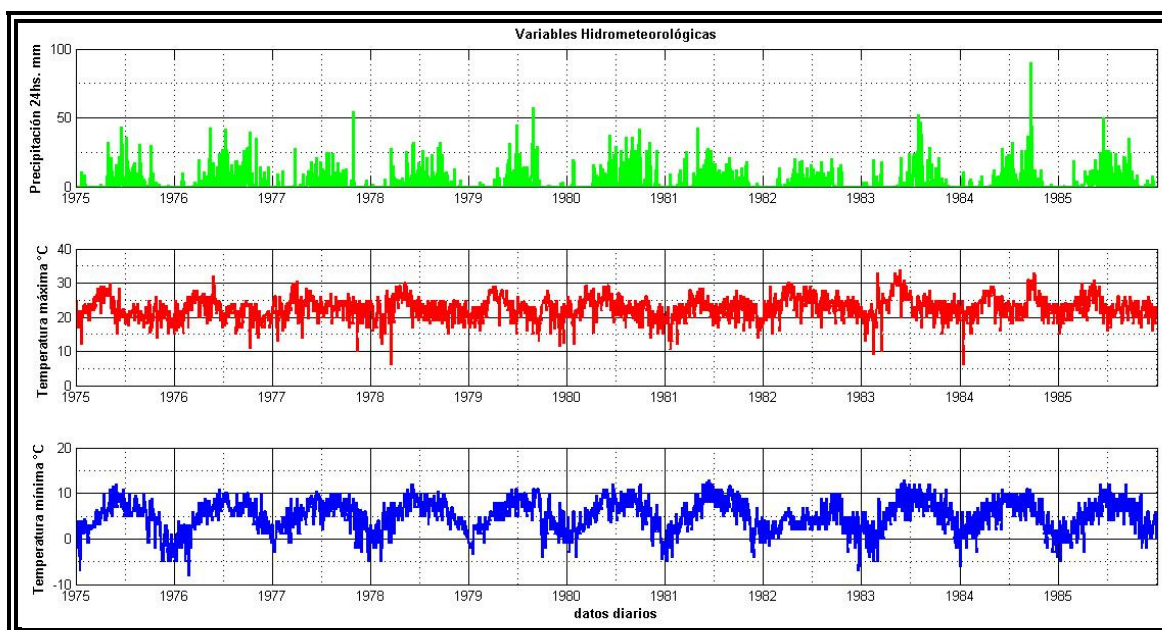


Figura 1.1 ST de las variables Precipitación, Temperaturas máxima y mínima
(Estación Meteorológica de Apizaco-Tlaxcala).

CAPÍTULO 1.

CALIDAD DE UNA BASE DE DATOS.

El gráfico de la ST, permite detectar algunos elementos, tales como los *valores extremos o valores atípicos (outliers)*. Son puntos que escapan a la tendencia general de la serie (ver figura 1.2), es decir, que distan en magnitud del resto de los datos. Con éstos, se debe reunir información adicional sobre posibles factores que afectaron algún proceso dentro del sistema en estudio, como la **Variabilidad Climática (VC)** y el **CC**.

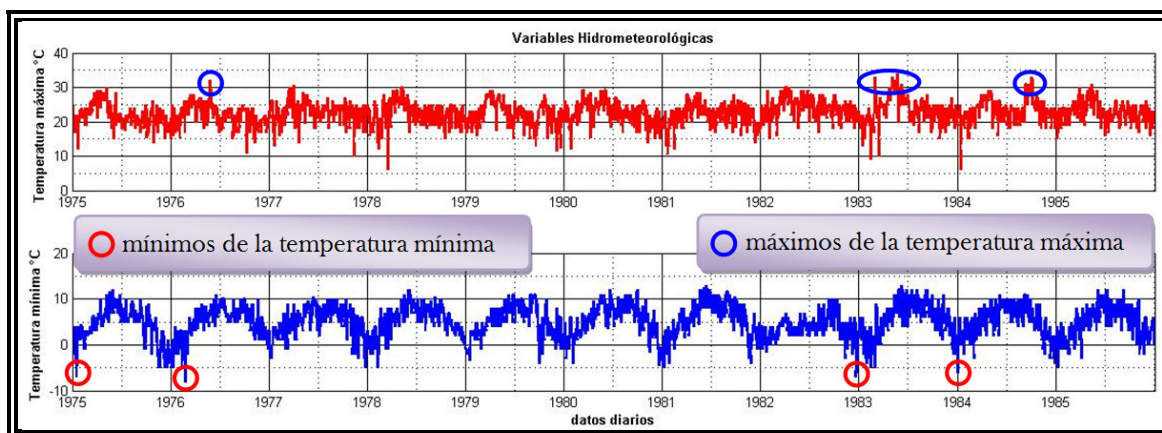


Figura 1.2 Valores extremos de las ST de las variables: Temperaturas máxima y mínima.

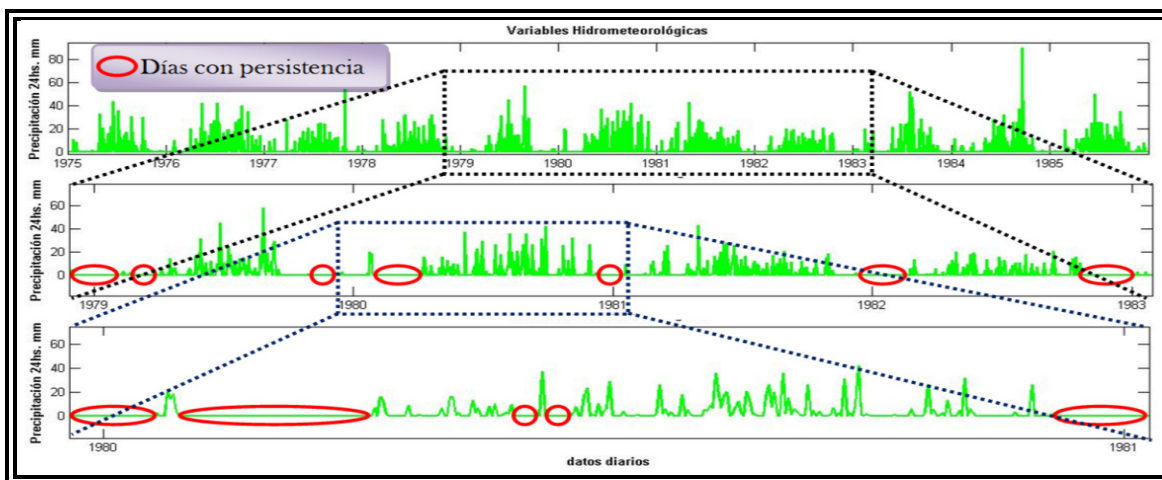


Figura 1.3 ST de Precipitación (pcp). Días con pcp=0 [mm].

Otro concepto sobre el que podemos hablar es la **persistencia**. ¿A qué nos referimos con **persistencia**?, es cuando tenemos valores repetidos (valores iguales) o que persisten durante un cierto periodo de tiempo; por definición, la persistencia es la acción y efecto de mantenerse constante en algo, durar por largo tiempo [Real Academia Española (RAE), 22.a edición]. Por ejemplo, podemos tener una persistencia de temperaturas máximas con el mismo valor de tres días; o bien, varios días en que no haya lluvia (precipitación igual a cero) (ver figura 1.3).

CAPÍTULO 1.

CALIDAD DE UNA BASE DE DATOS.

La persistencia constituye el método empírico de pronóstico más sencillo, cualquier modelo de pronóstico debe superarla, de no ser así, el modelo propuesto es inútil.

El análisis de ST estudia el comportamiento temporal de una o varias variables; y sus métodos, investigan secuencias de datos como función del tiempo. Existen varios objetivos posibles al analizar las ST (modificado de Chatfield, 1995):

- a) *Descripción*: permite obtener medidas descriptivas de las propiedades principales de la serie para identificar patrones, valores extremos, puntos de bifurcación, etc. Ya antes mencionadas.
- b) *Explicación*: al registrar dos o más variables de un sistema es posible usar la variación en una ST para explicar la variación de otra serie. Lo que puede llevar a un entendimiento mayor sobre el mecanismo que generó una determinada serie y así establecer ciertas propiedades del sistema, aunque esto depende del grado de acoplamiento entre las variables.
- c) *Predicción*: el entender el proceso subyacente (mecanismo) que produjo los **datos observados**, permite predecir futuros valores de una variable, a partir de una ST.
- d) *Control*: cuando se genera una serie temporal que mide la “calidad” de un proceso, el objetivo del análisis puede ser el control de los parámetros del primero.

Para los sucesos naturales, su comportamiento tiende a ser **no lineal y complejo**. En general, la mayoría de los sistemas naturales (reales) son **no lineales**, aunque frecuentemente se utiliza un *modelo lineal*, como primera aproximación a su dinámica. Este tipo de modelos son preferibles desde el punto de vista científico, porque el análisis matemático implicado, es menos complejo que para el caso de los **sistemas no lineales** [Peña Maciel, 2012].

Para el análisis de ST de **sistemas atmosféricos**, se han realizado varios estudios utilizando distintos métodos, con el fin de mejorar el entendimiento de sus procesos y demostrar que las variables que los describen tienen un **comportamiento no lineal**, como autores principales de algunos de los trabajos tenemos a: Meneses-Popocatl (1993), Hernández-Arriaga (1996), Tlaxcala-Castañeda (2005), Casillas-Lavín (2004), Sánchez-Pérez. (2006), Chamizo-Checa (2008), Morales-Acoltzi (2011) y Peña-Maciel (2012). Estos trabajos conciernen al Estado de Tlaxcala, área de estudio para el trabajo de tesis presente.

1.3 CALIDAD DE LA BD.

Al hablar de calidad de datos, se considera que éstos no presentan tendencias o **cambios no naturales** (*tendencias inducidas*), en el periodo de registro de los mismos. Pero generalmente, hay **datos faltantes** (*datos perdidos*) en las ST, debido principalmente a: cambios de sensores y/o de lugar del equipo, falta de calibración u operación (mantenimiento). Inclusive puede haber errores de captura (*error humano*), lo que implica la necesidad de aplicar un sistema de **calidad de datos**, específicamente una prueba de homogeneidad de datos, como la prueba de *Sved-Eisenhart* o una prueba de *igualdad de varianza*.

La **falta de datos** en la ST en cuestión, puede incluir la ausencia de uno, dos, o más de ellos, lo cual nos quita información, puesto que es necesario conocer a profundidad el comportamiento del sistema dinámico en estudio, para considerar los posibles valores extremos asociados a su evolución.

En el caso del sistema climático, el panel intergubernamental de **cambio climático** (IPCC)¹ propone un software para analizar la **calidad de datos** y el cálculo de índices², “RClimDex”.

1.3.1 RClimDex

ClimDex es un programa basado en Microsoft Excel que proporciona un paquete computacional, fácil de usar para el cálculo de *índices* de extremos climáticos para monitorear y detectar **CC**. Fue desarrollado por Byron Gleason del National Climate Data Centre (NCDC) de National Oceanic and Atmospheric Administration (NOAA) [Santos J.L., 2004].

¹ El Grupo Intergubernamental de Expertos sobre el **Cambio Climático** o Panel Intergubernamental del **Cambio Climático**, conocido por el acrónimo en inglés IPCC (Intergovernmental Panel on Climate Change), se estableció en el año 1988 por la Organización Meteorológica Mundial (OMM o por sus siglas en inglés WMO, World Meteorological Organization) y el Programa Ambiental de las Naciones Unidas (UNEP, United Nations Environment Programme).

Una de las principales actividades del IPCC es hacer una evaluación periódica de los conocimientos sobre el **CC**. <http://www.ipcc.ch/> [consulta: 6-01-12]

² Índices Climáticos. Caracterización numérica de las variables del clima y de sus relaciones. RClimDex calcula todos los 27 índices básicos recomendados por, “The Expert Team Climate Change Detection Monitoring and Indices” (ETCCDMI)

CAPÍTULO 1.

CALIDAD DE UNA BASE DE DATOS.

Se descubrió que el método utilizado por ClimDex para calcular índices de temperatura basados en percentiles³ resultaba en in-homogeneidad en las series. Una solución a este problema requiere un procedimiento "bootstrap"⁴ que se hace imposible de implementar en un ambiente Excel.

El usar *R* como plataforma para ClimDex, es debido a que *R* es un lenguaje y entorno de programación muy robusto y poderoso para gráficos y análisis estadístico. *R*, trata de un proyecto de "software libre", al igual que contribuye con la posibilidad de cargar diferentes bibliotecas o paquetes con finalidades específicas de cálculo o gráfico. Está disponible para los sistemas operativos Windows, Macintosh, Unix y GNU/Linux.

Uno de los principales objetivos del RClimDex es el control de calidad de datos climáticos diarios para que sean calculados los índices extremos y usarlos en estudios de monitoreo y detección de cambios climáticos. Esto requiere que los índices sean homogeneizados.

Es de suma importancia, que antes de utilizarse los datos de las ST, se efectúe un serio control de la calidad y preparación de la información.

Una metodología implementada para la verificación de la homogeneidad y consistencia de los datos, en este caso para las ST hidrometeorológicas, es el *test* de *Sved-Eisenhart*, el cual determina si existe alguna tendencia, o bien, si es muy oscilante. La tendencia u oscilación puede ser el resultado del cambio en el comportamiento atmosférico (lo que es el CC), o simplemente a errores sistemáticos de los instrumentos de medición.

3 Percentil. Mat. Valor que divide un conjunto ordenado de datos estadísticos de forma que un porcentaje de tales datos sea inferior ha dicho valor. Así, un individuo en el percentil 80 está por encima del 80% del grupo a que pertenece [RAE, 22.a edición]. Cada uno de los noventa y nueve segmentos que resultan de dividir algo en cien partes de igual frecuencia.

4 El bootstrapping (o bootstrap) es un método de remuestreo propuesto por Bradley Efron en 1979. Se utiliza para aproximar la distribución en el muestreo de un estadístico. Se usa frecuentemente para aproximar el sesgo o la varianza de un estadístico, así como para construir intervalos de confianza o realizar contrastes de hipótesis sobre parámetros de interés.

1.4 FALTA DE DATOS.

Debe considerarse que el software RClimDex, no completa los datos faltantes de las ST. Por ejemplo, se registró una precipitación igual a cero, pero llovió, al ser mínima, no fue perceptible para la estación meteorológica, lo que se conoce como *inapreciable*, lo que constituye el cálculo de un índice. Considerando también, que ésta variable sólo puede tomar valores correspondientes a cero y/o positivos.

Con lo mencionado, la BD de las ST hidrometeorológicas a trabajar (***datos observados***), han pasado por un control de la calidad y han sido homogeneizados mediante el software libre RClimDex(1.0). Con ello, se asegura que se ha disminuido el grado de incertidumbre de las mediciones presente en los datos. Para esta investigación, los ***datos observados*** fueron proporcionados, por el equipo de trabajo del PEACC-Tlaxcala⁵, ya con el control de calidad mencionado.

1.4.1 Zona de estudio.

Las variables hidrometeorológicas a utilizar en el presente trabajo son: **precipitación (pcp)**, **temperatura máxima (temp.máx.)** y **temperatura mínima (temp.mín.)**. Para las estaciones meteorológicas ubicadas en el **Estado de Tlaxcala: Apizaco, Tlaxco y Españita** (ver figura 1.4, 1.5 y tabla 1.1).

Tlaxcala, ha sido elegido zona de estudio para la evaluación a la vulnerabilidad y de adaptación al **cambio climático** (Morales-Acoltzi, 2009), consideradas en el proceso de planeación de desarrollo estatal y el desarrollo de políticas públicas sustentables (2009-2011); esto ha sido producto de varios estudios realizados, uno de ellos es por Morales-Acoltzi et al. (2012), denominado: “Variabilidad y Cambio Climático de la lluvia en Tres Sitios en el Estado de Tlaxcala”.

⁵ Programa Estatal de Acción ante el Cambio Climático, estado de Tlaxcala.

CAPÍTULO 1.

CALIDAD DE UNA BASE DE DATOS.

Tlaxcala presenta principalmente clima templado sub-húmedo. En primavera-verano presenta calores moderados que no pasan de los 30°C. En invierno hiela en casi todo el estado. En la cumbre del volcán La Malinche hay clima frío con nevadas ocasionales. Vientos dominantes en invierno y primavera procedentes del sur, en verano del noreste y en otoño del norte ⁶.

Las descripciones de los municipios fueron tomadas de la “Enciclopedia de los Municipios de México: Estado de Tlaxcala”⁷.

Apizaco. Ubicado en el Altiplano Central Mexicano a una altura media de 2 380 msnm. Su orografía presenta zonas planas, que comprenden el 45% de la superficie total; zonas accidentadas que abarcan el 30% y se localizan al norte del municipio; zonas semiplanas, que ocupan el 25% restantes de la superficie, se localizan en el sureste y parte norte del municipio.

Tlaxco. Ubicado en el Altiplano Central Mexicano a 2 540 msnm (altura media). Localizado al norte del estado, colinda al norte con el estado de Puebla, al sur con los municipios de Atlangatepec, Tetla y Muñoz de Domingo Arenas. Comprende el 60% sobre un terreno plano, ligeramente inclinado hacia oeste y sureste.

Españita. Ubicado en el Altiplano Central Mexicano con una altura media a 2 640 msnm. Colinda al poniente con el municipio de Sanctorum de Lázaro Cárdenas y el estado de Puebla. En general el municipio de Españita es accidentado; además se encuentran en él mesetas, llanuras y altiplanicies.

Ubicación de las estaciones meteorológicas										
Clave	Municipio	Latitud			Longitud			Altura	Coordenadas UTM ⁸	
		g	m	s	g	m	s		latitud	longitud
29002	Apizaco	19	25	41	98	7	51	2410.0	19.428055	98.130833
29032	Tlaxco	19	35	33	98	7	32	2530.0	19.59250	98.125555
29047	Españita	19	27	28	98	25	20	2570.0	19.457777	98.422222

Tabla 1.1 Localización de las estaciones meteorológicas.

[6] Tomado de: http://www2.ine.gob.mx/cclimatico/edo_sector/estados/tlaxcala.html [consulta: 6-01-12]

7 <http://www.e-local.gob.mx/work/templates/enciclo/tlaxcala/> [consulta: 30-03-12]

8 Universal Transverse Mercator (Universal Transversal de Mercator).

CAPÍTULO 1.
CALIDAD DE UNA BASE DE DATOS.

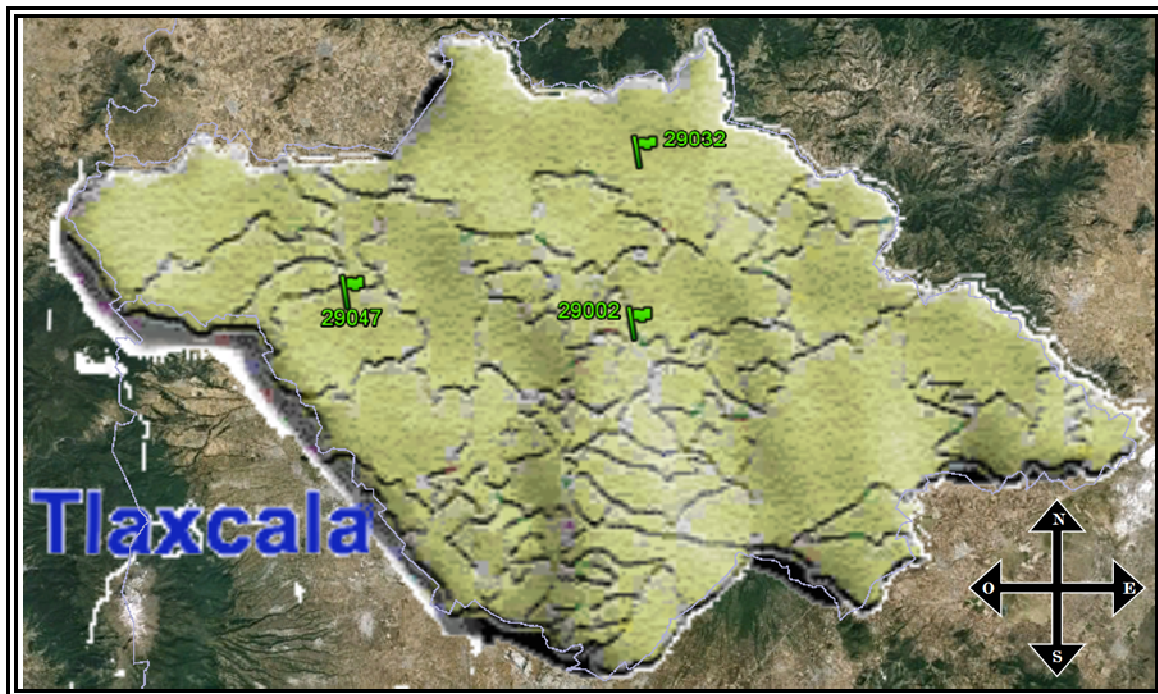


Figura 1.4 Localización de las estaciones.
Modificado de Google Earth.

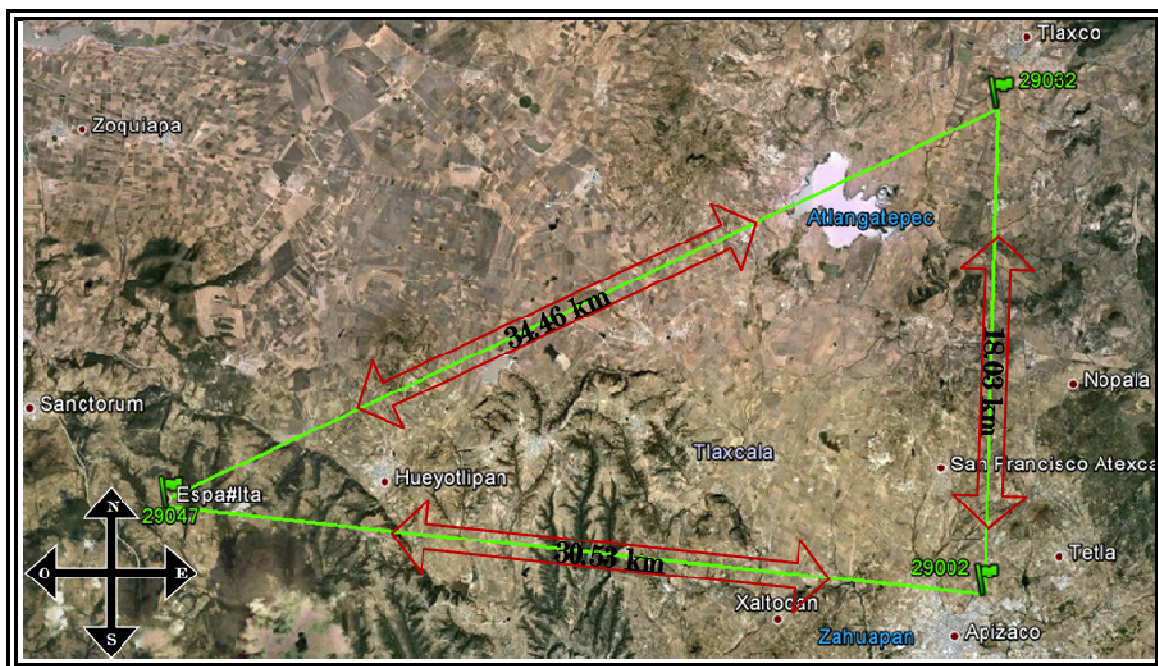


Figura 1.5 Imagen de Google Earth de la zona de Estudio.
Ubicación de las estaciones meteorológicas.

CAPÍTULO 1.
CALIDAD DE UNA BASE DE DATOS.

1.4.2 ST Incompletas.

Hemos descrito la zona de estudio con las estaciones meteorológicas a trabajar. En la siguiente tabla 1.2, se tiene el análisis sobre la BD para la estación de Apizaco.

ST APIZACO 01/01/1928 - 31/12/2010 datos diarios					
Precipitación		Temperatura máxima		Temperatura mínima	
Años de datos faltantes	Número datos faltantes	Años de datos faltantes	Número datos faltantes	Años de datos faltantes	Número datos faltantes
1928	30	1928	30	1928	30
1931	181	1931	151	1931	151
1932 - 1939	2922	1932 - 1950	6940	1932 - 1950	6940
1940	244				
		1951	150	1951	151
1970	1	1970	1	1970	1
1986	184	1986	184	1986	184
				1987	1
1989	1	1989	1		
1991	1	1991	1	1991	1
1994	2	1994	1		
2008	1	2008	2	2008	2
2009	5	2009	8	2009	7
2010	1	2010	1		
Precipitación		Temperatura máxima		Temperatura mínima	
ST APIZACO 01/01/1928 - 31/12/2010 datos diarios					

Tabla 1.2 Análisis sobre la BD, Apizaco.

Para esta estación, se tiene un periodo bastante amplio en el que no hubo datos para las tres variables, se presenta esta ausencia desde mediados del año 1931, hasta mediados de los 40's, para la **pcp** (9 años seguidos), y en ambas temperaturas hasta mitad del año 1951 (20 años consecutivos). El año 1986, a finales, medio año no se tiene datos; posterior a éste, hay ausencia de uno a ocho datos, en diferentes años.

En las siguientes tablas 1.3 y 1.4, se muestran los análisis de cada BD para las dos estaciones meteorológicas restantes.

CAPÍTULO 1.
CALIDAD DE UNA BASE DE DATOS.

ST TLAXCO 01/06/1941 - 31/12/2010 datos diarios					
Precipitación		Temperatura máxima		Temperatura mínima	
Años de datos faltantes	Número datos faltantes	Años de datos faltantes	Número datos faltantes	Años de datos faltantes	Número datos faltantes
1946	1				
1948	1				
1949	1				
1951	1				
		1950	2	1950	2
1953	5	1953	6	1953	6
		1957	2	1957	2
1958	31	1958	31	1958	31
				1960	1
1961	30	1961	30	1961	30
1962	3	1962	3	1962	3
1963	34	1963	32	1963	32
1964	92	1964	92	1964	92
1965	61	1965	61	1965	61
1966	153	1966	153	1966	153
1967	1				
1969	123	1969	123	1969	123
1985	31	1985	31	1985	31
1986	184	1986	184	1986	184
1987	1			1987	1
1989	4	1989	4	1989	3
1991	1			1991	2
		1992	1		
		2002	1	2002	1
2004	1				
2009	1	2009	1		
		2010	1	2010	1
Precipitación		Temperatura máxima		Temperatura mínima	
ST TLAXCO 01/06/1941 - 31/12/2010 datos diarios					

Tabla 1.3 Análisis sobre la BD, Tlaxco.

Para la estación Tlaxco, se tiene ausencia de datos en distintos años; tres años, marcados en color rojo, hay más de cien datos faltantes; independiente de la variable, hay un total máximo de veintiún años que presentan de uno a menos de cien datos faltantes.

CAPÍTULO 1.
CALIDAD DE UNA BASE DE DATOS.

ST ESPAÑITA 19/03/1977 - 31/12/2010 datos diarios					
Precipitación		Temperatura máxima		Temperatura mínima	
Años de datos faltantes	Número datos faltantes	Años de datos faltantes	Número datos faltantes	Años de datos faltantes	Número datos faltantes
1986	153	1986	153	1986	153
1988	3	1987	7	1987	5
		1988	181	1988	153
		1989	61	1989	62
		1990	1		
1991	1	1991	1	1991	1
1992	7	1992	7	1992	8
				1993	1
1994	1	1994	1	1994	1
1996	1	1996	1	1996	2
				2002	1
2003	15	2003	15	2003	14
		2005	1	2005	1
2006	15	2006	17	2006	17
2007	12	2007	11	2007	11
2008	8	2008	38	2008	37
2009	245	2009	245	2009	245
2010	1	2010	1	2010	1
Precipitación		Temperatura máxima		Temperatura mínima	
ST ESPAÑITA 19/03/1977 - 31/12/2010 datos diarios					

Tabla 1.4 Análisis sobre la BD, Española.

Españita presenta más de doscientos datos faltantes, en el año 2009 para las tres variables. En 1986 a 1988, superan los ciento cincuenta datos ausentes para ambas temperaturas y **pcp**, excepto 1988 para esta última. En distintos años, no hay de uno a sesenta y dos datos.

Las gráficas de las ST para cada estación y variable meteorológica, se muestran en el anexo A1.

CAPÍTULO 1.

CALIDAD DE UNA BASE DE DATOS.

Analizando las ST de las estaciones meteorológicas descritas, el ***problema a solucionar*** es la **generación de los datos faltantes** de estas ST hidrometeorológicas. El método que se empleará se describe en el capítulo 3, pero antes de ello, se buscaron datos válidos mundialmente, para poder obtener una relación desde un punto de vista regional a uno local, con el fin de obtener una base de ***datos teóricos*** y pensar que con ello podemos resolver la problemática que se tiene.

También debemos tener presente que la Organización Meteorológica Mundial (OMM), indica que se deben tener, al menos, 30 años de datos seguidos si queremos hablar de condiciones climatológicas.

CAPÍTULO 2

BASE DE DATOS DE REANÁLISIS.

Cuando se prueba una teoría y/o modelo de un proceso atmosférico-oceánico es deseable tener una plataforma *universal*, es decir, tanto computacionalmente como en las condiciones iniciales y/o de frontera del modelo (Morales-Acoltzi, 2011). Se debe de considerar *un problema bien puesto* desde el punto de vista matemático y computacional. Con respecto a las condiciones iniciales, se ha construido una BD conocida como “**BD de Reanálisis**”.

2.1 REANÁLISIS.

National Center for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR), es una página electrónica⁹ que forma parte de la NOAA. En ella, se encuentran disponibles datos del NCDC, desde el año de 1979 al presente.

La NOAA National Operational Model Archive and Distribution System (NOMADS), es un proyecto de servicios web basada en proporcionar acceso independiente al formato, tanto en tiempo real y retrospectivo, a los datos de los modelos tiempo-clima ^[10].

Para buscar los datos de interés, se ingresa al sitio web, están organizados por proveedor y tipos de datos. Los que nos son de utilidad, corresponden a la BD de la NCEP-NARR con resolución espacio-temporal de 32 km y 25 años.

⁹ <http://www.emc.ncep.noaa.gov/nmb/irean/> [consulta: 22-06-11]

^[10] <http://nomads.ncdc.noaa.gov/index.php> Última actualización 11/3/10
Creada por NOMADS.ncdc@noaa.gov [consulta: 22-06-11]

CAPÍTULO 2

BASE DE DATOS DE REANÁLISIS

Para la descarga de los datos de reanálisis, se efectuó en una plataforma GNU/Linux, en este caso Fedora 14, para ello, se tiene que preparar la descarga de datos dando selección al rango de datos y hora, esta última es la hora registrada secuencialmente para cada día y están disponibles cada tres horas (0000 0300 0600 0900 1200 1500 1800 2100); ejemplo:

Desde: Enero 01 de 1979 a las 00:00 UTC¹¹

Hasta: Marzo 28 de 2012 a las 00:00 UTC (esta última varía de acuerdo a la fecha en la que nos encontremos en la actualidad [consulta: 01-04-12]).

Se deben seleccionar los niveles de altura geopotencial (**ver Anexo B1**) y las variables¹² de interés. El nivel de altura geopotencial considerada, es a 2 metros sobre la superficie, lo que corresponde a la altura a la que se ubica una estación meteorológica.

Posteriormente, se extrae una subregión de los datos de reanálisis, ingresando los intervalos de longitud, ψ , y latitud, ϕ , de interés, por ejemplo las coordenadas que corresponden a la República Mexicana son aproximadamente, longitud: -120 W, -80 E; latitud: 10 S, 40 N.

Desafortunadamente, una variable no se tiene presente en esta base mundial, y es la **pcp**, por lo que la BD de reanálisis, serán de las variables de temperaturas mínima y máxima, las cuales, se encuentran en grados Kelvin (K), por lo que se hace una conversión a grados Celsius ($^{\circ}C$) debido a que los datos observados están en ésta escala:

$$t_c [^{\circ}C] = T [K] - 273.15 \quad (2)$$

Recordando que se tienen datos de reanálisis disponibles desde el año de 1979 a la fecha. Al extraer estos **datos teóricos**, hacemos un análisis de sus ST respecto a los **datos observados**.

¹¹ Tiempo universal coordinado, o UTC, es el tiempo de la zona horaria de referencia respecto a la cual se calculan todas las otras zonas del mundo. Sucesor del GMT (Greenwich Meridian Time: tiempo promedio del Observatorio de Greenwich, en Londres) aunque todavía coloquialmente algunas veces se le denomina así.

¹² Las claves de las variables y los niveles de altura geopotencial, pueden encontrarse en el documento merged_land_AWIP32, el cual está en formato pdf y puede descargarse de la dirección: http://www.emc.ncep.noaa.gov/mmb/treanl/merged_land_AWIP32.pdf [consulta 01-04-12].

CAPÍTULO 2

BASE DE DATOS DE REANÁLISIS

A continuación se muestran sus respectivas gráficas a una escala diaria, por año, donde se muestra en la parte superior su año, seguida de doce datos (ver figura 2.1, 2.2, 2.3, 2.5, 2.6 y 2.7), éstos, son episodios fríos y calientes de la temporada (temperaturas medias superficiales del océano, mensuales); son datos históricos que describen el evento climático **El Niño/ La Niña**.

Esta nueva BD se conoce como *analysis-monitoring* (*análisis de monitoreo* o simplemente *monitoring*), El Niño / Southern Oscillation (ENSO¹³), la cual se encuentra disponible en la página electrónica de la NOAA National Weather Service - Climate Prediction Center¹⁴. Con el *monitoring*, podemos conocer, para cada ST anual, la fase: **año NIÑO**, **año NIÑA**, o **neutro**; para No perder de vista lo que ocurre físicamente dentro de los eventos climáticos, que afecten tanto un nivel regional (Lo que sería la República Mexicana) a un nivel local (Edo. de Tlaxcala).

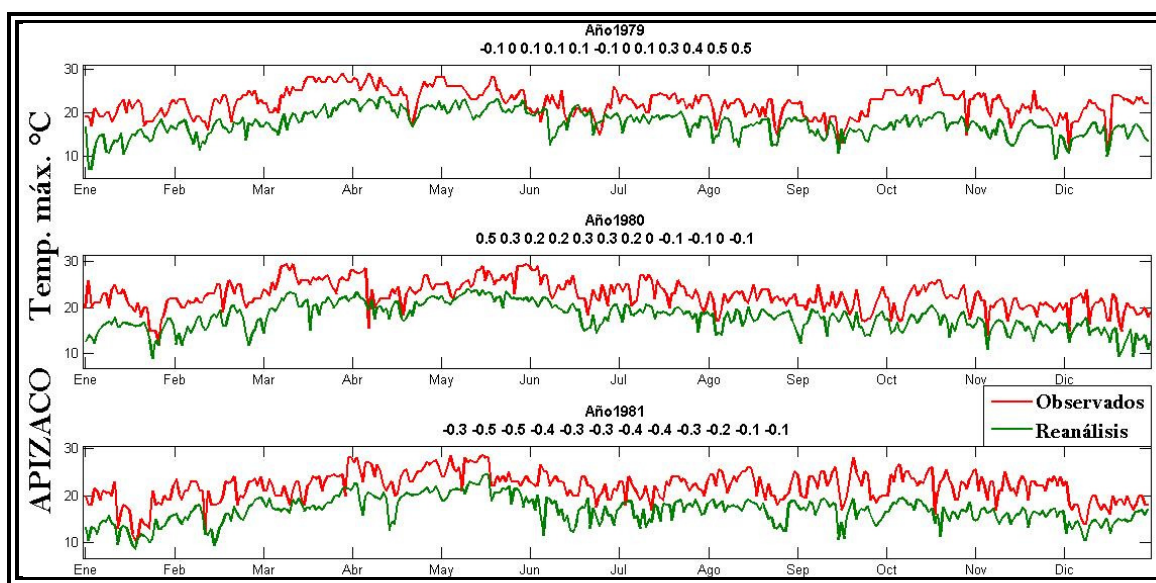


Figura 2.1 ST observados y teóricos, Apizaco-temp.máx.

En el anexo B2 se muestran las gráficas restantes de las ST observadas y reanálisis, a escala diaria, para las tres estaciones meteorológicas, teniendo la BD teóricos desde 1979 hasta el 2001 en ambas temperaturas.

¹³ El Niño, de acuerdo con la NOAA, se caracteriza por el calentamiento inusual de las capas superficiales del océano en el Pacífico Ecuatorial del este y está relacionado con una oscilación atmosférica global llamada la Oscilación del Sur, la cual altera la presión atmosférica entre el hemisferio occidental y oriental.

¹⁴ http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml [última consulta 03-04-12]
Page Author: Climate Prediction Center Internet Team. Page last modified: March 5, 2012

CAPÍTULO 2

BASE DE DATOS DE REANÁLISIS

Esta BD del *monitoring*, tiene contemplado desde el año de 1950 a la fecha. Para los fines históricos que se describen en este sitio web, son episodios fríos y calientes (números de color azul y rojo, respectivamente), lo cual significa que hubo **NIÑA**, por lo que el océano se encontró con temperaturas medias superficiales menores a -0.5°C , con cinco valores consecutivos para ambas fases, o **NIÑO**, con temperaturas superiores de $+0.5^{\circ}\text{C}$, los episodios de color negro son **neutros**.

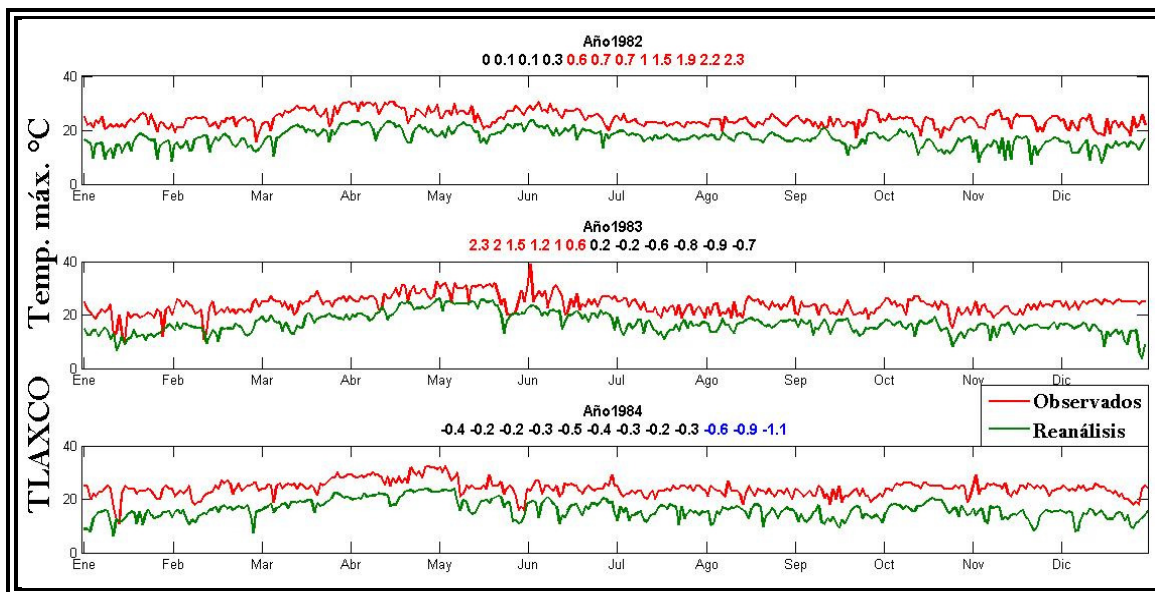


Figura 2.2 ST observados y teóricos, Tlaxco-temp.máx.

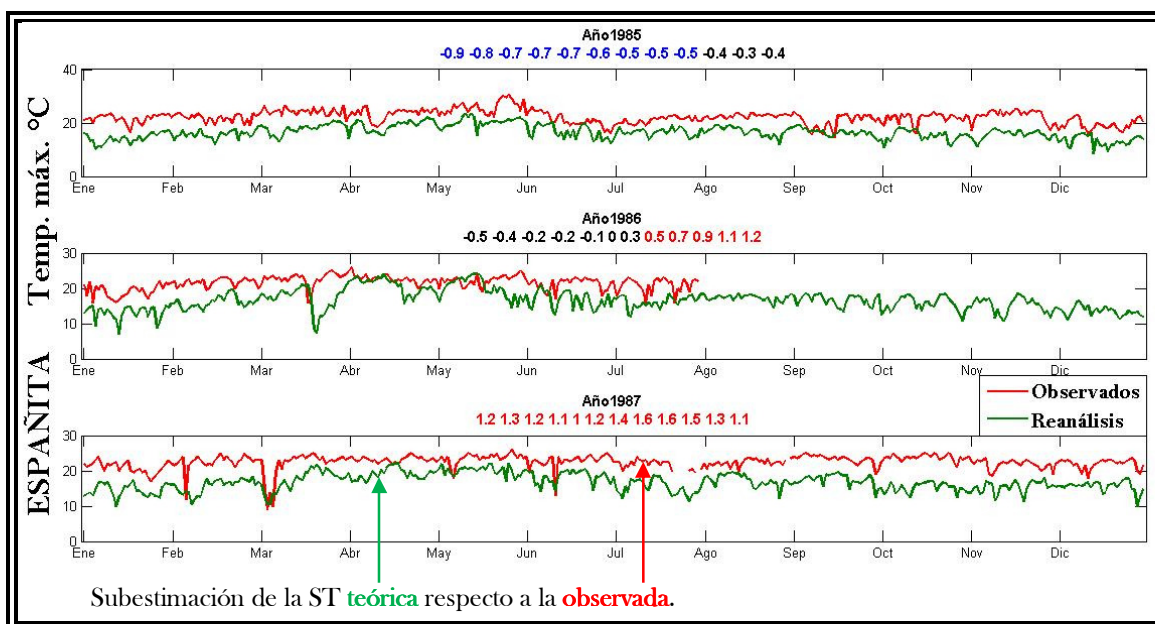


Figura 2.3 ST observados y teóricos, Españita-temp.máx.

CAPÍTULO 2

BASE DE DATOS DE REANÁLISIS

Se tiene, para las gráficas de **temp.máx.** mostradas en las figuras 2.1, 2.2 y 2.3, las **teóricas** (color verde) están **subestimadas** con respecto a las **observadas** (color rojo).

Una primera explicación de la **subestimación** es: debido a que las ST teóricas son de un nivel regional, la cual, se construye al realizar un promedio respecto a los valores de los puntos, o estaciones que caractericen la zona, formando así una cuadrícula o malla, donde en cada esquina de ésta, se encuentran las estaciones (ver figura 2.4).

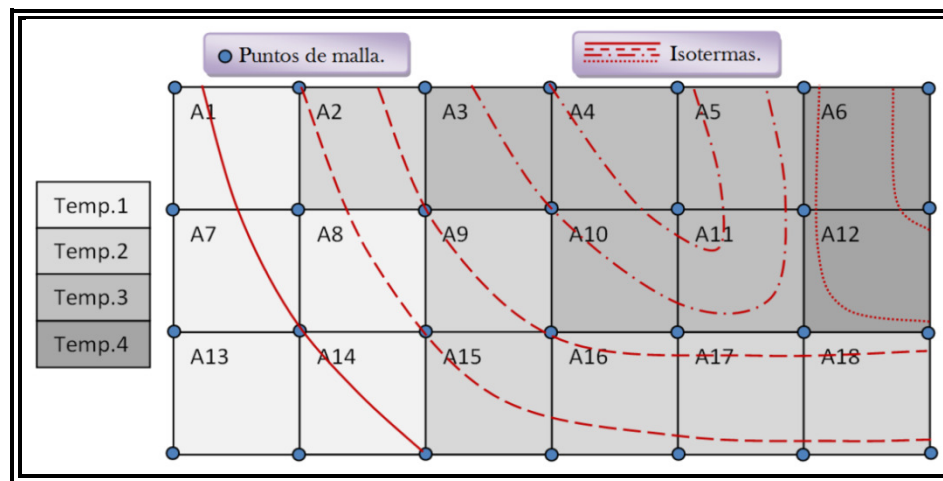


Figura 2.4 Región de dieciocho áreas, con cuatro temperaturas representativas.

Ejemplo, se podría hacer un gradiente o malla de isotermas (líneas de misma temperatura) a un nivel geopotencial de altura ***h***, con una distancia ***x*** entre puntos de malla, los valores de cada punto son promediadas para cubrir una área y formar así el gradiente (valor representativo de cada una de las áreas que construye la malla), pero, si quisiéramos saber el valor de un sólo punto en específico dentro de un área, nos arrojaría el valor que esté más cercano a un punto que lo conforma.

El ejemplo previo, lo encontramos al descargar los datos de reanálisis para la estación meteorológica de Atlangatepec, al norte de Apizaco y al sur-oeste de Tlaxco; ingresando las coordenadas que se tiene localizada la estación, nos arrojó valores idénticos a los correspondientes que se tienen para Apizaco, por lo que se descarto la estación de Atlangatepec, tomando así la estación meteorológica de Tlaxco.

Por lo que en un principio, se piensa que las ST teóricas están *subestimadas* respecto a las observadas debido al enfoque de un nivel regional a un punto local.

CAPÍTULO 2

BASE DE DATOS DE REANÁLISIS

Otra explicación de la *subestimación* del reanálisis respecto a las observadas es: la temperatura es la cantidad de energía, y para el caso de **temp.máx.**, es cuando hay **mayor** cantidad de energía, por lo que se tiene que las ST observadas se encuentren *sobrestimadas*, respecto a las teóricas, y tomando en cuenta que la de reanálisis son series promediadas para un nivel regional.

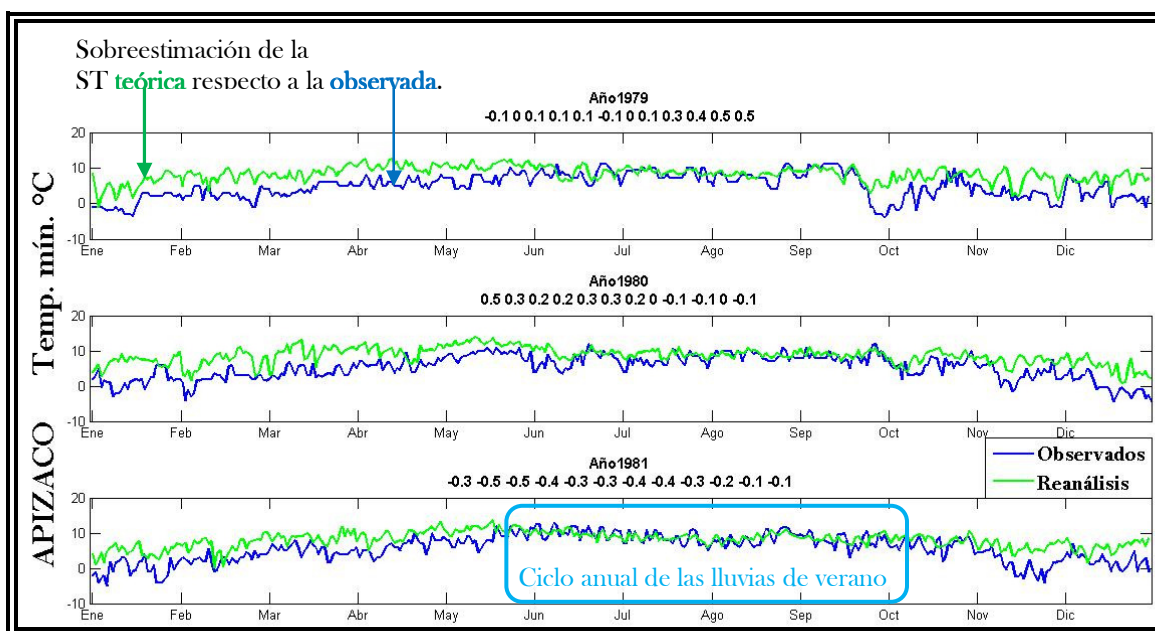


Figura 2.5 ST observados y teóricos, Apizaco-temp.mín.

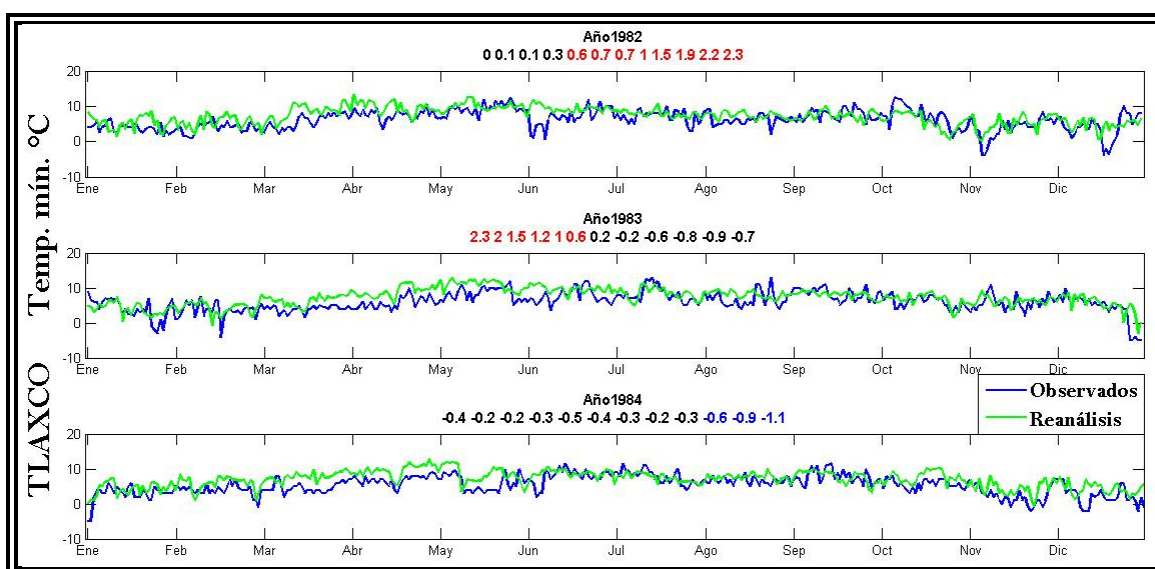


Figura 2.6 ST observados y teóricos, Tlaxco-temp.mín.

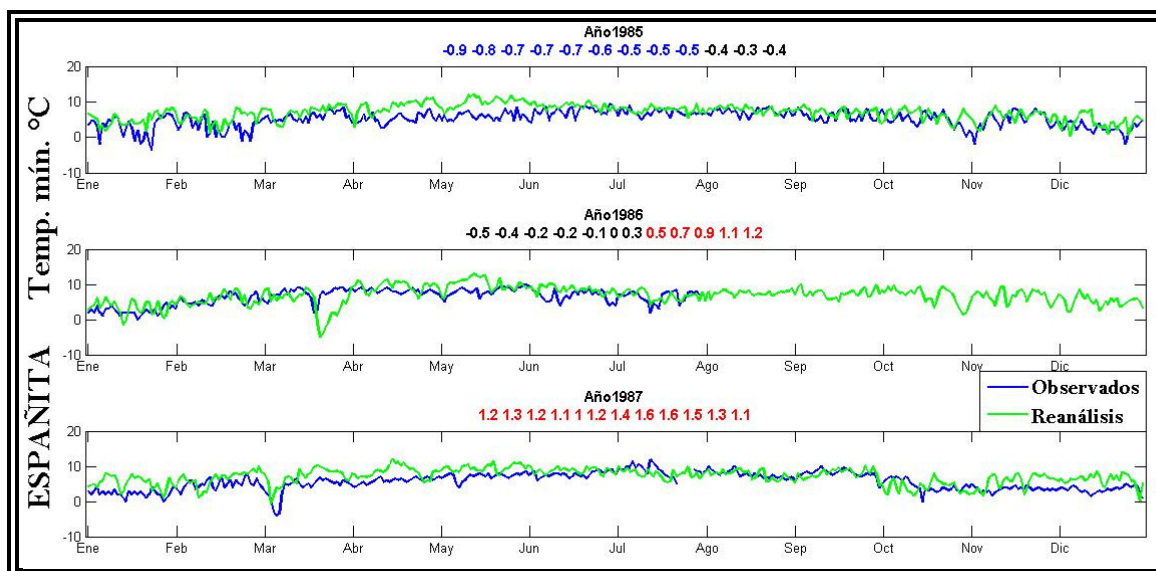


Figura 2.7 ST observados y teóricos, Espanita-temp.mín.

Para la ST de **temp.mín.**, figuras 2.5, 2.6 y 2.7, tenemos lo contrario a la observación descrita de la **temp.máx.** Se tiene que la de reanálisis (*color azul*), se encuentran **sobreestimadas**, respecto a las ST observadas (*color verde*). Se menciona que, la temperatura es la cantidad de energía, y para la **temp.mín.**, es la **menor** cantidad de energía registrada en el transcurso del día; tal como es para la **temp.máx.**, pero siendo ésta la mayor cantidad de energía. Por ello, se tiene que la ST teórica está *sobreestimada* respecto a la observada para la **temp.mín.**, teniendo en cuenta también, que la de reanálisis es una serie promediada de los puntos de malla que la describen.

Para las dos restantes estaciones meteorológicas, se tiene el mismo comportamiento descrito, una *subestimación* de las ST teóricas respecto a las observadas para la **temp.máx.**, y una *sobreestimación*, con la misma relación, pero de la **temp.mín.**.

Lo que se puede observar, de estas gráficas de punto de las ST para la **temp.mín.**, es que independientemente de cualquier estación meteorológica, se presenta una *sobreestimación* de las teóricas respecto a las observadas. También se puede observar un mejor ajuste respecto al de **temp.máx.** (es decir, menor distancia entre las ST de reanálisis y las de las estaciones meteorológicas). Es importante señalar que en algunas gráficas, sin importar la estación, el ajuste mejora para los periodos de Junio a Octubre, aproximadamente. Este

comportamiento se debe a la presencia de lluvias, que provocan una baja en la temperatura de la superficie terrestre, y de esta forma es registrado por las estaciones.

El periodo referido previamente, forma parte del ciclo anual de las lluvias de verano, en la región centro-sur de México, en la cual aparecen dos máximos en la **pcp** de verano, uno en Junio y otro en Septiembre, por tanto, existe un mínimo relativo entre Julio y Agosto, conocido como **canícula**¹⁵ [Magaña Rueda et al., 1999]. Ésta, se ve reflejada en las ST de **temp.mín.**, observando una subestimación de los datos teóricos respecto a los observados, un ejemplo, es en la ST del año de 1981 para la estación meteorológica de Españaita, presentando dicho comportamiento en el periodo de **canícula** (ver figura2.8).

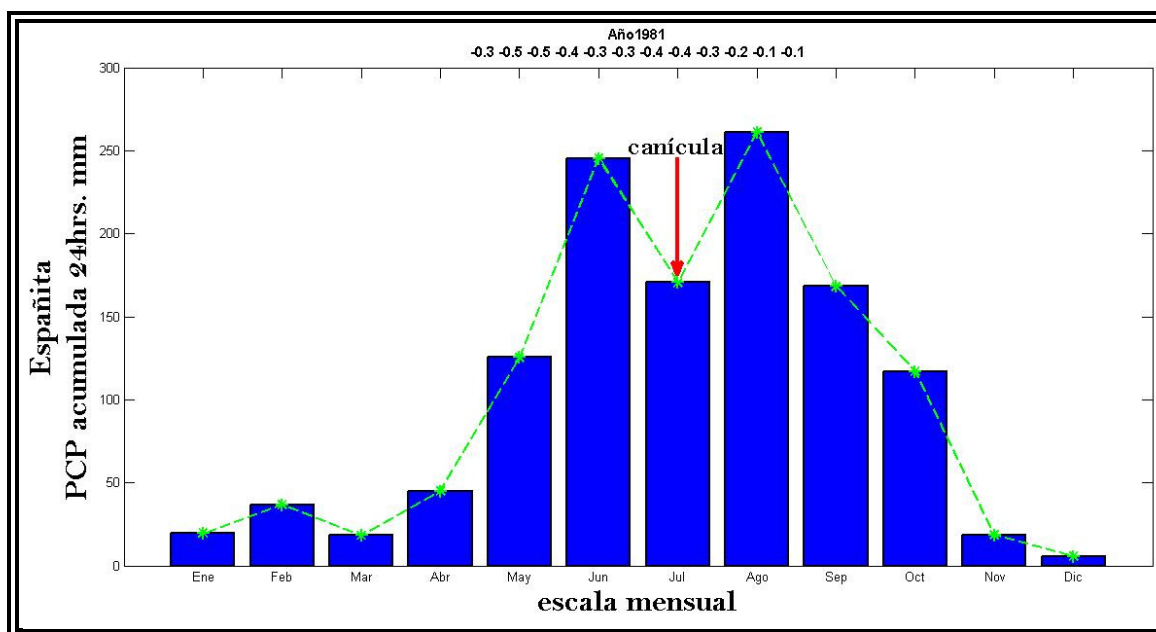


Figura 2.8 Periodo de canícula.

Al analizar las ST de los **datos teóricos** respecto a los **datos observados**, se observa una diferencia de magnitud para cada variable, por lo que se realizó una diferencia de distancias, o bien, una función coste, en este caso se obtuvo el error cuadrático medio (ver Anexo B3), por año, de cada estación y variable.

¹⁵ La **canícula** corresponde a una disminución en las lluvias durante Julio y Agosto, provocando el aumento de la temperatura. Este periodo canicular, no se presenta en todas las regiones de México [Magaña Rueda et al., 1999].

CAPÍTULO 2

BASE DE DATOS DE REANÁLISIS

Error Cuadrático Medio						
Año	Apizaco		Tlaxco		Españita	
	temp.máx.	temp.mín.	temp.máx.	temp.mín.	temp.máx.	temp.mín.
1979	129913	70754	160764	13194	956527	65254
1980	101370	66757	151181	13357	359310	13872
1981	139582	44735	205355	7369	104544	2244
1982	196246	158848	437828	19874	119335	5168
1983	189800	61547	544723	14771	173727	4923
1984	444294	60209	685963	14785	234838	12695
1985	165244	77741	NaN	NaN	220620	18725
1986	NaN	NaN	NaN	NaN	NaN	NaN
1987	133739	NaN	621847	NaN	NaN	NaN
1988	144381	88875	589324	15170	NaN	NaN
1989	NaN	103901	NaN	NaN	NaN	NaN
1990	204923	55030	694208	10923	NaN	22769
1991	NaN	NaN	541409	NaN	NaN	NaN
1992	132426	51142	NaN	8687	NaN	NaN
1993	193581	101571	691582	11538	162453	NaN
1994	NaN	101658	481216	10883	NaN	NaN
1995	148537	89025	521081	12659	71563	5862
1996	155241	128762	598128	23001	NaN	NaN
1997	185056	83950	896015	14217	134794	12699
1998	258023	129128	912116	14467	180014	7030
1999	234316	129308	751509	15839	130236	13908
2000	232616	94342	870969	30065	161760	5897
2001	106458	61853	441970	11613	95866	14172

Tabla 2.1 Función coste de los datos de reanálisis, con respecto de los datos de las estaciones meteorológicas.

Se utilizó el error cuadrático medio, debido a que esta función coste, será empleada para la metodología que se describe en el siguiente capítulo, y así tener un registro estadístico de errores, bajo una misma función.

Como se observa en tabla 2.1, hay años donde se tienen NaN¹⁶, esto se explica porque son las ST que están incompletas, recordando que sólo se tienen datos faltantes en los datos observados, descrito en el capítulo anterior (apartado 1.4.2).

Este tipo de análisis, es para tener una estadística de errores de las ST de los *datos teóricos* respecto a los *datos observados*.

Los “errores”, mostrados en la tabla 2.1, son demasiado grandes, por lo que se pensaría, ¿De qué nos sirve la BD de reanálisis? Como bien se mencionó, el buscar datos reconocidos mundialmente, es para poder obtener una relación, desde un punto de vista regional (punto de malla), a un punto local (estación meteorológica), lo cual se ha conseguido; y es de esperar

¹⁶ Not-a-Number. Para nuestro caso, NaN se utiliza cuando hay ausencia de datos. Se hace uso del software comercial, MatLab (Matrix Laboratory), y un NaN se obtiene matemáticamente como resultado de operaciones indefinidas como: $0.0/0.0$ e $inf - inf$. Al realizar cualquier operación con un valor o cifra respecto a un NaN, o viceversa, da como resultado NaN.

CAPÍTULO 2

BASE DE DATOS DE REANÁLISIS

la diferencia de magnitudes, debido, a que estamos pasando de algo general a algo más detallado, pues la función de error lo confirma y se puede apreciar en un ejemplo ilustrativo como es la línea de una costa real; a medida que aumenta la resolución, surgen más entrantes y salientes, como bahías y cabos, por lo que la longitud aumenta, tenemos un “proceso que depende de la escala” (ver figura 2.9).



Figura 2.9 Costa británica, tomado de Montesdeoca P. (2005)

Pero, ¿realmente soluciona nuestra problemática esta base de **datos teóricos**? En un principio, basándonos con la tabla 2.1, pensaríamos que no, por eso, se hace el análisis de ST y se grafican (primer paso y más importante). Es de apreciar, gráficamente, que el comportamiento, o bien, la tendencia, se asemejan en ambas bases de datos, lo cual es de gran utilidad, debido a que éstas se presentan en los elementos de algunas ST, que son los **valores extremos**, aunque no son de la misma magnitud, pero se tiene este comportamiento.

Las ST teóricas, no solucionan por completo el problema que se tiene, pero **sí**, nos podrán ayudar a generar los datos faltantes de las ST observadas. Esto, se realiza mediante la metodología descrita a continuación, para posteriormente pasar al capítulo 4, que son los experimentos numéricos realizados con el método propuesto.

METODOLOGÍA

REDES NEURONALES ARTIFICIALES

3.1 INTRODUCCIÓN.

Las Redes Neuronales Artificiales (**RNA**), son la implementación de modelos matemáticos idealizados de la función de las neuronas biológicas. Éstos son viables para resolver una amplia variedad de problemas, tales como la clasificación, estimación y optimización de patrones; reconocimientos de visión, del habla y de caracteres; funciones de aproximación, **modelado de sistemas no lineales**, robótica, control, procesamiento de señal, predicción, economía, defensa, bioingeniería, etc. [Martín del Brío, 2007; Hilera González, 1995].

Una **RNA** es un sistema de procesamiento de información que está formado por elementos “**no-lineales**” llamados neuronas. Estas neuronas están interconectadas con ellas mismas y son distribuidas en capas, de tal forma que emulan en forma simple la estructura/función neuronal del cerebro.

Las **RNA** intentan emular de forma simplificada el funcionamiento del cerebro humano [Corchado J.M., 2000], no tratan de reproducirlo idénticamente, sino que se centran en mecanismos de resolución a problemas individuales. Por ejemplo, tratan de imitar el proceso de almacenar información en patrones y utilizarla para resolver cierto tipo de problemas.

Para cada modelo de neurona, es capaz de realizar algún tipo de procesamiento a partir de estímulos de *entrada (input)* y ofrecer una *respuesta (salida-output)*, por lo que las **RNA**, en conjunto, funcionan como redes distribuidas de cálculo en paralelo, similares a los sistemas cerebrales biológicos. Sin embargo, a diferencia de las computadoras convencionales, las cuales son programas para realizar tareas específicas, las **RNA**, tal como los sistemas biológicos, deben ser *entrenados*. Un atributo significativo de las neuronas es su habilidad

CAPÍTULO 3. METODOLÓGÍA. REDES NEURONALES ARTIFICIALES

para aprender, ya sea interactuando con el medio en el que se encuentran o bien con información de entrada.

Los primeros modeladores de las neuronas artificiales estuvieron interesados, no solamente en las relaciones *estímulo-respuesta* de las neuronas biológicas, sino también, en la representación de la estructura interna de la neurona. Sin embargo, fueron las propiedades computacionales de las neuronas artificiales lo que atrajo la atención de los científicos en varias otras disciplinas.

El creciente interés en las **RNA** fue, es y estará motivado por el deseo de construir una nueva clase de computadoras y de algoritmos más potentes para resolver problemas que han demostrado ser difícilmente manejables con las herramientas tradicionales de cálculo y procesamiento de la información. Esas tareas son análogas a aquellas realizadas típicamente, en forma natural y rápidamente por el cerebro humano, sin embargo, ellas están mucho más allá del alcance de las computadoras convencionales y de los denominados *Sistemas Expertos*¹⁷ (**SE**).

Por ejemplo, un humano puede normalmente reconocer una cara familiar en aproximadamente 200 milisegundos; el ojo humano puede ajustar los niveles de la intensidad de la luz en 7 niveles diferentes de magnitud; ningún sistema de procesador de imágenes puede aproximarse siquiera al desempeño de este sistema humano.

El trabajo hecho en el campo de las **RNA**, fue motivado inicialmente por el interés de entender la forma en la que trabaja el cerebro de cualquier ser vivo, pues es completamente diferente a la forma en la que trabajan las computadoras. Por ello, es de gran y suma importancia, que para entender el interés por la **RNA**, es útil y necesario, el entender los fundamentos de los conceptos neurobiológicos.

17 Sistemas Expertos. Son llamados así porque emulan el comportamiento de un experto en un dominio concreto. Se busca una mejor calidad y rapidez en las respuestas dando así lugar a una mejora de la productividad del experto (**ver Anexo C1**).

3.2 REDES NEURONALES BIOLÓGICAS (RNB).

Las **RNB** están inspiradas en la estructura y funcionamiento de los sistemas nerviosos, donde la neurona biológica es el elemento fundamental.

El cerebro humano está constituido por un arreglo complejo de neuronas; se estima que el sistema nervioso contiene alrededor de “**cien mil millones de neuronas**” que dan lugar a un **sistema cognitivo robusto**, con un diseño distinto al de las computadoras. Vistas al microscopio, este tipo de células puede presentarse en múltiples formas, aunque muchas de ellas presentan un aspecto similar muy peculiar (ver figura 3.1), con un cuerpo celular o **soma** (de entre 10 y 80 micras de longitud), que es donde se localiza el núcleo, del que surge un denso árbol de ramificaciones (árbol dendrítico compuesto por dendritas); las **dendritas** es por donde las neuronas reciben o atrapan las señales provenientes de las otras neuronas, y del cual parte una fibra tubular denominada **axón** (cuya longitud varía desde las 100 micras hasta el metro, en el caso de las neuronas motoras que es lo que constituye al nervio), ésta se extiende desde el cuerpo de la célula hasta eventualmente ramificarse en filamentos y sub-filamentos, que también se ramifica en su extremo final para conectar con otras neuronas (cada neurona puede conectarse con otras diez mil neuronas, en promedio).

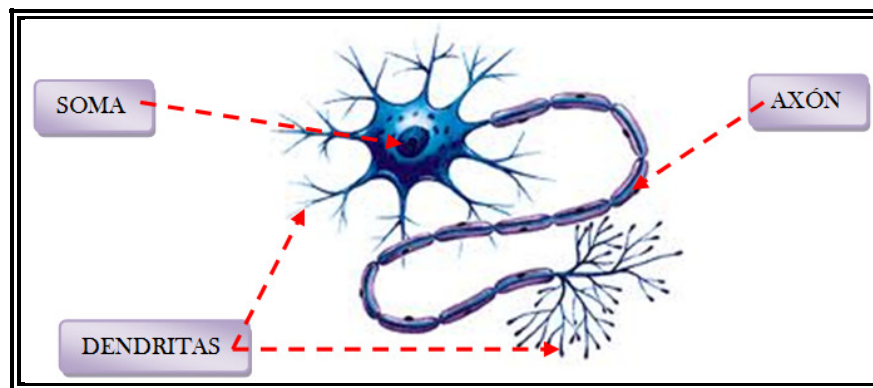


Figura 3.1 Estructura General de una Neurona Biológica Típica ¹⁸

En los extremos de las ramificaciones del axón se encuentran las uniones sinápticas o **sinapsis** y representa el aprendizaje adquirido hacia las otras neuronas. Las uniones sinápticas juegan el doble papel de terminaciones *transmisoras* en una célula y terminaciones

¹⁸ Figura modificada toma de: La Neurona - Estructura, clasificación.
<http://psicologiavprado.blogspot.mx/> [consulta: 21-01-12]

CAPÍTULO 3. METODOLÓGÍA. REDES NEURONALES ARTIFICIALES

dendritas como en el propio cuerpo de la célula. El axón de una neurona típica realiza algunos miles de sinapsis con otras neuronas.

En 1888 Ramón Cajal¹⁹ demostró que el sistema nervioso estaba compuesto por una red de células individuales, “las neuronas”, ampliamente interconectadas entre sí. No sólo se observó al microscopio los pequeños vacíos que separaban unas neuronas de otras, sino que también estableció que la información fluye en la neurona desde las dendritas hacia el axón, atravesando el soma.

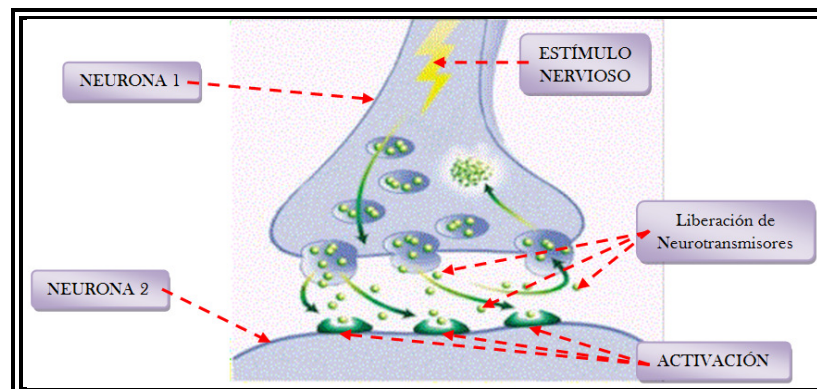


Figura 3.2 Transmisión de una señal (Sinapsis), de la Neurona 1 a la Neurona 2 ²⁰

La transmisión de una señal (típicamente en forma de un tren de pulsos) de una célula a otra a través de la sinapsis, se realiza mediante un proceso químico complejo, en la que algunas sustancias específicas²¹ favorecen la transmisión, y son emitidas del lado de la neurona emisora con el objetivo de incrementar o disminuir el potencial eléctrico dentro del cuerpo de la célula receptora. Si este potencial alcanza un umbral determinado, un pulso de duración e intensidad fija es enviado a través del axón. Es entonces cuando se dice que la neurona ha disparado. El pulso se ramifica a través de las arborizaciones del axón hasta alcanzar las uniones sinápticas con las otras células.

¹⁹ Santiago Ramón Cajal. Científico aragonés el cual inicia la historia de las redes neuronales artificiales, descubridor de la estructura neuronal del sistema nervioso.
http://www.enciclopedia-aragonesa.com/monograficos/biografias/santiago_ramon_y_cajal/default.asp [consulta: 21-01-12].

²⁰ Figura modificada toma de: Sinapsis - las neuronas se organizan en redes y sistemas.
<http://iibce.edu.uy> [consulta: 21-01-12].

²¹ La neurona presináptica libera unas sustancias químicas complejas denominadas neurotransmisores como el glutamato o la adrenalina. El glutamato es el neurotransmisor excitatorio por excelencia de la corteza cerebral humana.

CAPÍTULO 3. METODOLÓGÍA. REDES NEURONALES ARTIFICIALES

Después de disparar la célula tiene que esperar un intervalo de tiempo conocido como periodo de refracción antes de que pueda volver a disparar otra vez. Cabe aclarar que una sinapsis, en realidad, no es una conexión física (ver figura 3.2). En resumen, la sinapsis se convierte de una señal eléctrica a una señal química y luego nuevamente en una señal eléctrica, o de forma más simple, una sinapsis es una conexión que puede **excitar** o **inhibir** a otra neurona, pero no ambas.

Las Redes Neuronales Biológicas (RNB) pueden establecerse como grupos de neuronas activas especializadas en tareas como: cálculos matemáticos, posicionamiento y memoria [Simpson P.K., 1990].

3.3 COMPARACIÓN DE LAS '*RNB*' CON LAS '*RNA*'.

Como se menciona anteriormente, el campo de las **RNA** fue motivado por entender la forma en la que trabaja el cerebro humano así como la de cualquier ser vivo, a diferencia en la forma que trabajan las computadoras.

Se sabe que el cerebro está compuesto de neuronas, mientras que en una computadora se tienen componentes electrónicos como compuertas lógicas. Ahora, si comparamos el tiempo que dura un evento en una compuerta de silicón ($\sim 10^{-9}$ s) con la duración de un evento en una neurona ($\sim 10^{-5}$ s), vemos que hay una diferencia entre cinco y seis órdenes de magnitud. Sin embargo, el cerebro compensa esta falta de rapidez, teniendo una enorme cantidad de neuronas, junto con una cantidad aún más impresionante de conexiones o sinapsis entre ellas.

Mientras el cerebro es capaz de procesar la información que recibe a través de los ojos, para que logremos interactuar con nuestro entorno, o reconocer caras familiares en un ambiente cualquiera, una computadora puede tomar días en cálculos mucho menos complejos. Además, el cerebro resulta ser una máquina mucho más eficiente energéticamente hablando, usando alrededor de 10^{-16} J/s por operación, mientras que muchas computadoras gastan alrededor de 10^{-6} J/s por operación. En conclusión, si quisiéramos tener una computadora

CAPÍTULO 3. METODOLÓGÍA. REDES NEURONALES ARTIFICIALES

comparable con el cerebro, tendría que ser una máquina extremadamente *compleja, no lineal y capaz de operar en paralelo*²² [Simon Haykin, 1994] (ver Anexo C2).

Una **RNA** debe asemejarse al cerebro en dos aspectos, el conocimiento debe ser adquirido mediante un proceso de aprendizaje y debe usar las conexiones neuronales (sinapsis) para almacenar la información, para después hacerla o tenerla disponible para algún uso.

La **RNA** está compuesta por elementos simples *operando en paralelo*. Esta idea fue inspirada en las **RNB**, en las que se realiza algún aprendizaje al crear distintas conexiones entre los elementos que lo conforman. Al procedimiento usado para llevar a cabo el proceso de aprendizaje se llama *algoritmo de aprendizaje*, que consiste en simplemente modificar los pesos de las conexiones sinápticas, de igual manera, en las **RNA** se ajustan las conexiones (pesos) entre sus neuronas, permitiéndonos hacer que dada una entrada (input) la red esté entrenada para dar cierto tipo de salida (output).

3.4 LA NEURONA ARTIFICIAL.

3.4.1. Modelo General.

La respuesta de las neuronas biológicas es de tipo “*no lineal*”, característica que es emulada en las **RNA**. La formulación de la neurona artificial, como dispositivo *no lineal*, constituye de las características más destacables y una de las que proporciona un mayor interés de las **RNA**, debido al tratamiento de problemas altamente *no lineales*, el cual no suele ser fácil de abordar mediante las técnicas convencionales.

En un modelo general de la neurona artificial, se le denomina como un dispositivo simple de cálculo y que, a partir de un vector de entrada (input) procedente del exterior o de otras neuronas, proporciona una única respuesta o salida (output). El modelo de la neurona formal se inspira en la operación de la neurona biológica, en el sentido de integrar una serie de entradas y proporcionar una cierta respuesta, que se propagan por el axón.

²² Se basa en el principio de que los problemas grandes se pueden dividir en partes más pequeñas, llevándolas a cabo a un mismo tiempo.

Sea una neurona i , los elementos que la constituyen son los siguientes:

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

De este modo, la operación de la neurona i puede expresarse como:

$$y_i(t) = F_i \left(f_i \left[a_i(t-1), \sigma_i \left(w_{ij}, x_j(t) \right) \right] \right) \quad (3)$$

(ver figura 3.3)

3.4.2 Modelo Estándar.

La neurona artificial estándar consiste en:

- Un conjunto de entradas $x_j(t)$ y pesos sinápticos w_{ij} .
- Una regla de propagación $h_i(t) = \sigma(w_{ij}, x_j(t))$; $h_i(t) = \sum w_{ij} x_j$. Con frecuencia se añade al conjunto de pesos de la neurona un parámetro adicional θ_i , al que se denomina umbral²³,

²³ En algunos modelos se le es llamado *threshold*, es decir, “umbral”, y en otras *bias*, que no tiene una traducción clara para este caso, motivo por el que llamaremos umbral a este parámetro (aunque a veces no haga el papel de umbral) [Martín del Brío, 2007].

CAPÍTULO 3. METODOLÓGÍA. REDES NEURONALES ARTIFICIALES

que se resta del potencial postsináptico, por lo que el argumento de la función queda $\sum_j w_{ij}x_j - \theta_i$, lo que representa añadir un grado de libertad adicional a la neurona.

- Una función de activación $y_i(t) = f_i(h_i(t))$, que presenta simultáneamente la salida de la neurona y su estado de activación.

Para que los índices i, j comiencen en cero, se define que $w_{i0} \equiv \theta_i$ y $x_0 \equiv 1$ (cte), con lo que el potencial postsináptico se obtiene realizando la suma desde $j = 0$

$$y_i(t) = f_i\left(\sum_{j=0}^n w_{ij}x_j\right) \quad (4)$$

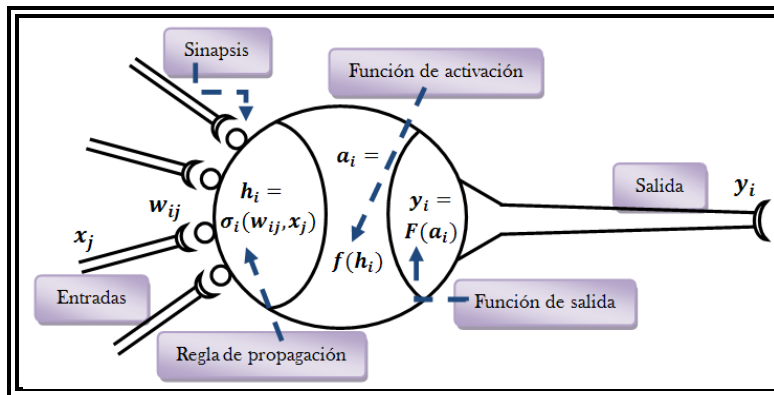


Figura 3.3 Modelo General de una Neurona Artificial (Rumelhart 1986²⁴)

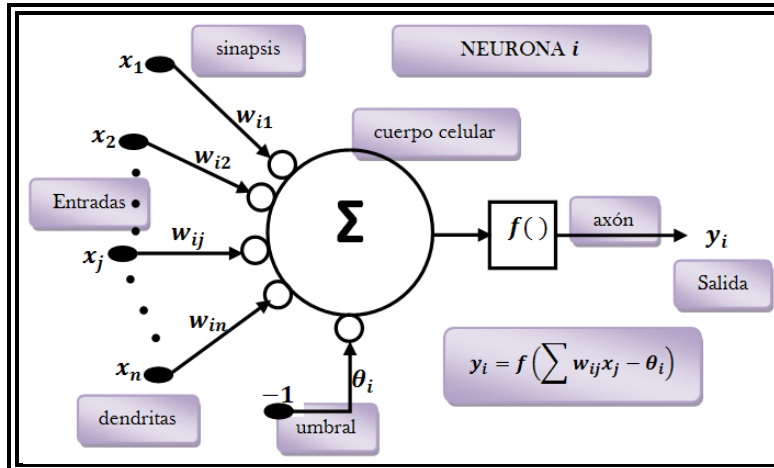


Figura 3.4 Modelo de una Neurona Estándar [modificada de Martín del Brío, 2001]

²⁴ David Rumelhart Everett (1942 - 2011). Psicólogo norteamericano que hizo muchas contribuciones al análisis formal de la cognición humana, que trabaja principalmente dentro de los marcos de la psicología matemática, la inteligencia artificial simbólica, y procesamiento distribuido en paralelo. En 1986, publica "Parallel distributed processing" (procesamiento distribuido en paralelo): Exploraciones en la microestructura de la cognición. Realizó dicho trabajo con James McClelland, el cual describió su creación de simulaciones por ordenador de la percepción, dando a los científicos de la computación sus primeros modelos comprobables del procesamiento neural, y que ahora es considerado como un texto central en el campo de la ciencia cognitiva.

3.4.3 Conceptos de la Estructura de una Neurona Artificial.

Entrada (input) – Salida (output).

Variables que pueden ser binarias (digitales) o continuas (analógicas), dependiendo del modelo y aplicación. Por ejemplo, un perceptrón multicapa MLP (Multilayer Perceptron) admite ambos tipos de señales. Así para tareas de clasificación poseería salidas digitales $\{0, +1\}$, mientras que para un problema de ajuste funcional de una aplicación multivariable continua, se utilizarán salidas continuas pertenecientes a un cierto intervalo.

Peso sináptico.

El peso sináptico w define en este caso la intensidad de interacción entre la neurona presináptica j y la postsináptica i (ver figura 3.5). Dada una entrada positiva (procedente de un sensor o simplemente la salida de otra neurona), si el “**peso es positivo**” tendrá a “**excitar**” a la neurona postsináptica, si el “**peso es negativo**” tendrá a “**inhibirla**”. Así se habla de *sinapsis excitadoras* (de peso positivo) e *inhibidoras* (de peso negativo).

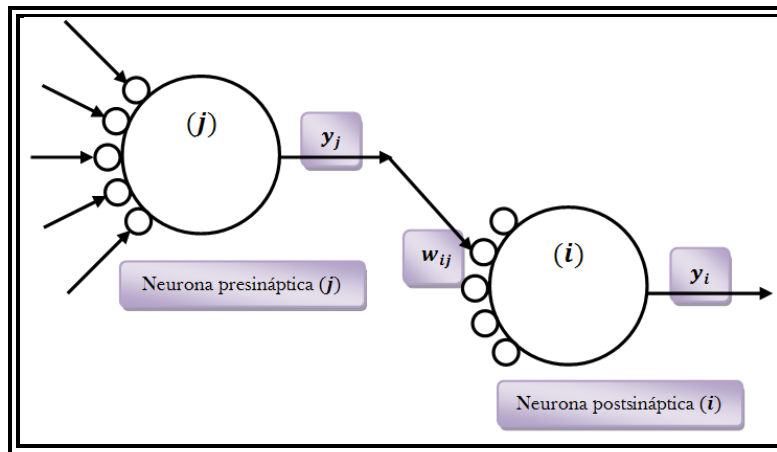


Figura 3.5 Interacción entre una neurona presináptica y otra postsináptica.
[modificada de Martín del Brío, 2001]

Regla de propagación.

Permite obtener, a partir de las *entradas* y los *pesos*, el valor del potencial postsináptico de la neurona:

$$h_i(t) = \sigma_i(w_{ij}, x_j(t)) \quad (5)$$

CAPÍTULO 3. METODOLÓGÍA. REDES NEURONALES ARTIFICIALES

La función más habitual es de tipo lineal, y se basa en la “*suma ponderada*” de las entradas con los pesos sinápticos:

$$h_i(t) = \sum_j w_{ij} x_j \quad (6)$$

que normalmente, también puede interpretarse como el producto escalar de los vectores de *entrada* y *pesos*

$$h_i(t) = \sum_j w_{ij} x_j = \mathbf{w}_i^T \cdot \mathbf{x} \quad (7)$$

Función de activación o de transferencia.

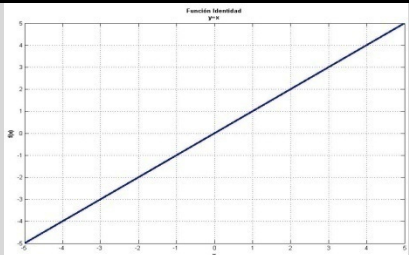
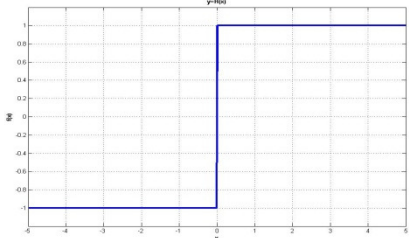
Esta función proporciona el estado de activación actual $\mathbf{a}_i(t)$ a partir del potencial postsináptico $\mathbf{h}_i(t)$ y del propio estado de activación anterior $\mathbf{a}_i(t - 1)$;

$$\mathbf{a}_i(t) = f_i(\mathbf{a}_i(t - 1), \mathbf{h}_i(t)) \quad (8)$$

Sin embargo, en muchos modelos de las **RNA** se considera que el estado actual de la neurona no depende de su estado anterior, sino únicamente del actual

$$\mathbf{a}_i(t) = f_i(\mathbf{h}_i(t)) \quad (9)$$

La función de activación se suele considerar determinista, y en mayor parte de los modelos es monótona creciente y continua, como se observa habitualmente en las neuronas biológicas. En ocasiones los algoritmos de aprendizaje requieren que la función de activación cumpla la condición de ser derivable. Las más empleadas en este sentido son las funciones de tipo sigmoideo, la gaussiana y funciones sinusoidales (ver tabla 3.1).

	Función	Rango	Gráfica
Identidad	$y = x$	$[-\infty, +\infty]$	
Escalón	$y = \text{signo}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$	

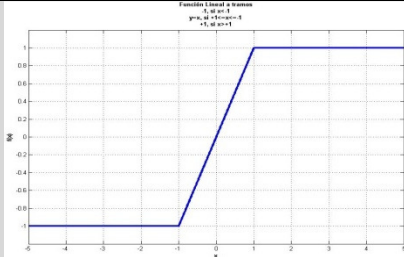
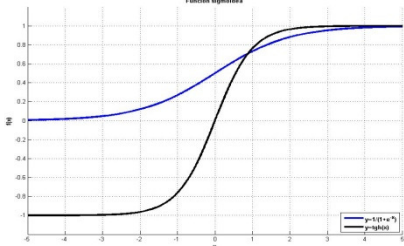
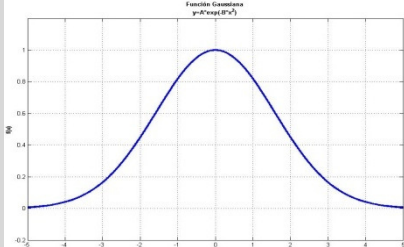
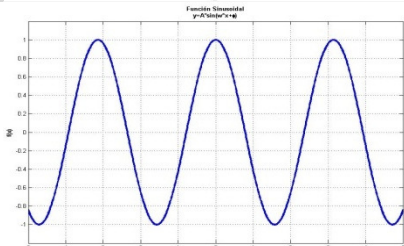
	Función	Rango	Gráfica
Lineal a tramos	$-1, si \ x < -l$ $y = x, si \ -l \leq x \leq +l$ $+1, si \ x > +l$	$[-1, +1]$	
Sigmoidea	$y = \frac{1}{1 + e^{-x}}$ $y = tgh(x)$	$[0, +1]$ $[-1, +1]$	
Gaussiana	$y = A \cdot e^{-Bx^2}$	$[0, +1]$	
Sinusoidal	$y = A \cdot sen(\omega x + \varphi)$	$[-1, +1]$	

Tabla 3.1 Funciones de activación.

El separar el concepto de regla de propagación y función de activación, permite considerar desde un punto de vista unificado varios modelos, que de otra manera, habría de tratar como casos especiales de una neurona estándar.

Función de salida.

Esta función proporciona la salida global de la neurona $y_i(t)$ en función de su estado de activación actual $a_i(t)$:

$$y_i(t) = F_i(a_i(t)) \quad (10)$$

(ver figura 3.5)

3.5 VENTAJAS DE LAS RNA

Las ventajas que ofrecen son las siguientes:

No linealidad.	El procesador neuronal es básicamente no lineal y, por consecuencia, la red neuronal también.
Transformación entrada (input) – salida (output).	El proceso de aprendizaje consiste básicamente en presentar a la red un ejemplo y modificar sus pesos sinápticos de acuerdo con su respuesta.
Adaptabilidad.	La red tiene la capacidad de adaptar sus parámetros, aun en tiempo real.
Tolerancia a fallas.	Debido a la interconexión masiva, la falla de un procesador no altera seriamente la operación.
Uniformidad en el análisis y diseño.	Esto permite garantizar características precisas.
Analogía con las redes biológicas.	Esto permite la utilización mutua del conocimiento de las dos áreas.

[Sánchez Camperos, 2006]

3.6 ARQUITECTURAS DE LAS REDES NEURONALES.

El número de neuronas, la disposición y el tipo de conexiones entre ellas determinan su estructura o patrón, denominada arquitectura o topología de la red neuronal. Los nodos de las **RNA** se conectan por medio de la sinapsis, esta estructura de conexiones sinápticas determina el comportamiento de la red.

Las conexiones sinápticas son direccionales, es decir, la información solamente puede propagarse en único sentido, desde la neurona presináptica a la neurona postsináptica (ver **Figura 3.5**). Por lo general, las neuronas se suelen agrupar en unidades estructurales que comúnmente son denominadas *capas*; y éstas a su vez, se pueden agrupar formando grupos neuronales llamados como *clusters*.

En conclusión, para el conjunto de una o más capas constituye la “**red neuronal**”. La **Figura 3.6** muestra la topología general de una red neuronal, formada por tres capas de neuronas, la capa de entrada o sensorial, la capa intermedia u oculta y la capa de salida.

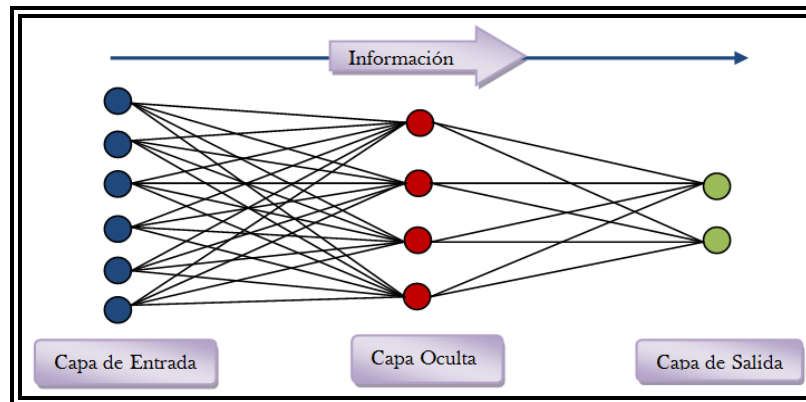


Figura 3.6 Arquitectura Unidireccional de Tres Capas de una RNA

La capa de entrada recibe datos o señales procedentes del exterior. La capa de salida es la que proporciona la respuesta de la red neuronal. La capa oculta, no tiene una conexión directa del exterior, recibe la información de la capa de entrada y proporciona el resultado que se envía a otra capa oculta o a la capa de salida.

Se pueden establecer diferentes tipos de arquitecturas neuronales, de acuerdo con el número de capas. Así, en relación a su estructura, se hablan de **redes monocapa**, son aquellas compuestas por una única capa de neuronas; y **redes multicapa** (*layered networks*), son aquellas cuyas neuronas se organizan en varias capas.

El tipo de conexiones o bien el flujo de datos en la red neuronal, se hablan de **redes unidireccionales** (*feedforward*), en las que su información circula en un único sentido, desde la neurona de entrada hacia la neurona de salida (ver figura 3.6); y redes **recurrentes** o **realimentadas** (*feedback*), cuya información puede circular entre las capas en cualquier sentido (ver figura 3.7). Esta última se da en aquellos casos en que la salida de una neurona de un nivel realimenta las neuronas de un nivel anterior [Corchado J.M., 2000].

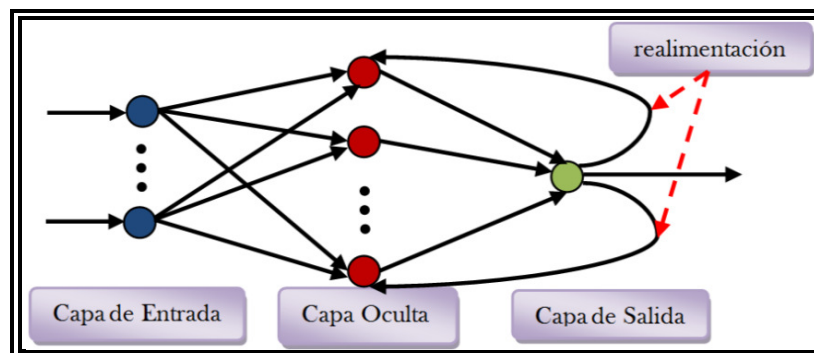


Figura 3.7 RNA con realimentación [modificada de Corchado J.M., 2000]

3.7 MECANISMO DE ENTRENAMIENTO (APRENDIZAJE).

Otro elemento de suma importancia de las **RNA**, es la **regla de aprendizaje**. Ésta es una de las características más importantes de una red neuronal y se puede definir como el proceso por el que se produce un ajuste de los parámetros libres de la misma, a partir de una estimulación por el entorno que le rodea.

El cerebro, es capaz de lograr aprender a realizar sus procesos de manera rutinaria, que es mediante la experiencia; biológicamente se suele aceptar que la información memorizada en el cerebro, está más relacionada con los valores sinápticos de las conexiones entre neuronas que con ellas mismas; es decir, el conocimiento se encuentra en la sinapsis.

En el caso de las **RNA**, se puede considerar que el conocimiento reside en los pesos de las conexiones entre neuronas; un peso representa la intensidad de asociación entre dos unidades, por lo que para la red, un peso representa qué tan frecuente una unidad receptora ha estado activa simultáneamente con la unidad emisora.

Para el proceso de aprendizaje (también reciben el nombre de algoritmos de aprendizaje o modo aprendizaje) implica un cierto número de cambios para determinar las conexiones entre neuronas, de esta forma, el cambio en el peso entre dos unidades depende de la frecuencia con que las dos unidades estén teniendo una salida positiva en forma simultánea.

En realidad, se puede decir que el aprendizaje consiste simplemente en determinar un conjunto de pesos sinápticos, modificando los valores de éstos, el cual permita que la red realice correctamente el tipo de procesamiento deseado.

Cuando se construye un sistema neuronal, se parte de un cierto modelo de neurona y de una determinada arquitectura de red, estableciéndose los pesos sinápticos iniciales como aleatorios.

El entrenamiento o aprendizaje más convencional es el de modelado de las sinapsis, que consiste en modificar o ajustar los pesos sinápticos siguiendo una cierta regla de aprendizaje, construida normalmente a partir de la optimización de una función de error, que mide la capacidad actual de la operación de red.

Ejemplo de un mecanismo de ajuste para los pesos sinápticos es el siguiente; si se denomina $w_{ij}(t)$ al peso que conecta la neurona presináptica j con la postsináptica i en la iteración t , el algoritmo de aprendizaje, en función de las señales que en el instante t llegan procedentes del entorno, proporcionará el valor $\Delta w_{ij}(t)$ que da la modificación que se debe incorporar en dicho peso y el cual quedará actualizado de la forma:

$$\Delta w_{ij}(t + 1) = w_{ij}(t) + \Delta w_{ij}(t) \quad (11)$$

El proceso de aprendizaje es usualmente iterativo, actualizándose los pesos de manera anterior, una y otra vez, hasta que la red neuronal alcanza el rendimiento deseado [Martín del Brío, 2001].

Se encuentran **dos tipos básicos de aprendizaje** que son el *supervisado* y el *no supervisado*. Ambos pretenden estimar funciones entrada/salida multivariable o densidades de probabilidad, la diferencia está en que el supervisado dispone de información sobre salida objetivo o deseada de la red, el cual se le proporciona cierta información sobre estas funciones (como la distribución de las clases, etiquetas de los patrones de entrada o salidas asociadas a cada patrón), hecho que no ocurre en el no supervisado.

3.7.1 Aprendizaje supervisado.

En el aprendizaje supervisado, cada patrón o dato de entrada que recibe la red, es comparado junto con la salida u objetivo deseado (dato específico), de tal manera que la red modifica los pesos iterativamente hasta que su salida tienda a ser la deseada. El proceder de este algoritmo, para cada paso de la fase de entrenamiento, se define una función objetivo o error a minimizar, el cual se actualizan los pesos hasta que el error entre la salida de la red y los datos establecidos se reduce.

3.7.2 Aprendizaje no supervisado o autoorganizado.

En este tipo de aprendizaje, involucra el agrupamiento o detección de similitudes entre elementos de un conjunto de entrenamiento a su entrada, lo cual presentan a la red una multitud de patrones sin adjuntar la respuesta que deseamos.

CAPÍTULO 3. METODOLÓGÍA. REDES NEURONALES ARTIFICIALES

La red, por medio de la regla de aprendizaje, modifica su comportamiento en base a mecanismos de la estimación de la función de densidad de probabilidad $p(\mathbf{x})$ que describe la distribución $\mathbf{x}$, pertenecientes al espacio de entrada $\mathbf{R}^n$ a partir de muestras (ejemplos), del cual pueden reconocerse regularidades en el conjunto de entradas, extraer rasgos, o agrupar patrones según su similitud, lo que determina la salida de la red.

3.7.3 Aprendizaje reforzado (reinforcement learning).

Para este tipo de aprendizaje, la red toma el error entre la entrada y la salida para maximizar los valores esperados de una función criterio, llamada *señal reforzadora* [Hassoon M.H., 1995] (solamente le indicamos lo bien o lo mal que está actuando, pero sin proporcionar más detalles), con este criterio la red se modifica actualizando los pesos de cada neurona así como su salida, repitiéndose el ciclo hasta obtener un error mínimo.

La similitud que tiene con el caso del no supervisado, es la de no suministrar explícitamente la salida deseada. En ocasiones se denomina *aprendizaje por premio-castigo*.

3.8 RED NEURONAL SUPERVISADA.

En este trabajo se tomará la red neuronal supervisada para los propósitos que se centra en esta tesis, que es la generación de datos faltantes de series de tiempo (ST).

¿Por qué una red neuronal supervisada? Se habla en el Capítulo 2 de una base de datos de reanálisis, el cual como se ha mencionado, las ST de esta base, tienen un comportamiento o bien una tendencia similar a la base de las ST observadas, la cual nos es de mucha ayuda para cumplir el objetivo primordial de este trabajo.

Antes de continuar el procedimiento que se llevó de cómo se ha utilizado estas ST para la red neuronal, se hablará primero, como un antecedente, del Perceptrón así como el Perceptrón Multicapa (MLP – Multi-Layer Perceptron), y éste último con un aprendizaje por retropropagación de errores (BP – Back-Propagation). La cuál es el modelo que se ha utilizado para el procesamiento sobre las ST mencionas.

3.8.1 El Perceptrón.

Como un pequeño antecedente, cuando se iniciaron las investigaciones sobre las RNA, algunos investigadores se destacaron por sus contribuciones pioneras, los cuales se hablan de:

- McCulloch y Pitts (1943) por introducir la idea de redes neuronales como máquinas de cálculo.
- Hebb (1949) por postular la primera regla de aprendizaje auto-organizado.
- Ronsenblat (1958) por proponer al perceptrón como el primer modelo para aprendizaje supervisado.

[Sánchez Camperos, 2006]

El Perceptrón, o también conocido como Perceptrón Simple, es un modelo unidireccional, compuesta por dos capas, una de entrada y otra de salida (ver figura 3.8), con pesos ajustables y un umbral. Es la forma más simple de una red neuronal utilizada para la clasificación de patrones linealmente separables (es decir, cuyas regiones de decisión pueden ser separadas mediante una única condición lineal o hiperplano).

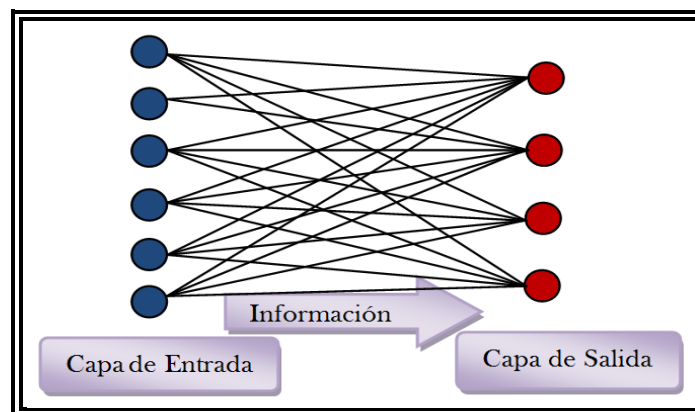


Figura 3.8 Perceptrón simple.

La importancia histórica del perceptrón radica en su carácter entrenable, pues el algoritmo de aprendizaje introducido por Rosenblatt, permite determinar automáticamente los pesos sinápticos que clasifican un conjunto de ejemplos etiquetados, lo cual este algoritmo de aprendizaje es de los denominados **por corrección de errores**, el cual ajustan los pesos en proporción a la diferencia existente entre la salida actual de la red y la salida deseada, con el objetivo de minimizar el error actual de la red.

La operación de un perceptrón con n neuronas de entrada y m de salida, puede escribirse de esta manera:

$$y_i = H\left(\sum_{j=1}^n w_{ij}x_j - \theta_i\right), \forall i, 1 \leq i \leq m \quad (12)$$

La función de activación o de transferencia de las neuronas de la capa de salida es de tipo escalón $H(\cdot)$.

La forma más habitual de la regla de aprendizaje o bien la **regla del perceptrón** se expresa:

$$\Delta w_{ij}^\mu(t) = \varepsilon \cdot (t_i^\mu - y_i^\mu)x_j^\mu \quad (13)$$

y el ajuste de los pesos en la iteración t será:

$$w_{ij}(t+1) = w_{ij}(t) + \sum_{\mu=1}^p \Delta w_{ij}^\mu(t) \quad (14)$$

Donde p es el conjunto de patrones $x^\mu, \mu = 1, \dots, p$, y con t^μ que son sus salidas deseadas. El ritmo de aprendizaje ε es de gran utilización práctica, puesto que un valor pequeño implica un aprendizaje lento, mientras que uno excesivamente grande puede conducir a oscilaciones en el entrenamiento, al introducir variaciones en los peso excesivamente amplias.

3.8.2 El Perceptrón Multicapa.

Si se añade capas intermedias (ocultas) a un perceptrón simple, se obtiene un perceptrón multicapa o MLP. Inicialmente fue desarrollado por Werbos P. J. (1974), debido a las desventajas que presentaba el perceptrón simple la cual es su incapacidad para resolver problemas de clasificación que *no* sean linealmente separables [Haykin S., 1999]. La arquitectura del MLP suele entrenarse mediante el algoritmo denominado **retropropagación de errores** o BP (back-propagation), pero en muchas ocasiones el conjunto *arquitectura MLP + aprendizaje BP* suele denominarse '**red de retropropagación**', o simplemente BP [Martín del Brío, 2007] (ver figura 3.9). Fue Werbos P. J. quien introdujo por primera vez el BP en su tesis doctoral en 1974 [Werbos P. J., 1974].

La descripción de la arquitectura más común del MLP mostrado en la figura 3.9 se puede describir de la siguiente forma:

x_i serán las entradas de la red, así como sus pesos w_{ji} y umbrales θ_j para la sinapsis con la capa oculta; y_j son las salidas de la capa oculta y z_k las de la capa final, por lo que sus pesos

CAPÍTULO 3. METODOLÓGÍA. REDES NEURONALES ARTIFICIALES

y umbrales para la conexión de estas últimas capas serán w_{kj} y θ_k ; por otro lado t_k son las salidas objetivo (target). Todo esto se expresa matemáticamente de la siguiente manera:

$$z_k = \sum_j w_{kj} y_j - \theta_k = \sum_j w_{kj} f \left(\sum_i w_{ji} x_i - \theta_j \right) - \theta_k \quad (15)$$

Siendo la función de transferencia $f(\cdot)$ de tipo sigmoideo (ver tabla 3.1), y usualmente, como se observa en la expresión matemática anterior, se utilizan neuronas de *salida lineal*, la cual se considera identidad (función de transferencia de tipo lineal) como se han descrito en un apartado de este capítulo sobre las funciones de activación o de transferencia.

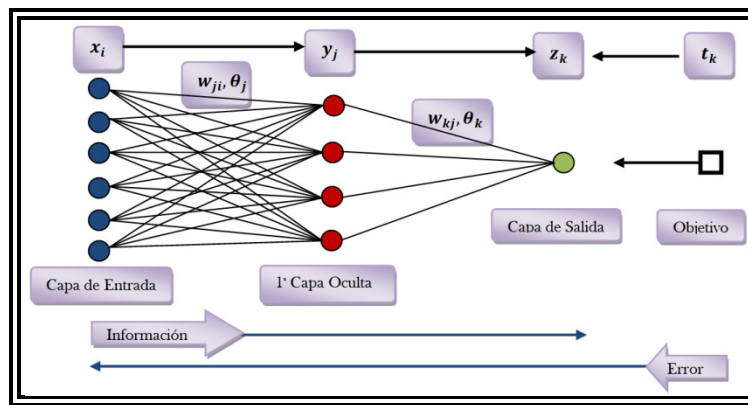


Figura 3.9 Arquitectura de una red neuronal del tipo MLP

Este tipo de red, han sido aplicadas satisfactoriamente para resolver muy diversos y difíciles problemas, y se podría resumir en dos etapas:

- 1) Etapa hacia adelante. Se fijan los parámetros de la red y se presenta una señal de entrada a la red, que se propaga hacia adelante para producir la salida.
- 2) Etapa hacia atrás. El error entre la salida y el objetivo, la red se propaga hacia atrás cuyos parámetros se modifican para minimizar dicho error

Las características distintivas del MLP son:

1. El modelo de cada neurona en la red incluye una función de activación no lineal. Lo importante de esta característica es la no linealidad la cual es suave, en otras palabras, que en cualquier otro punto existen todas sus derivadas.

2. La red contiene una o más capas ocultas que *No* son parte de las entradas o salidas de la red. Estas neuronas ocultas permiten que la red aprenda tareas complejas por la extracción progresiva de las características principales de los patrones de entrada.

[Sánchez Camperos, 2006]

3.8.2.1 Aprendizaje por BP

Las reglas por corrección de errores, son manipuladas sobre los datos a comparar que originalmente se proponen como unidades de entrenamiento simples. Estas reglas manejan el error de la salida hasta aproximarlos a cero. Algunos ejemplos son la regla del perceptrón, la regla de mínimos cuadrados promedio (Least Mean Square, LMS) y la regla delta generalizado [Hasson M.H., 1995].

La regla de aprendizaje LMS es otra regla de aprendizaje que se utiliza para encontrar los pesos óptimos $\mathbf{w}$ y se desarrolla al iniciar la fase de entrenamiento, se inicializa el vector de pesos de manera aleatoria, después se calcula la suma de los pesos por las entradas, se compara el resultado con el valor esperado y se calcula el error. Una vez calculado el error, éste se utiliza para ajustar los pesos.

Para el algoritmo BP aparece como una consecuencia natural de extender el algoritmo LMS, para ello se hace uso de la regla de la cadena, esto debido a que la función de error estará en función, tanto los pesos de salida y los pesos de las neuronas de la capa oculta. Basados en la estructura de la figura 3.9, con las entradas, salidas, pesos y umbrales de las neuronas, ya anteriormente descritas se tiene:

$$\mathbf{z}_k = \sum_j \mathbf{w}_{kj} \mathbf{y}_j - \theta_k = \sum_j \mathbf{w}_{kj} f(\sum_i \mathbf{w}_{ji} \mathbf{x}_i - \theta_j) - \theta_k \quad (16)$$

La función de error (función coste) de la cual parte del error cuadrático medio

$$E(\mathbf{w}_{ji}, \theta_j, \mathbf{w}_{kj}, \theta_k) = \frac{1}{2} [\sum_k \mathbf{t}_k - \mathbf{z}_k]^2 \quad (17)$$

Un problema para la optimización, o bien la búsqueda de un mínimo de una función, aparecen una serie de procedimientos de búsqueda que, generalmente, se dividen en dos grupos: métodos de *búsqueda global* en donde se busca el *mínimo global* de la función objetivo proporcionando los pesos $\mathbf{w}$ que dan el valor más pequeño de dicha función sobre todo su dominio y, métodos de *búsqueda local*, que buscan el *mínimo más cercano* de la

CAPÍTULO 3. METODOLOGÍA. REDES NEURONALES ARTIFICIALES

función objetivo en relación al estado inicial de los pesos al comenzar dichos algoritmos (ver figura 3.10).

En diversas experimentaciones se obtienen unos pesos peores para diferentes métodos, pero hay algoritmos mucho más rápidos en la obtención de la solución. Entre estos métodos se encuentran todos los basados por el *descenso por gradiente*, gradientes conjugados (que pueden ser la de direcciones conjugados, conjugación Gram-Schmidt, propiedades de los residuales, método de los gradientes conjugados versión lineal y no lineal), descenso más inclinado, etc. [Pérez Pérez, 2001]

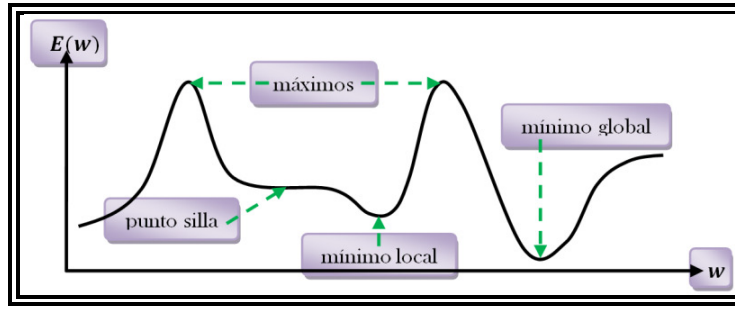


Figura 3.10 Búsqueda de un mínimo de una función

La minimización del error del aprendizaje BP se llevará a cabo mediante el **descenso por el gradiente** (ver figura 3.11); el cual se tendrá un gradiente respecto de los pesos de la capa de salida w_{kj} y otro, respecto de los de la oculta w_{ji} :

$$\delta w_{kj} = -\varepsilon \frac{\partial E}{\partial w_{kj}} \quad ; \quad \delta w_{ji} = -\varepsilon \frac{\partial E}{\partial w_{ji}} \quad (18)$$

Las expresiones de actualización de los pesos se obtienen derivando, como se muestra en la ecuación anterior, teniendo en cuenta las dependencias funcionales y aplicando adecuadamente la regla de la cadena:

$$\delta w_{kj} = -\varepsilon \Delta_k y_j \quad ; \quad \Delta_k = [t_k - g(h_k)] \frac{\partial g(h_k)}{\partial h_k} \quad ; \quad h_k = \sum w_{kj} y_j - \theta_k \quad (19)$$

$$\delta w_{ji} = -\varepsilon \Delta_j x_i \quad ; \quad \Delta_j = [\sum_k \Delta_k w_{kj}] \frac{\partial g(h_j)}{\partial h_j} \quad ; \quad h_j = \sum w_{ji} x_i - \theta_j \quad (20)$$

donde h_k y h_j son los potenciales post-sinápticos (regla de propagación). Las actualizaciones de los umbrales se realiza haciendo uso de las expresiones anteriores.

Con estas expresiones está implícito el concepto de propagación hacia atrás de los errores (BP). Primero se calcula la expresión Δ_k , que se denominara *señal de error*, por ser proporcional al error de la salida actual de la red, con el que se calcula la actualización δw_{kj} de los pesos de la capa de salida. A continuación se propaga la señal de error hacia atrás, proporcionando así las señales de error Δ_j correspondientes a las sinapsis de la capa oculta; con éstas se calcula la actualización δw_{ji} . El algoritmo puede extenderse fácilmente a arquitecturas con más de una capa siguiendo el mismo esquema.

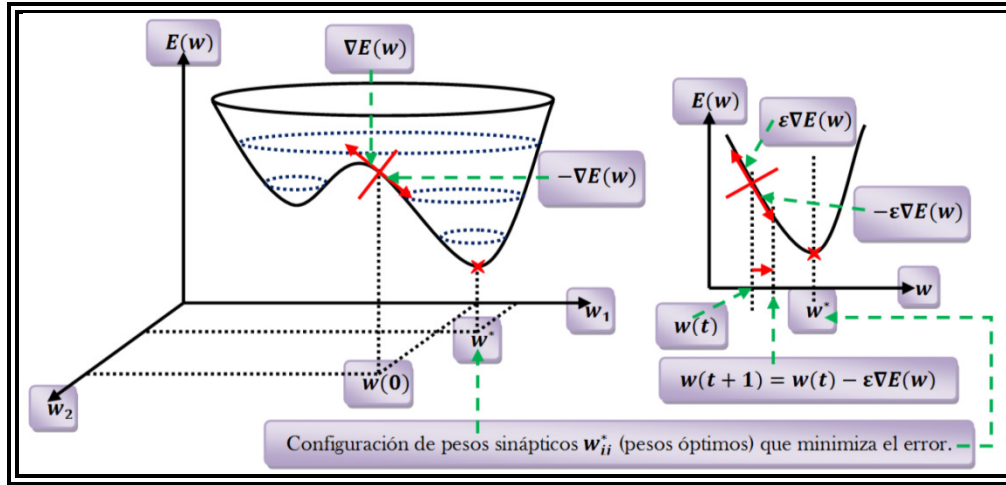


Figura 3.11 Superficie de error $E(w)$ en el espacio de los pesos w , y descenso por el gradiente hacia un mínimo local.

3.8.2.2 Aceleración del BP

Debe tenerse en cuenta que el algoritmo BP no garantiza alcanzar el mínimo global de la función de error, tan sólo en un mínimo local, por lo que el proceso de aprendizaje puede estancarse en uno de éstos mínimos locales.

Ahora bien, para resolver algunos de los inconvenientes del BP es el problema de su lenta convergencia. La primera variante la propusieron Rumelhart D.E. en 1986, quien con un grupo de investigación en **RNA**, popularizaron el modelo BP [Rumelhart D.E., 1986], la cual hacen incluir en el algoritmo un término de inercia (momentum), consistente en añadir al cálculo de la variación de los pesos un término adicional proporcional al incremento de la iteración anterior, que proporciona una cierta inercia al entrenamiento:

$$\delta w_{kj}(t+1) = -\varepsilon \frac{\partial E}{\partial w_{kj}} \Big|_t + \alpha \cdot \delta w_{kj}(t-1) \quad (21)$$

$$\delta w_{ji}(t+1) = -\varepsilon \left. \frac{\partial E}{\partial w_{ji}} \right|_t + \alpha \delta w_{ji}(t-1) \quad (22)$$

con α un parámetro entre 0 y 1, que se suele tomar próximo a 1 ($\alpha \approx .9$).

Es ésta una manera de aumentar el ritmo de aprendizaje efectivo en determinadas direcciones.

La razón por la que se usan funciones sigmoides como funciones de transferencia o de activación para el BP, es que este algoritmo busca minimizar la función de error para los pesos mediante el cálculo del gradiente en dirección descendente (ver figura 3.11). Como se debe calcular el gradiente de la función de error en cada iteración, se tendrá que garantizar su continuidad y diferenciabilidad en todo momento.

La derivada de la sigmoide es muy importante para el funcionamiento del algoritmo, ya que consiste en calcular el gradiente de la función error compuesta con la sigmoide para moverse en la dirección en la que este disminuya. Este estilo de minimizar la función se podría pensar como un proceso físico equivalente al poner una canica en la superficie formada por la función de error y dejarla rodar sobre ella hasta encontrar el mínimo [Rojas Raul, 1996].

Esto siempre lleva en una dirección más o menos correcta, porque la derivada de la sigmoide siempre es positiva, marcándonos la dirección adecuada para movernos; con la desventaja de que algunas veces el gradiente es muy grande y otras muy pequeño, haciendo que pueda resultar un poco complicado seguir la dirección óptima.

3.8.3 Capacidad de generalización de la red.

3.8.3.1 Validación cruzada (cross-validation)

La forma habitual de entender los sistemas supervisados, es dividir los ejemplos disponibles en dos conjuntos: entrenamiento y generalización, normalmente en una relación **2/3** a **1/3**. Usualmente, de todo un conjunto de entrenamiento se emplea aproximadamente en **80%** de los patrones para entrenar, reservándose un **20%** como *conjunto test* [Haykin S., 1999].

Con el primer conjunto de datos se entrena la red, esto es, se ajustan las conexiones de la red o pesos, y con el segundo se comprueba su capacidad de generalización. Se entiende por

CAPÍTULO 3. METODOLÓGÍA. REDES NEURONALES ARTIFICIALES

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

Hay que tener en cuenta que una de las principales características de las redes neuronales, que las hacen especiales frente a otros métodos, es su capacidad de generalización, es decir, ante entradas desconocidas son capaces de dar salidas aproximadas a las deseadas.

En un proceso de entrenamiento, se debe considerar, por una parte, un **error de aprendizaje**, que se suele calcular como el error cuadrático medio de los resultados proporcionados por la red. Para una red suficientemente grande, puede reducirse tanto como se quiera sólo con llevar a cabo *más iteraciones (épocas)*.

Por otra parte, existe un **error de generalización** (medida de error en test), que se puede medir empleando un conjunto representativo, diferentes a los utilizados en el entrenamiento. De esta manera, se puede entrenar una red neuronal haciendo uso de un conjunto de aprendizaje, y comprobar su eficiencia real, o error de generalización, mediante un *conjunto test*.

Si representamos a la vez el error en aprendizaje y el error en test durante el transcurso del aprendizaje (ver figura 3.12), este procedimiento consistente en entrenar y validar, que es para encontrar y detenerse en el punto óptimo, el cual se le denomina **validación cruzada** (cross validation), y es ampliamente utilizada en la fase de desarrollo de una red neuronal supervisada, tal es el caso para el MLP.

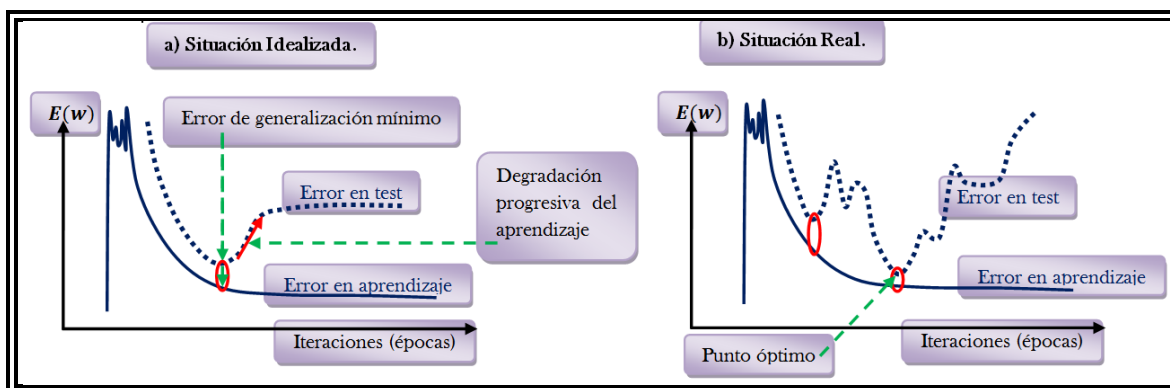


Figura 3.12 Evolución del error de aprendizaje y del error de generalización [Prechelt L., 1998].

CAPÍTULO 3. METODOLÓGÍA. REDES NEURONALES ARTIFICIALES

Para una situación idealizada (Figura 3.12a), tras una fase inicial, en la que pueden aparecer oscilaciones en el valor del error, el de aprendizaje tiende a disminuir monótonamente, mientras que el de generalización al principio la red se adapta progresivamente al conjunto de aprendizaje, adaptándose al problema y mejorando la generalización, pero a partir de cierto punto comienza a incrementarse.

Para una situación más realista (Figura 3.12b), pueden presentarse varios mínimos para el conjunto test, debiendo detener el aprendizaje en el punto óptimo de mínimo error de generalización, y no quedarnos en el primer mínimo que aparezca en el error en test.

La **validación cruda** es utilizada por el denominado **sobre-ajuste** o **sobre-aprendizaje** (overfitting o overtraining). Dado que el sistema se ajusta demasiado, aprendiendo demasiado, incluso el ruido presente en los ejemplos empleados, por lo que crece el error de generalización.

El potencial de las componentes descritas, será investigado en los experimentos numéricos con ST de *datos teóricos* reconocidos mundialmente (*datos del reanálisis*) y de *datos observados*, en el siguiente capítulo.

CAPÍTULO 4

EXPERIMENTOS NUMÉRICOS.

Una de las principales inquietudes que ha tenido la historia de la humanidad, ha sido estimar el futuro, utilizando información del presente y del pasado. Esto, se llama **predecir**. Es evidente que en las diversas áreas de investigación, requieren conocer el comportamiento futuro de ciertos sucesos, en este caso eventos naturales, con el fin de planificar y prevenir.

La visión genérica de los modelos, es que **un modelo**, es una **representación de la realidad**; mientras que la realidad puede ser arbitrariamente *compleja*. El modelo, por necesidad, debe ser más simple, además, se construye no directamente de la realidad, sino de las limitadas observaciones que de ella se tienen, por lo que un modelo es tan bueno, como las observaciones lo son. Con ello, reafirmamos la necesidad de tener una buena calidad, en la base de datos, además de tener éstas sin ninguna ausencia de ellos.

4.1 MODELO EMPLEADO.

Entendida la metodología, la cual es empleada para la generación de los datos faltantes en las ST de las estaciones meteorológicas. Se construyó un modelo de RNA, basados en la arquitectura del MLP con el algoritmo de aprendizaje BP, en otras palabras, se utilizará una **red de retropropagación**. La minimización del error del BP se llevará a cabo mediante el **descenso por el gradiente**, y a su vez, se utilizará la **validación cruzada (CV)**. Esta última, además de reducir el error de generalización, como se explicó anteriormente, se utilizará para apartar un cierto número de datos sin que hayan pasado por el algoritmo de aprendizaje, y observar que tan bueno va siendo el modelo de la RNA y así poder generar los datos (**predicción**), comparándola con un ST completa y saber que error se obtiene, así podremos comprobar el potencial que se obtiene al hacer uso de las RNA.

CAPÍTULO 4.

EXPERIMENTOS NUMÉRICOS.

Para el modelo empleado, se programó en el software comercial, MatLab, versión 7.11.0 (R2010b). Respecto al uso de éste software, es debido a que es un programa amigable al programar diversos algoritmos, como lo fue para este modelo de RNA; con gran variedad de comandos incluidos en sus carpetas, el cual nos son útiles y facilitan el manejo de este software. Además cuenta con su propia interfaz gráfica, la cual es de gran ayuda, pues nos facilita la visualización de las ST. MatLab, de sus siglas en inglés **Matrix Laboratory**, es una gran calculadora científica, la cual puede realizar operaciones matriciales, como bien su nombre lo describe, y esto es muy beneficioso para nuestro algoritmo, debido a que se manejan vectores y matrices; éstas son comúnmente asimétricas, por lo que al realizar operaciones con ellas, en el entorno de MatLab, se facilita su realización.

Existen los llamados *toolbox*, incluidos en este software, el cual incluye uno de *redes neuronales*. Aclarando, de que no se manejó el *toolbox* que contiene MatLab, se programó el algoritmo basados en la explicación y ecuaciones mostradas en el capítulo correspondiente. Sólo se manejaron los comandos más empleados en este entorno de programación, como son: la media (*mean*), la desviación estandard (*std*), cargar datos (*load*), entre otros.

A continuación se mostrarán los experimentos numéricos, realizados por estación, basados en una estadística de errores, la cual, es de gran ayuda para el entrenamiento de la RNA, y así, explotar su potencial del modelo y obtener el mínimo error o bien un error óptimo con fines de generar los datos faltantes.

Refiriéndonos a la estadística de errores, aplicando la función del error cuadrático medio, tenemos la siguiente analogía:

Tomando la estación de Apizaco para la **temp.máx**, como ejemplo (ver **tabla 4.1**), tenemos que para la 2a columna, es el error de las ST teóricas respecto a las observadas, éste también mostrado en la **tabla 2.1**.

Se obtienen las anomalías de las ST, tanto teóricas, como observadas, ¿a qué nos referimos con sus anomalías? Para las **ciencias atmosféricas**, obtener su anomalía, o trabajar sólo con su anomalía, es simplemente restar la media (media aritmética) de la ST para cada uno de los datos de la misma. En este trabajo, se utiliza una ventana

CAPÍTULO 4.

EXPERIMENTOS NUMÉRICOS.

anual, o sea, se obtienen las anomalías de cada año. Posteriormente, se obtiene el error de las anomalías mostradas en la columna 3 de la tabla 4.1. Se observa que el error disminuye y es debido a que al trabajar con anomalías, tenemos que la media de éstas son iguales a cero, teniendo así ST estacionarias²⁵.

La estandarización es otra forma de trabajar las anomalías, como éstas tienen ya media nula, basta con dividir por la desviación estándar. Por lo que todas las variables estandarizadas tienen media nula y desviación estándar igual a 1. Aplicando, por separado, esta estandarización a las anomalías de las ST teóricas y observadas se obtiene el error mostradas en la columna 4 de la tabla en cuestión.

Es de gran importancia *normalizar* las ST, para poder emplear la RNA. Al normalizar la ST, estamos dejando un rango de magnitudes entre 0 y +1, o bien de -1 a +1, porque al aplicar la función de transferencia o de activación en la RNA, está en un rango igual al mencionado, tal como sus respectivas derivadas de éstas; por lo que siempre se ha de normalizar cualquier ST al trabajar con una RNA. Teniendo en cuenta esto, se está normalizando las anomalías estandarizadas, y así sacar los errores mostrados en la columna 5.

Para la última columna de la tabla 4.1, se obtiene el error de las ST teóricas respecto a las observadas, sólo, aplicando su normalización de las ST desde un principio, esto es, sin haber obtenido sus anomalías ni estandarización, y así aplicar la función coste.

Al hacer este análisis, se tiene que el error sufrió cambios; observando que reduce su error de inicio (columna dos, tabla 4.1). Corresponde a la RNA reducirlo mucho más, con el objetivo de generar los datos que hagan falta a las correspondientes ST. El basarnos de este análisis, es para saber si se utiliza la normalización de las anomalías estandarizadas o sólo la normalización de las ST (sin obtener sus anomalías ni estandarización), con el fin de ayudar a la RNA a buscar un mínimo u óptimo error.

25 Estacionariedad.- Los valores de una ST, son estadísticamente similares (misma media y varianza). La estacionariedad implica que la media de la serie de datos, no cambia con el tiempo. Cualitativamente, porciones diferentes de una ST estacionaria son estadísticamente parecidas, aún cuando los valores individuales pueden ser muy diferentes.

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

años	Error Cuadrático Medio - Temp.máx.				
	ST	Anomalías	Estandarización (Estd.)	Normalización de la Estd.	Normalización de la ST
1979	129913	11344	130.2096	0.9084	0.0308
1980	101370	5674	67.5574	0.6704	0.0157
1981	139582	7514	94.1387	1.2158	0.0513
1982	196246	8916	140.9493	1.7217	0.0339
1983	189800	17662	93.8388	0.8369	0.0251
1984	444294	42408	419.1343	3.4394	0.0620
1985	165244	5072	95.3612	1.0536	0.0106
1986	NaN	NaN	NaN	NaN	NaN
1987	133739	6974	183.5627	1.6678	0.0231
1988	144381	7734	76.1634	0.7679	0.0177
1989	NaN	NaN	NaN	NaN	NaN
1990	204923	10074	185.7726	1.4079	0.0304
1991	NaN	NaN	NaN	NaN	NaN
1992	132426	17090	207.4008	1.8129	0.0344
1993	193581	11352	295.0685	1.6079	0.0361
1994	NaN	NaN	NaN	NaN	NaN
1995	148537	6089	67.6376	0.6692	0.0117
1996	155241	8574	133.2528	1.2851	0.0197
1997	185056	9293	258.9528	1.5552	0.0234
1998	258023	28327	102.5310	1.2650	0.0242
1999	234316	5144	55.2040	0.9724	0.0279
2000	232616	4910	89.5623	1.0308	0.0399
2001	106458	4067	107.1272	1.2335	0.0394

Tabla 4.1 Análisis de errores de **temp.máx.** (ST de la Estación de Apizaco).

Se hace uso de una RNA supervisada (*red de retropropagación*), definiendo los *datos teóricos* como la serie de entrada (*input*) y los *datos observados* serán el objetivo (*target*).

Se trabajó con la función de transferencia sigmoidea de rango [-1 1] (tabla 3.1). Un ritmo de aprendizaje $\varepsilon = 0.6$; en la literatura, un valor pequeño ($\varepsilon = 0.1$) implica un aprendizaje lento, por lo que el proceso tomaría bastante tiempo, mientras que uno grande ($\varepsilon = 0.9$) puede conducir a oscilaciones en el entrenamiento de la RNA, introduciendo variaciones en los pesos excesivamente amplias. El parámetro de aceleración del aprendizaje BP fue $\alpha = 0.9$. Estos parámetros mencionados, se aplicó para ambas temperaturas y cada estación meteorológica.

4.2 ESTACIÓN DE APIZACO.

En esta estación, para la **temp.máx.**, se efectuó su normalización de la ST, sin obtener sus anomalías ni estandarización, teniendo esta ST normalizada como entrada para la RNA.

CAPÍTULO 4.

EXPERIMENTOS NUMÉRICOS.

El año de datos faltantes que se trabajará, es el de 1986, la cual hay ausencia de 184 datos (al final del año), por lo que se toma las ST desde el año de 1979 hasta mitades de 1986, tomando en cuenta que de 1979 se tienen *datos teóricos*, y así, se realizaron experimentos reduciendo las ST, esto es, de 1980 a 1986, de 1981 a 1986, etc. (ver figura 4.1).

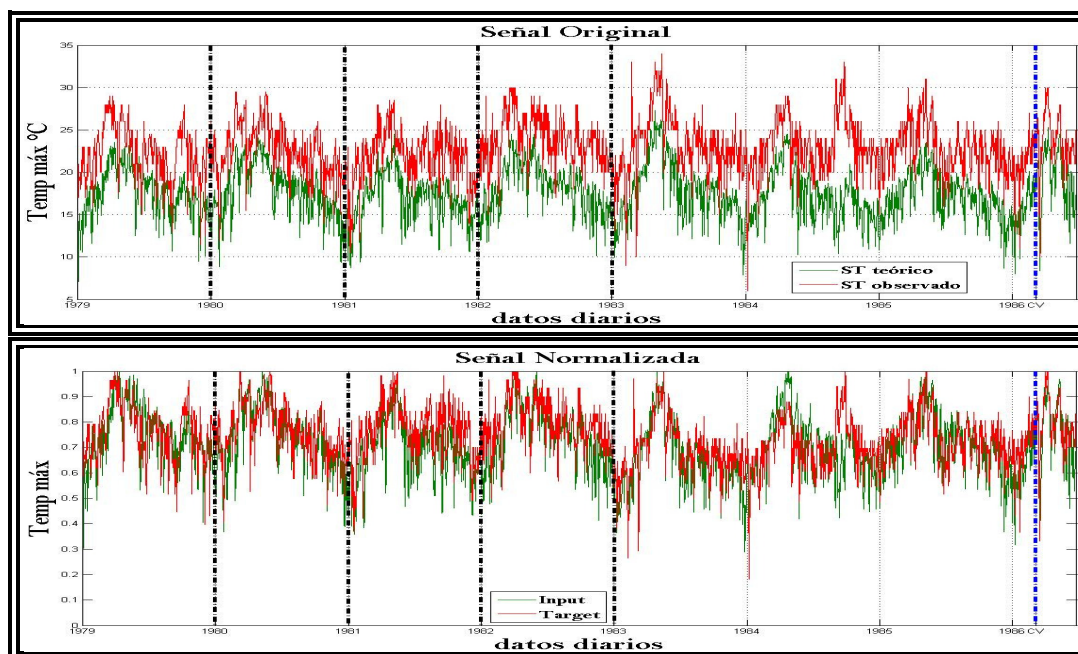
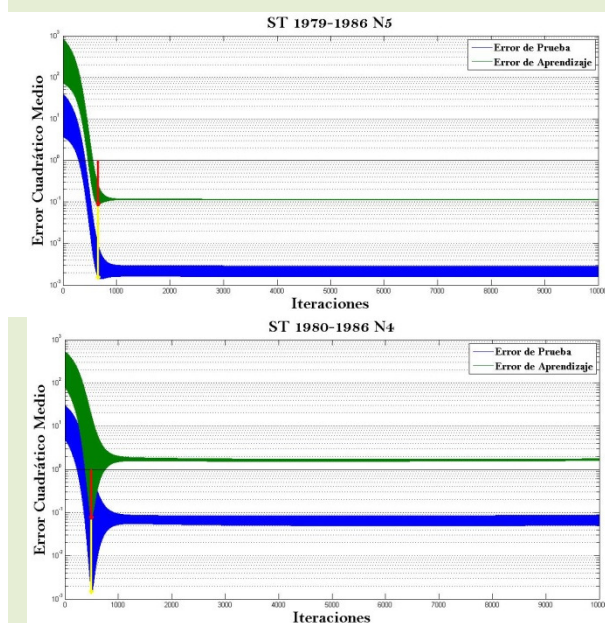


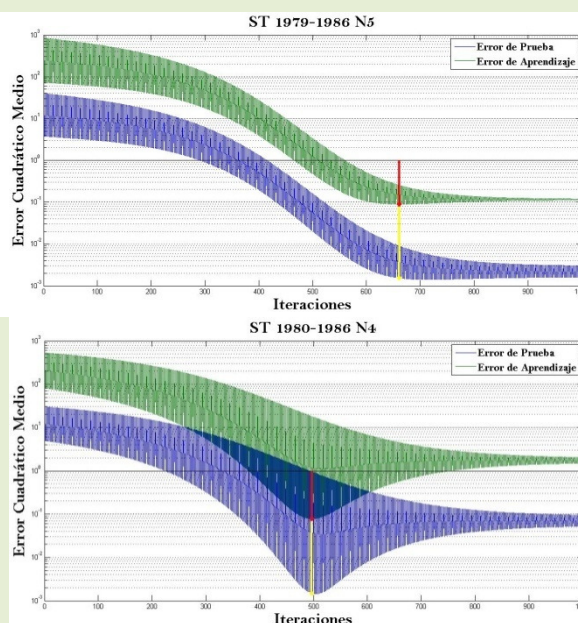
Figura 4.1 ST para Entrenar (temp.máx. estación de Apizaco).

Búsqueda del error mínimo.

Gráfica de errores.

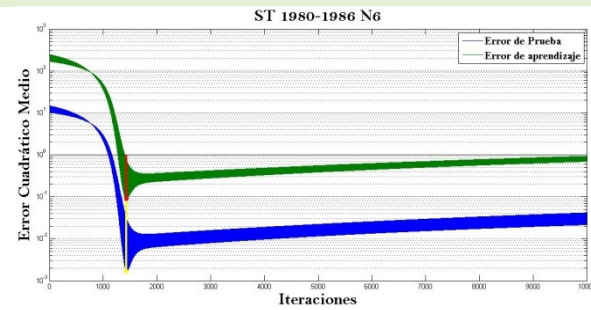


Zoom del mínimo error.

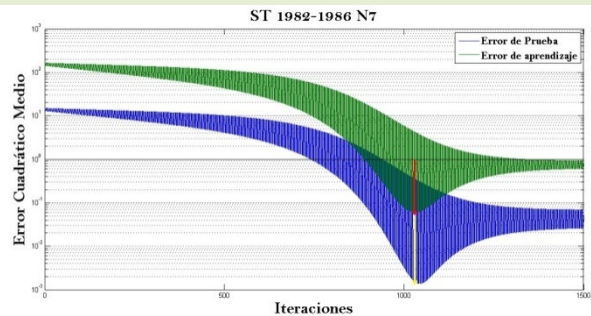
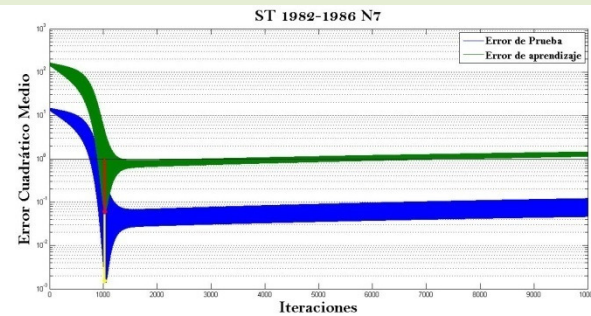
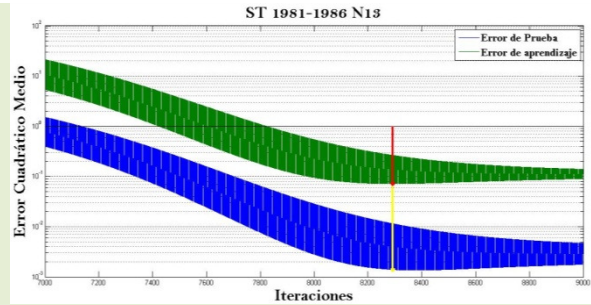
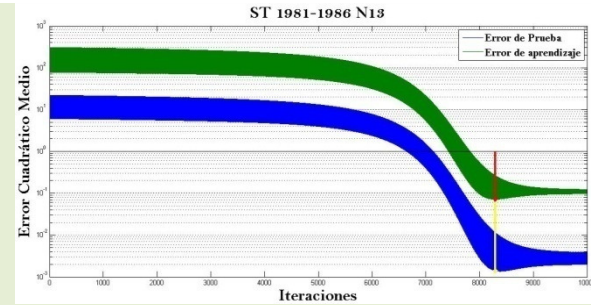
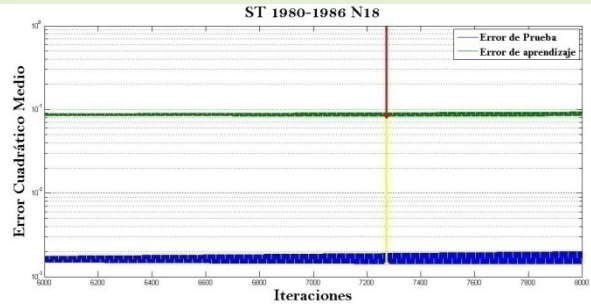
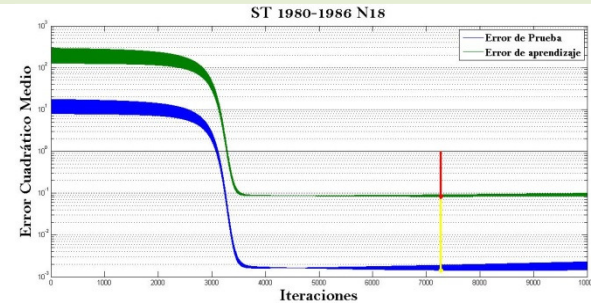
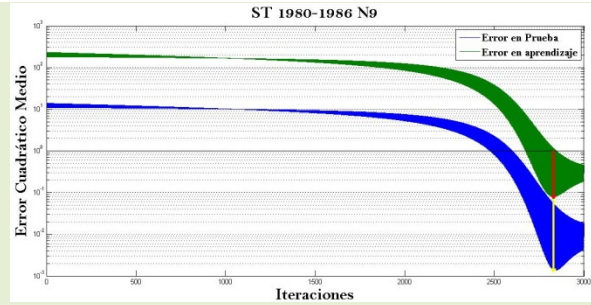
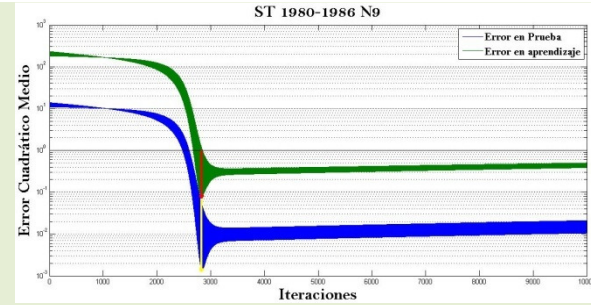
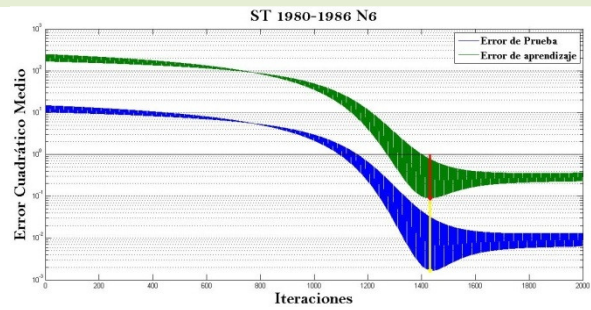


Búsqueda del error mínimo.

Gráfica de errores.

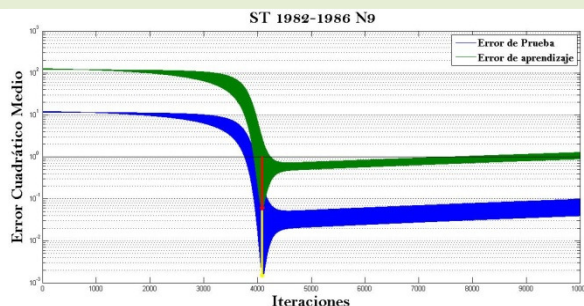


Zoom del mínimo error.



Búsqueda del error mínimo.

Gráfica de errores.



Zoom del mínimo error.

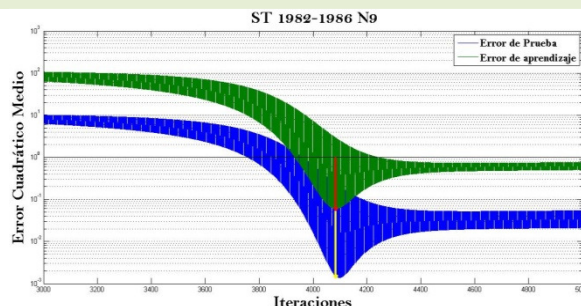


Tabla 4.2 Gráficas del Error de Aprendizaje y Prueba (**temp.máx.** Estación de Apizaco).

- E.P. Error de Prueba o Test, 110 datos Reservados de la ST sin entrenar (CV).
 E.A. Error de Aprendizaje o Entrenamiento.
 E.S.O. Error con respecto a la ST Original (de la fig.4.1, ST desde 1979 a 1986, incluyendo los 110 reservados para la CV).
 E.S.C. Error respecto a una ST Completa sin haber sido Entrenada o Reservada (ST-1987).

ST Entrenada	# Neuronas en la C.O.	Tiempo de Entrenamiento (s)	E.P. 110 datos.	E.A.	E.S.O.	E.S.C.
1979-1986	5	420.661625	0.0015	0.0892	0.0528	0.0083
1980-1986	4	284.487898	0.0015	0.0777	0.0534	0.0084
	6	397.321223	0.0014	0.0802	0.0548	0.0083
	9	432.772808	0.0014	0.0795	0.0524	0.0082
	18	438.131490	0.0014	0.0836	0.0563	0.0088
1981-1986	13	264.723990	0.0014	0.0718	0.0529	0.0082
1982-1986	7	224.509196	0.0015	0.0603	0.0505	0.0082
	9	246.029660	0.0015	0.0601	0.0506	0.0083

Tabla 4.3 Análisis del Aprendizaje (**temp.máx.** Estación de Apizaco).

El error, de la ST de entrada respecto al de objetivo, ambas normalizadas y sin haber pasado por el entrenamiento de la RNA (ver figura 4.1), es de 0.0977. La ST es desde 1979 hasta donde se tiene de 1986.

Para la ST completa de 1987, normalizado, se tiene un error de 0.0231 (ver figura 4.2).

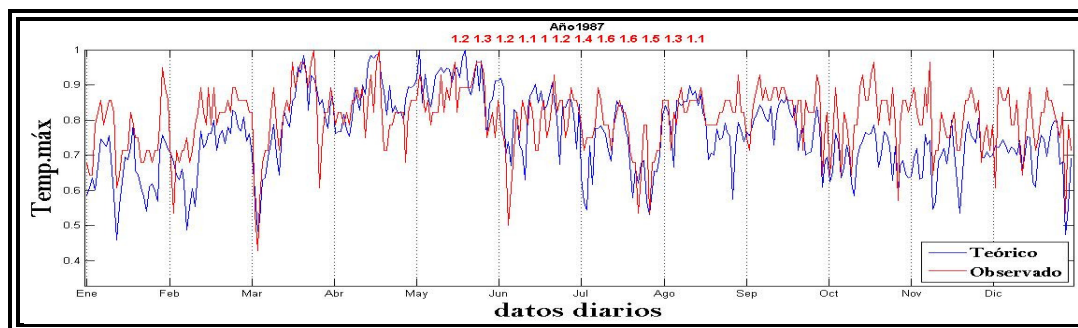


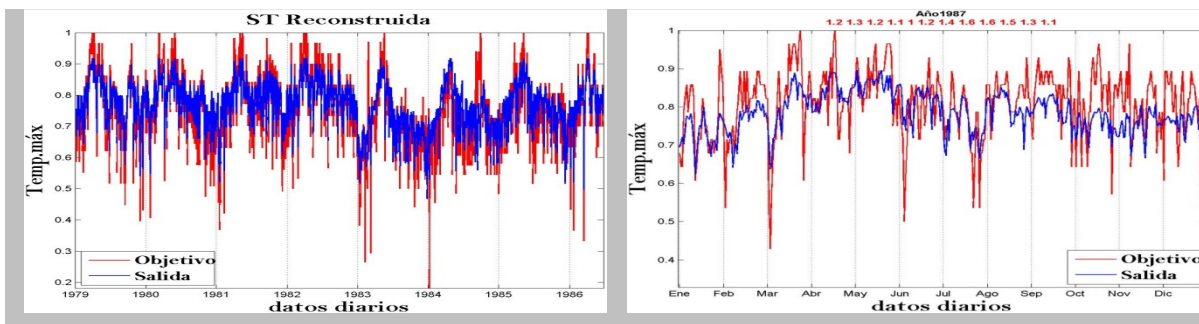
Figura 4.2 Prueba para una ST Completa, 1987.

Reconstrucción de la ST Original
Normalizada, utilizada en la RNA.

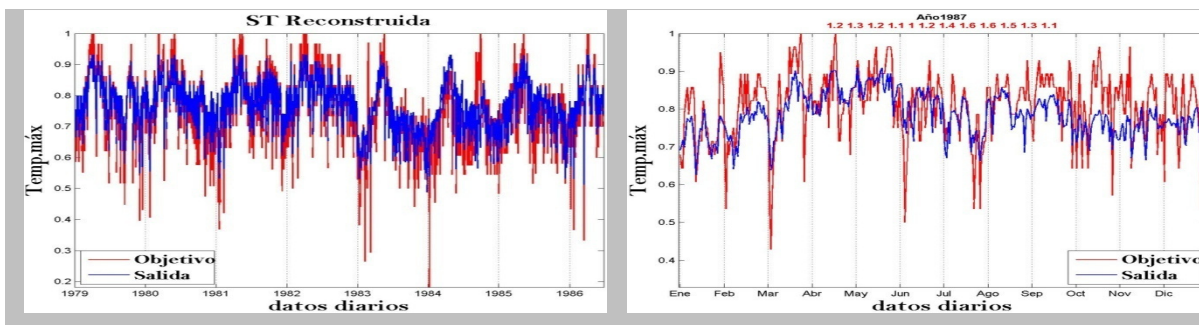
Reconstrucción de una ST Completa,
año de 1987

Con los Pesos y Umbrales Óptimos.

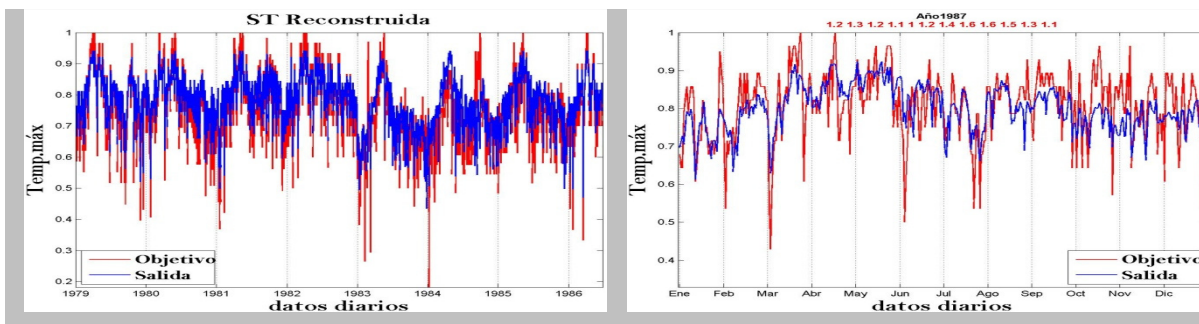
5 neuronas en la capa oculta.



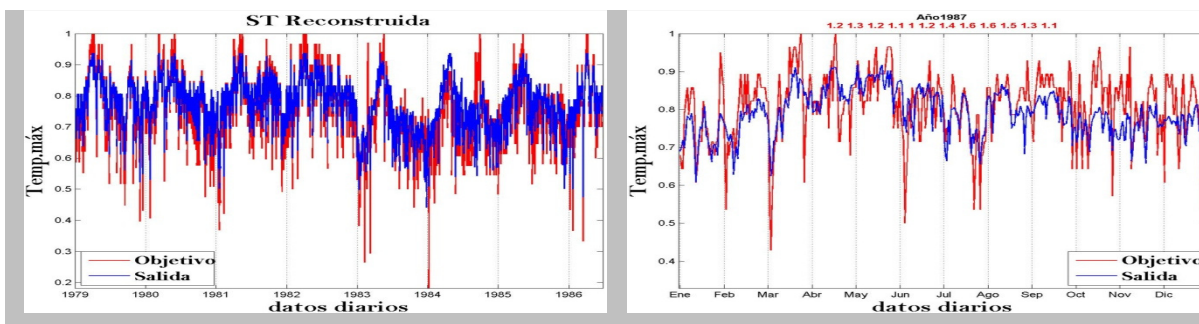
4 neuronas en la capa oculta.



6 neuronas en la capa oculta.



9 neuronas en la capa oculta.

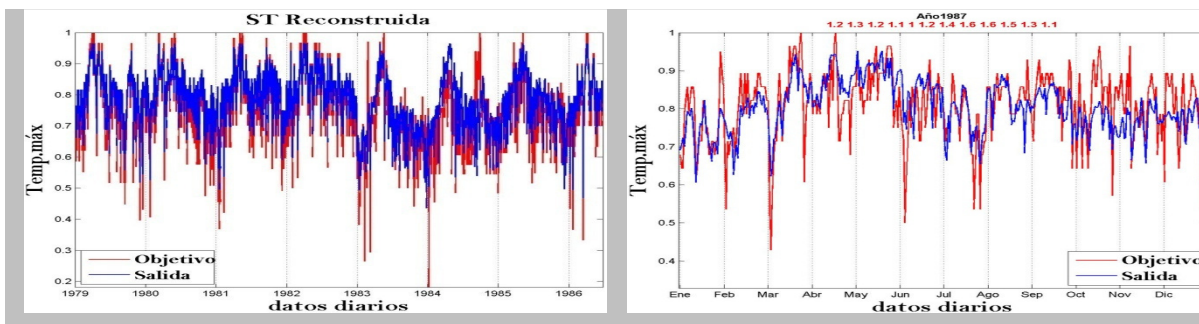


Reconstrucción de la ST Original
Normalizada, utilizada en la RNA.

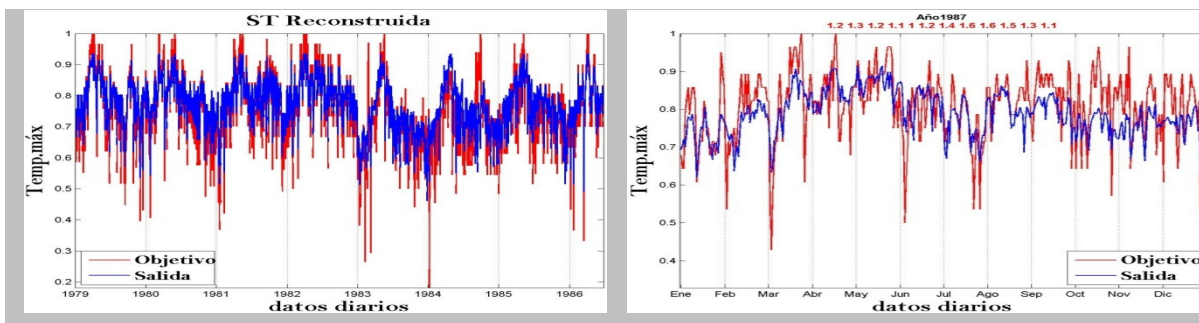
Reconstrucción de una ST Completa,
año de 1987

Con los Pesos y Umbrales Óptimos.

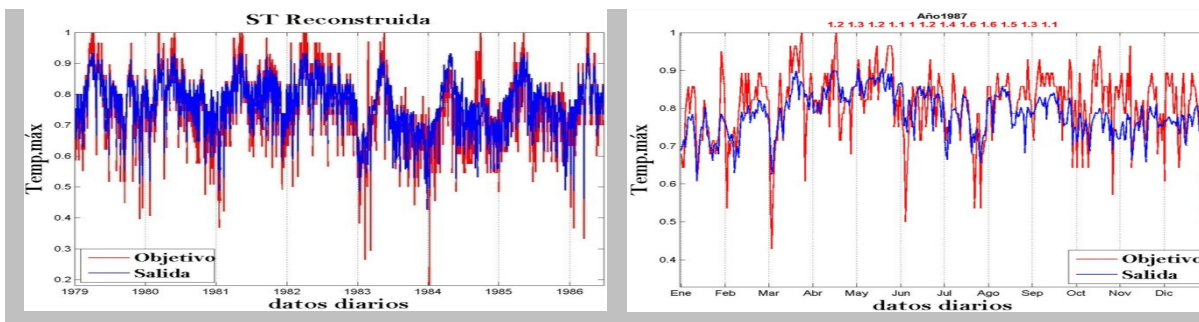
18 neuronas en la capa oculta.



13 neuronas en la capa oculta.



7 neuronas en la capa oculta.



9 neuronas en la capa oculta.

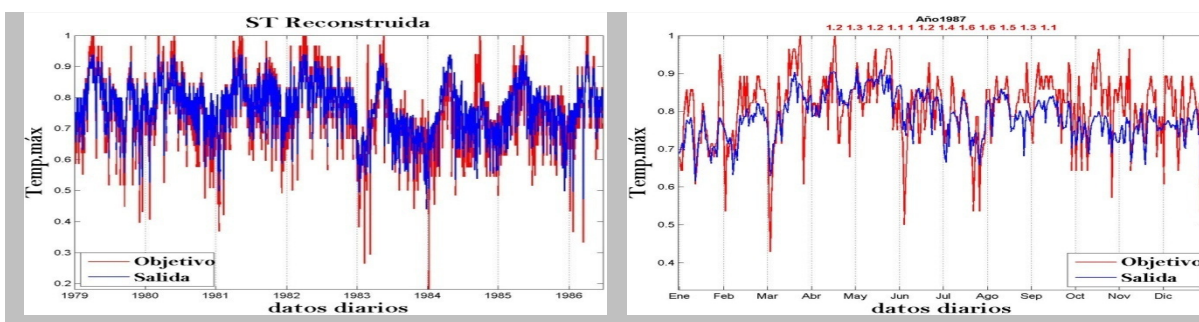


Tabla 4.4 Gráficas de Reconstrucción de la ST Original y de una Completa.

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

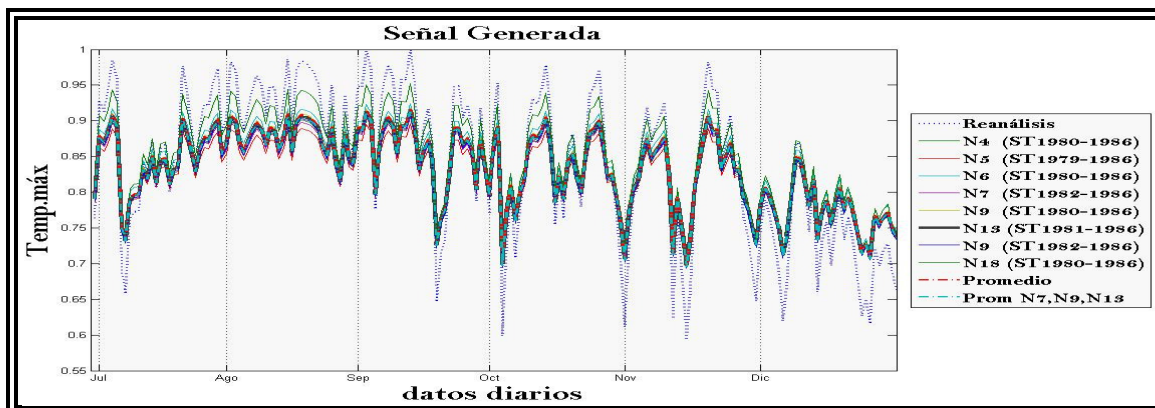


Figura 4.3 Datos Generados para el año de 1986.

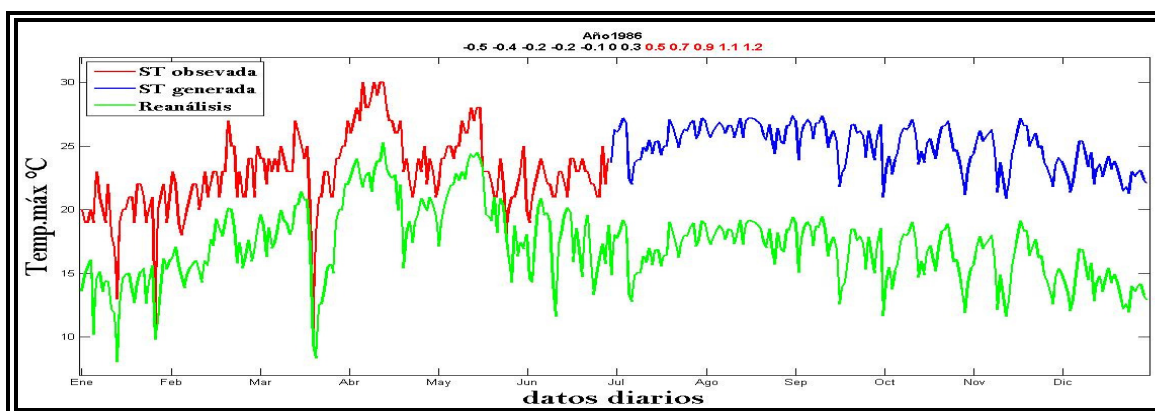


Figura 4.4 ST Reconstruida del año de 1986.

Para el año de 1989 hace falta un dato en el mes de Julio, graficando la ST de 10 días antes y después respecto al dato faltante, tenemos el siguiente comportamiento, mostrando también los mismos días para los datos de reanálisis.

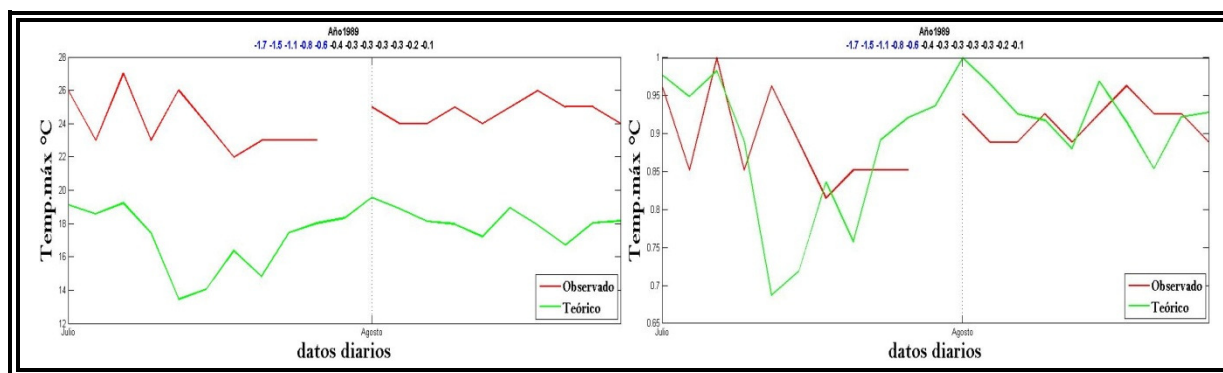


Figura 4.5 ST del año de 1989; ST Original (izquierda); ST Normalizada (derecha).

Observados	26	23	27	23	26	24	22	23	23	23	NaN	25	24	24	25	24	25	26	25	25	24
Reanálisis	19.129	18.57	19.226	17.422	13.444	14.056	16.362	14.819	17.46	18.022	18.333	19.568	18.887	18.123	17.961	17.216	18.948	17.907	16.706	18.044	18.158

Tabla 4.5 Diez datos Antes y Después del dato Faltante, año de 1989.

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

Empíricamente, podemos reemplazar el dato ausente, promediando los datos, anterior y posterior al faltante. Por lo que tendríamos el valor de 24°C; este método es adecuado y usualmente empleado para hacer pronóstico cuando hay persistencia (ver tabla 4.4). Pero al generar este dato faltante con la RNA tendríamos lo siguiente:

Utilizando los pesos y umbrales óptimos de 13 neuronas, utilizadas en la capa oculta del entrenamiento previo, tenemos que el valor generado para el dato faltante es de 23.8221. Considerando que los datos observados sólo puede haber números enteros, o bien, fraccionarios con una décima de 5; se redondea el valor generado obteniendo el valor de 24°C; mismo valor obtenido al reemplazar el dato faltante empíricamente.

En la siguiente tabla, 4.5, se muestran ST de valores previos y posteriores a los datos faltantes tal como en la tabla 4.4. Siguiendo el registro que se tiene en el capítulo1 en el apartado de ST incompletas para la estación en cuestión; se reemplazan los datos ausentes empíricamente, al igual que el generarlos con la RNA, como se mostró previamente.

	Datos Observados.	Datos generados empíricamente.	Datos generados con RNA
1991	20 22 22 20 12 19 19 19 19 19 NaN Fin del mes de Dic.		
	22 20 21 21 23 22 21 20 21 22 Inicio del mes de Ene. 1992.	20.5	19.9111 - 20
	24 21 22 22 22 24 21 22 21 21 NaN Fin del mes de Dic.		
1994	24 20 22 22 20 20 25 24 24 24 Inicio del mes de Ene. 1995.	22.5	22.4650 - 22.5

Tabla 4.6 Generación de los Datos Faltantes de 1991 y1994 (**temp.máx.** Estación de Apizaco).

Para la **temp.mín.** se escoge la ST para entrenar de 1979 hasta mitades de 1986, tal como la de **temp.máx.**, sólo que no se redujo tanto la ST, esto se muestra en la tabla de análisis del aprendizaje para la **temp.mín.**

Para el entrenamiento en la RNA, de la ST de **temp.mín.** se efectúo su normalización, tal como la variable anterior. Teniendo que para todas las ST de entrada, para cada estación y variable, se entrena sin obtener sus anomalías ni estandarización, simplemente su normalización.

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

años	Error Cuadrático Medio - Temp.mín.				
	ST	Anomalías	Estandarización (Estd.)	Normalización de la Estd.	Normalización de la ST
1979	70754	14261	149.9528	2.3189	2.6255
1980	66757	11869	115.1842	1.4728	1.5672
1981	44735	13156	97.6323	1.1055	1.2667
1982	158848	14489	356.9820	3.1660	3.1012
1983	61547	13725	82.9258	1.1502	1.5457
1984	60209	12256	164.9383	1.4061	1.6336
1985	77741	17442	223.6065	2.7459	2.6901
1986	NaN	NaN	NaN	NaN	NaN
1987	NaN	NaN	NaN	NaN	NaN
1988	88875	16460	134.4446	1.6287	2.3337
1989	103901	51854	206.2918	2.9440	3.8020
1990	55030	14161	148.9920	1.8696	1.7671
1991	NaN	NaN	NaN	NaN	NaN
1992	51142	10552	146.9188	1.4176	1.4846
1993	101571	15294	238.7074	6.1800	4.2548
1994	101658	13624	177.6077	2.3883	3.0865
1995	89025	18155	201.0102	2.8753	3.2705
1996	128762	22867	149.1740	1.5409	5.4143
1997	83950	13873	165.5193	1.1491	2.1350
1998	129128	29357	171.1520	3.1311	3.4172
1999	129308	24125	124.4821	1.4911	5.1196
2000	94342	19218	200.7343	1.7033	3.8894
2001	61853	10444	255.1621	3.4244	1.6331

Tabla 4.7 Análisis de Errores de **temp.mín.**

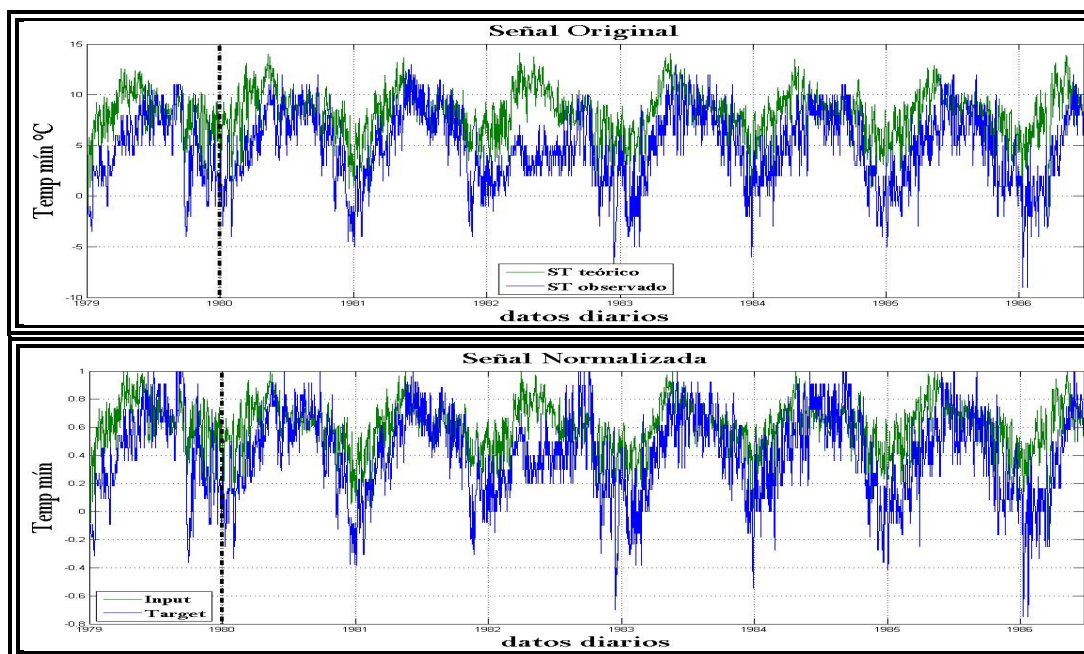


Figura 4.6 ST para Entrenar (**temp.mín.** estación de Apizaco).

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

ST Entrenada	# Neuronas en la C.O.	Tiempo de Entrenamiento (s)	E.P. 20% ST	E.A.	E.S.O.	E.S.C.
1979-1986	11	401.8743	1.6518	4.8495	3.7474	0.6029
	17	545.3456	1.7407	4.8892	3.7440	0.6290
1980-1986	6	238.7167	1.4692	4.0795	3.7687	0.6158
	9	259.8431	1.4458	3.9652	3.6345	0.6084
	10	287.3175	1.5419	4.3557	4.0854	0.6397

Tabla 4.8 Análisis del Aprendizaje (**temp.mín.** Estación de Apizaco).

El error, de la ST de entrada respecto al de objetivo, ambas normalizadas y sin haber pasado por el entrenamiento de la RNA (**ver figura 4.6**), es de 8.6390. La ST es desde 1979 hasta donde se tiene de 1986. Para la ST completa de 1988, normalizado, se tiene un error de 2.3510 (**ver figura 4.7**).

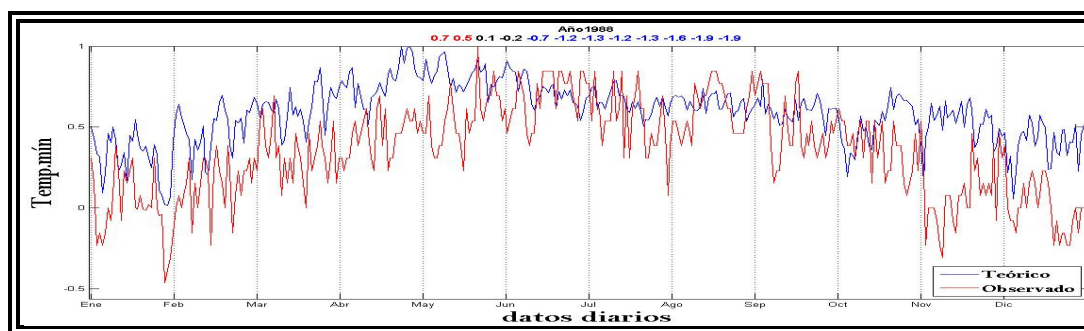
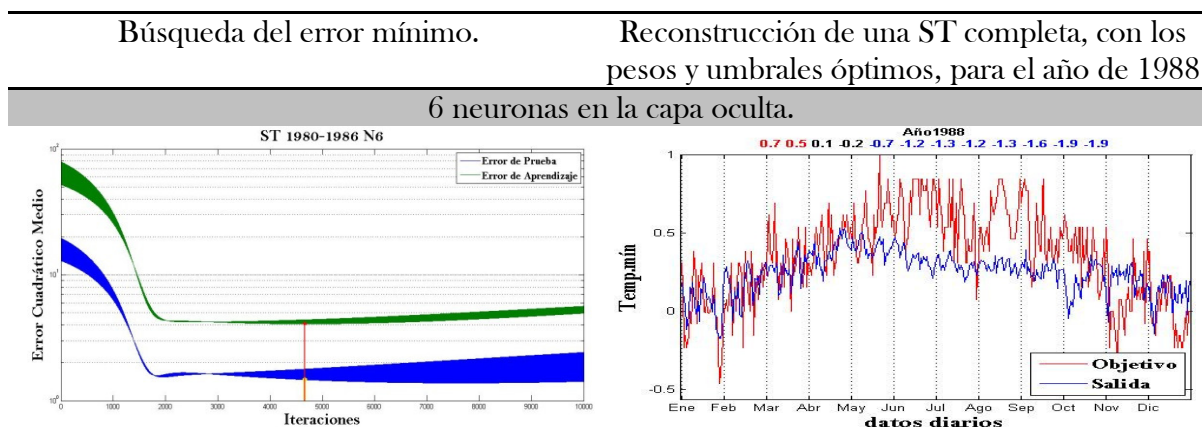


Figura 4.7 ST completa para probar los pesos y umbrales óptimos obtenidos en la RNA.

Mostrando en la tabla 4.9 las gráficas del error de aprendizaje y prueba, como la reconstrucción de la serie completa, de cada ST de entrenamiento con las neuronas utilizadas en la capa oculta para dicho proceso.



CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

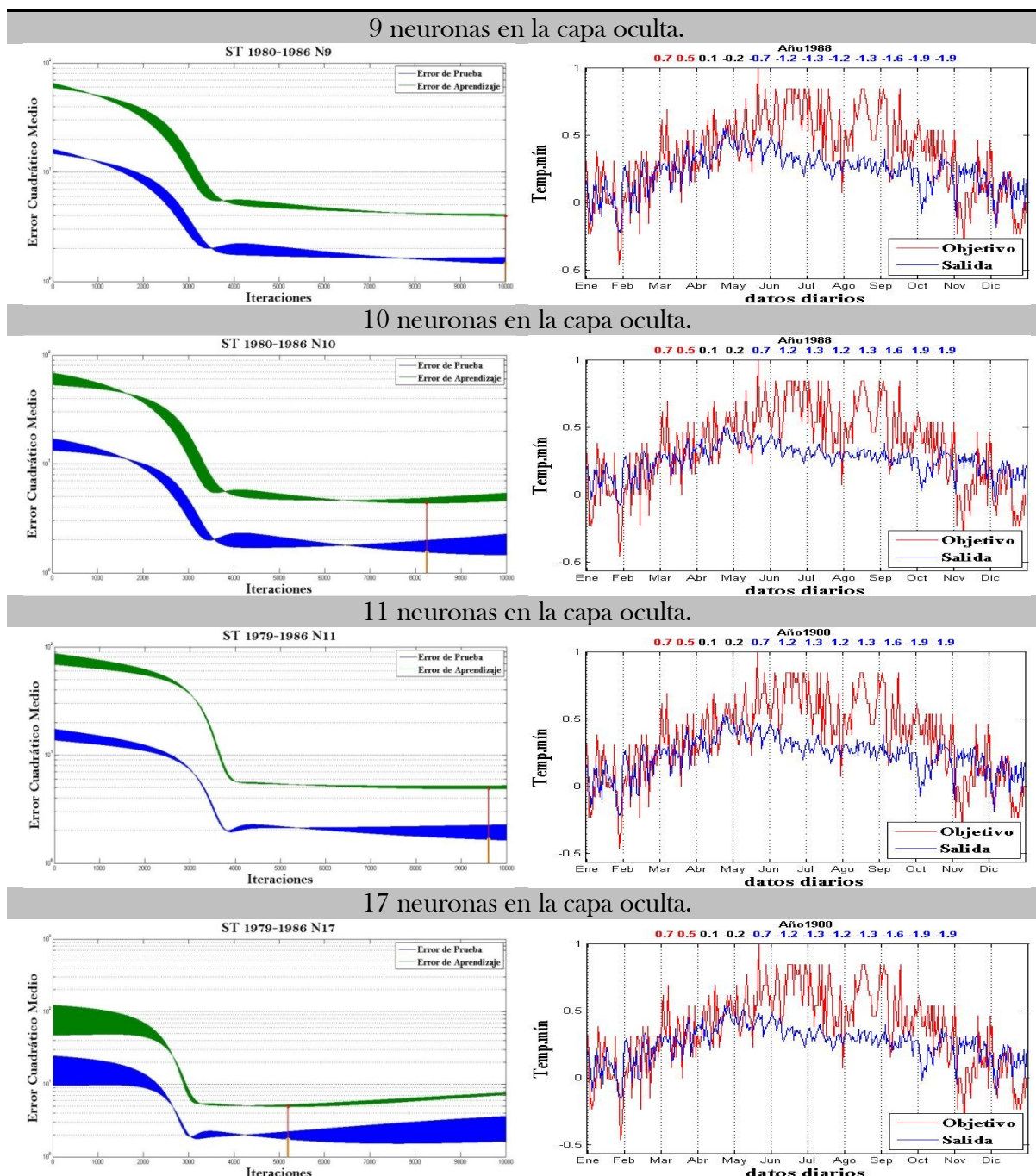


Tabla 4.9 Gráficas de Error y de Reconstrucción.

A continuación se muestran las señales generadas, para los años de datos faltantes, con los diferentes pesos y umbrales que se obtuvieron para cada número de neuronas; seguido de su reconstrucción del año correspondiente, tomando la de nueve neuronas para ésta.

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

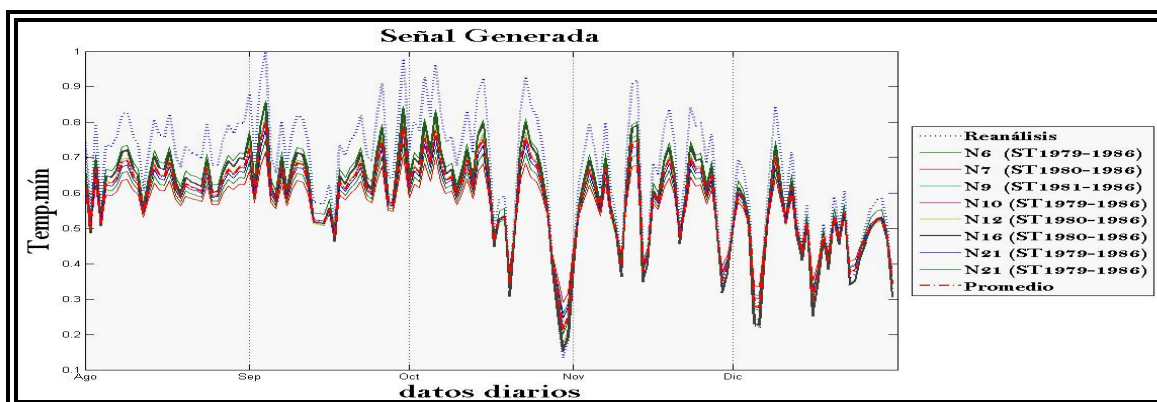


Figura 4.8 Datos Generados para el año de 1986.

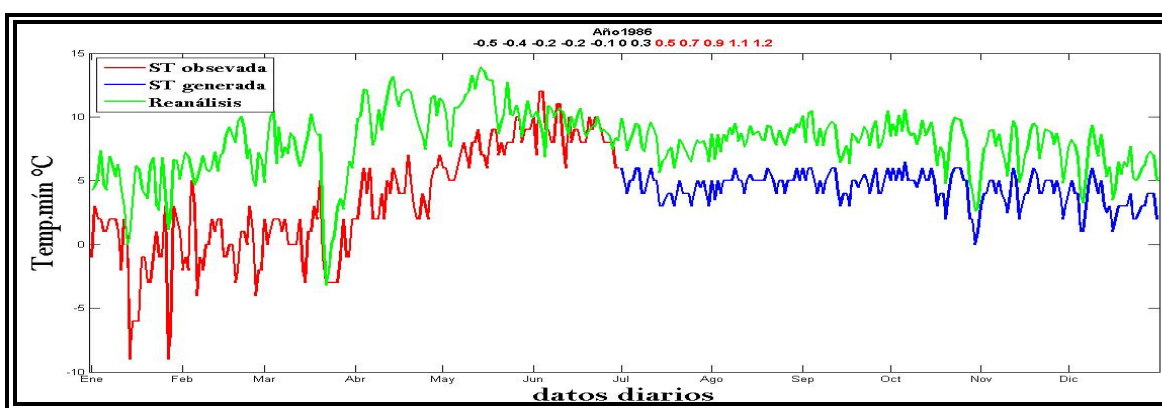


Figura 4.9 ST Reconstruida del año de 1986.

En la siguiente tabla, 4.10, se muestran los años de datos ausentes de un solo dato, reemplazándolos empíricamente y generados por los pesos y umbrales obtenidos en la RNA de nueve neuronas, como la previa reconstrucción.

	Datos Observados.	Datos generados empíricamente.	Datos generados con RNA
1987	NaN 3 3 1 2 2 1 1 2 2 4 Inicio del mes de Ene.	3	5.1267 - 5
	1 2 4 5 5 6 4 0 -2 -1 NaN Fin del mes de Dic.		
1991	0 1 2 4 3 2 2 3 3 4 Inicio de mes de Ene. 1992.	-0.5	3.0388 - 3

Tabla 4.10 Generación de los Datos Faltantes de 1987 y1991 (**temp.mín.** Estación de Apizaco).

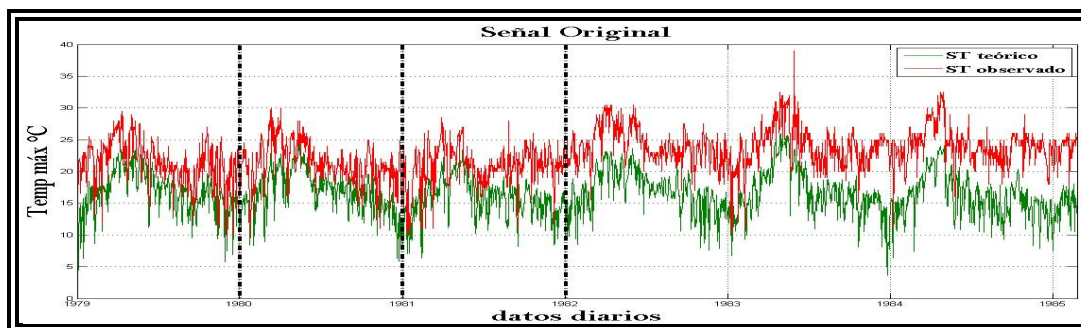
Para la **temp.máx.** y la **temp.mín.**, se mostrarán, para las siguientes estaciones, sus comparaciones de datos generados empíricamente y con la RNA, tal como esta estación.

4.3 ESTACIÓN DE TLAXCO.

Error Cuadrático Medio				
años	Temp.máx.		Temp.mín.	
	Normalización de la Estd.	Normalización de la ST	Normalización de la Estd.	Normalización de la ST
1979	0.3872	0.0509	0.7582	0.5296
1980	1.1484	0.0578	0.7641	0.4561
1981	2.1932	0.0878	0.7365	0.2963
1982	2.1127	0.0959	2.9630	0.6308
1983	1.1949	0.0665	0.7103	0.5511
1984	1.9801	0.1149	1.1060	0.5112
1985	NaN	NaN	NaN	NaN
1986	NaN	NaN	NaN	NaN
1987	1.2363	0.0845	NaN	NaN
1988	0.5770	0.0985	0.7673	0.5723
1989	NaN	NaN	NaN	NaN
1990	0.9836	0.1309	1.2796	0.4944
1991	0.5635	0.0158	NaN	NaN
1992	NaN	NaN	1.3894	1.1746
1993	0.8841	0.0794	2.1562	0.6655
1994	0.7320	0.0379	3.0369	0.4538
1995	1.1530	0.0629	0.7254	0.4659
1996	0.9259	0.0608	0.8433	1.0121
1997	6.0542	0.1728	0.9982	0.4991
1998	0.5724	0.0307	1.0472	0.4772
1999	0.6991	0.0431	1.4971	0.8170
2000	0.4524	0.0386	1.8171	1.3225
2001	1.9842	0.0339	1.7734	0.5057

Tabla 4.11 Análisis de errores de **temp.mín.** y **temp.máx.**

Para la **temp.máx.** se tiene la siguiente figura de la ST para entrenar, que va desde 1979 hasta finales de Febrero de 1985, reduciéndola año con año, por lo que quedarán las ST para entrenar mostradas en la tabla 4.12., y en ésta misma, se señala (color rojo) la ST y número de neuronas óptimas para la reconstrucción de los datos faltantes de los correspondientes años para esta estación de Tlaxco.



CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

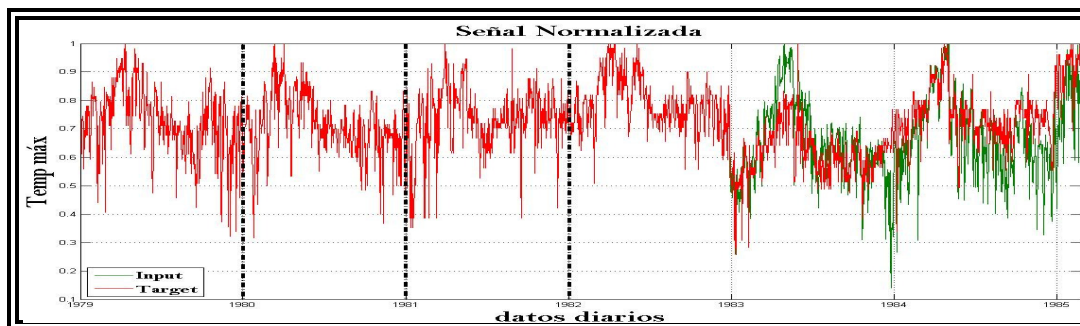


Figura 4.10 ST para Entrenar (**temp.máx.** estación de Tlaxco).

El error de la ST de entrada respecto al de objetivo, ambas normalizadas, es de 0.2224 (ST original, E.S.O.)

ST Entrenada	# Neuronas en la C.O.	Tiempo de Entrenamiento (s)	E.P. 20% ST	E.A.	E.S.O.	E.S.C.
1979-1985	15	392.571411	0.0335	0.1028	0.0818	0.0232
1980-1985	5	220.249266	0.0310	0.1030	0.0801	0.0194
	7	233.746476	0.0279	0.1015	0.0941	0.0267
	8	237.883500	0.0285	0.0978	0.0821	0.0236
	10	254.938497	0.0304	0.0987	0.0785	0.0225
1981-1985	4	153.654706	0.0252	0.0871	0.0844	0.0247
	8	199.802775	0.0269	0.0890	0.0789	0.0210
	16	266.278431	0.0250	0.0939	0.1043	0.0277
1982-1985	7	106.436118	0.0241	0.0668	0.0976	0.0289
	9	121.890954	0.0244	0.0657	0.0940	0.0289
	13	129.165423	0.0252	0.0631	0.0844	0.0270
	17	160.281018	0.0266	0.0626	0.0809	0.0270

Tabla 4.12 Análisis del Aprendizaje (**temp.máx.** Estación de Tlaxco).

La ST completa para probar los pesos y umbrales óptimos es del año de 1987, el error de esta ST de los datos observados respecto a los teóricos, ambas normalizadas, es de 0.0845.

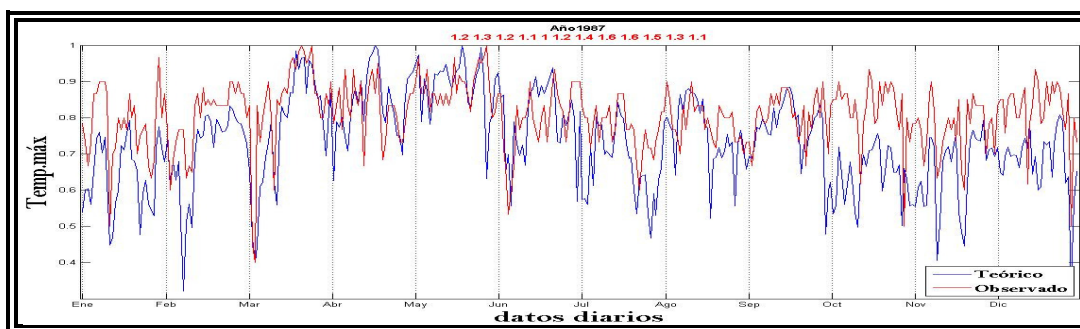


Figura 4.11 ST completa para probar los pesos y umbrales óptimos obtenidos en la RNA.

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

Para el año de 1985 se tiene ausencia de 31 datos del mes de Marzo, generando los siguientes:

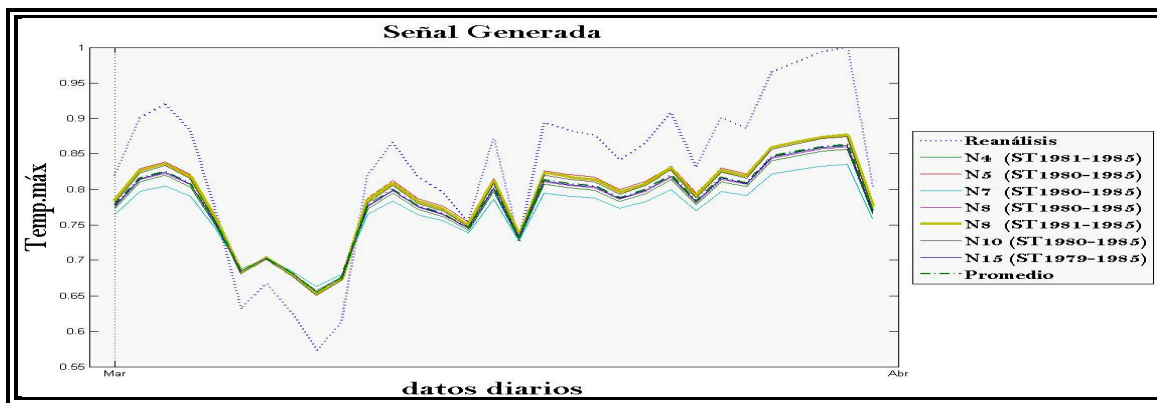


Figura 4.12 Datos generados de los faltantes del año de 1985.

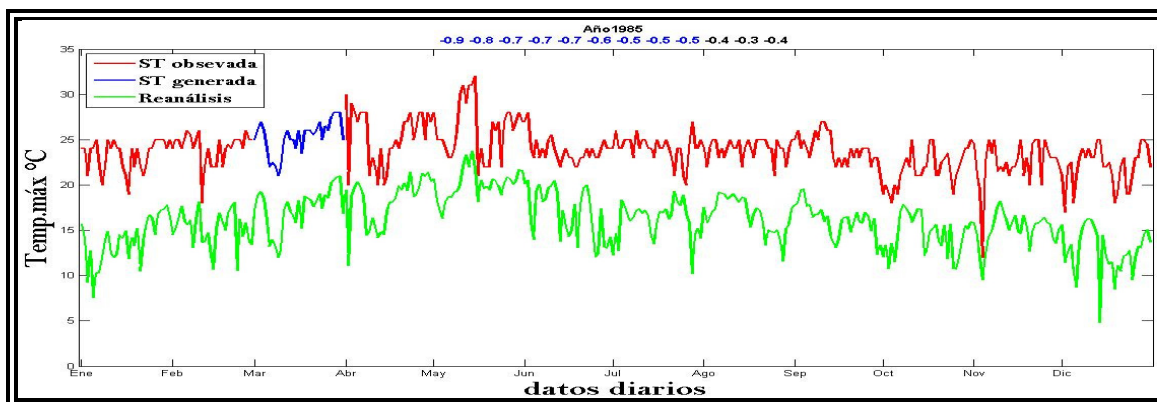


Figura 4.13 ST Reconstruida del año de 1985.

En 1986 se tiene ausencia de 184 datos del mes de Julio a finales del año, generando los siguientes:

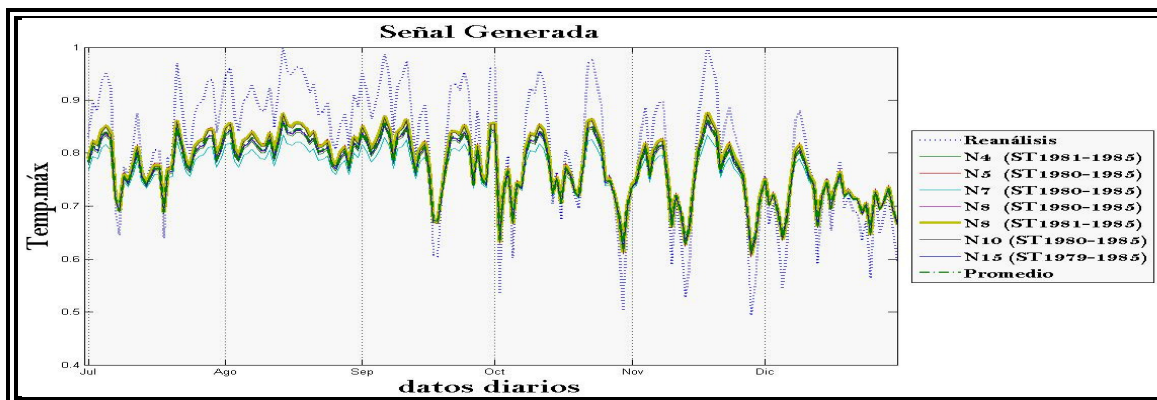


Figura 4.14 Datos generados de los faltantes del año de 1986.

CAPÍTULO 4.
EXPERIMENTOS NUMÉRICOS.

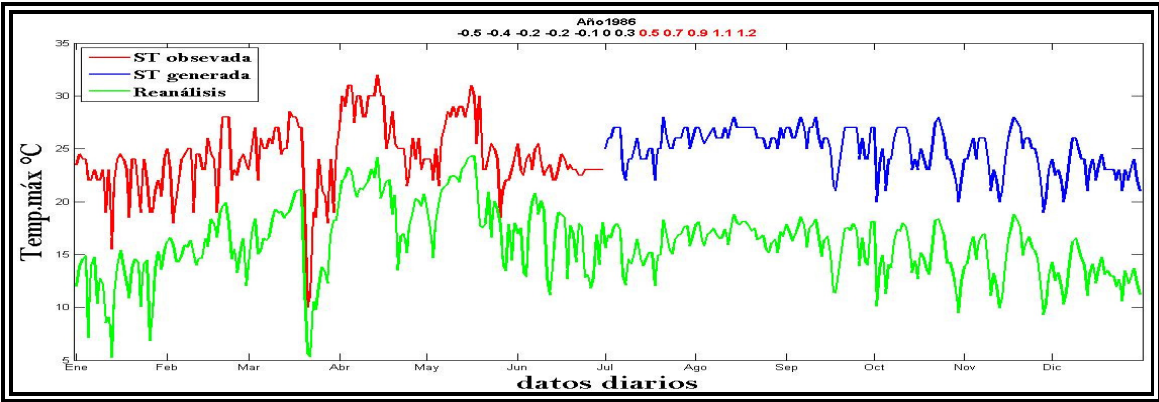


Figura 4.15 ST Reconstruida del año de 1986.

Para los años 1989 y 1992, hacen falta 4 y 1 datos respectivamente, reemplazando éstos empíricamente y generándolos con la RNA, tal como se mostró en la estación de Apizaco, para la generación de estos datos se utiliza los pesos y umbrales óptimos de 8 neuronas en la capa oculta, obtenidas en el entrenamiento de la ST 1981-1985.

	Datos Observados.	Datos generados empíricamente.	Datos generados con RNA
1989	24 26 24.5 26 18 16 21 23 23 27 NaN Fin del mes de Feb.		
	28.5 28 26.5 24.5 26 22 22 24.5 19 22.5 Inicio del mes de Mar.	28	24.8734 - 25
	24 26 24 26 24 25 25.5 28.5 29 25.5 Fin del mes de Mar.	26	23.7141 - 24
	NaN NaN NaN 26.5 21 25.5 25 26 28 26 24 26 24.5	25 26	25.0865 - 25 25.4143 - 25
	Inicio del mes de Abr.		
1992	23 26 25 22 28 21.5 24 23 21.5 24.5 NaN Fin del mes de Feb.		
	26.5 28 28 26.5 28.5 28 28 29 29 22.5 Inicio del mes de Mar.	25.5	22.4921 - 22.5

Tabla 4.13 Generación de los Datos Faltantes de 1989 y1992 (temp.máx. Estación de Tlaxco).

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

En la **temp.mín.** se tiene la siguiente ST para entrenar, que va desde 1979 hasta finales de Febrero de 1985, tal como la **temp.máx.** de esta estación, sólo que al reducir la ST para entrenar, mostradas en la tabla 4.14., se tienen solo dos series para entrenar, señalando, de misma forma como la variable anterior, el número de neuronas óptimas para la reconstrucción de los datos faltantes de los correspondientes años.

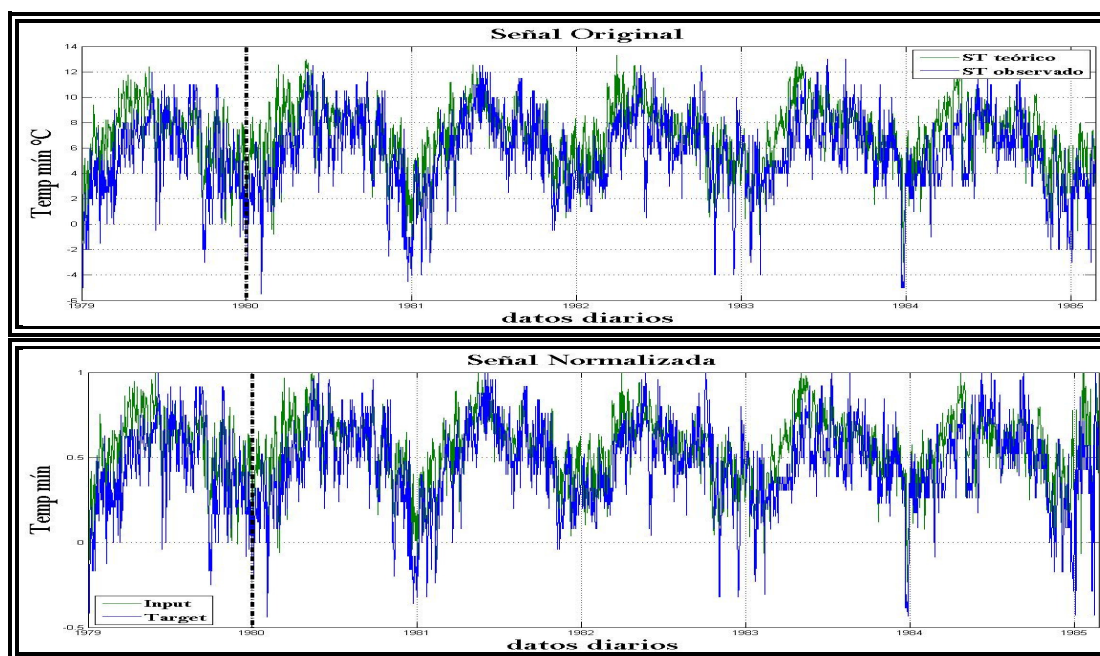


Figura 4.16 ST para entrenar (**temp.mín.** estación de Tlaxco).

ST Entrenada	# Neuronas en la C.O.	Tiempo de Entrenamiento (s)	E.P. 20% ST	E.A.	E.S.O.	E.S.C.
1979-1985	6	397.7765	1.2068	1.4620	1.4271	0.1572
	6	219.6130	1.1500	1.5133	1.4406	0.1302
1980-1985	8	177.3617	0.9510	1.3410	1.3870	0.1275
	12	209.3039	0.9489	1.3610	1.4102	0.1324
	17	253.5758	0.9308	1.3832	1.3932	0.1158

Tabla 4.14 Análisis del Aprendizaje (**temp.mín.** Estación de Tlaxco).

El error, de la ST de entrada respecto al de objetivo, ambas normalizadas (ver figura 4.16), es de 2.1321. Para probar los pesos y umbrales obtenidos en la RNA se utiliza la ST completa de 1992, normalizado, se tiene un error de 1.1882. A continuación se muestran las señales generadas, para los años de datos faltantes, con su reconstrucción de los años correspondientes, tomando la de ocho neuronas para esta variable y estación.

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

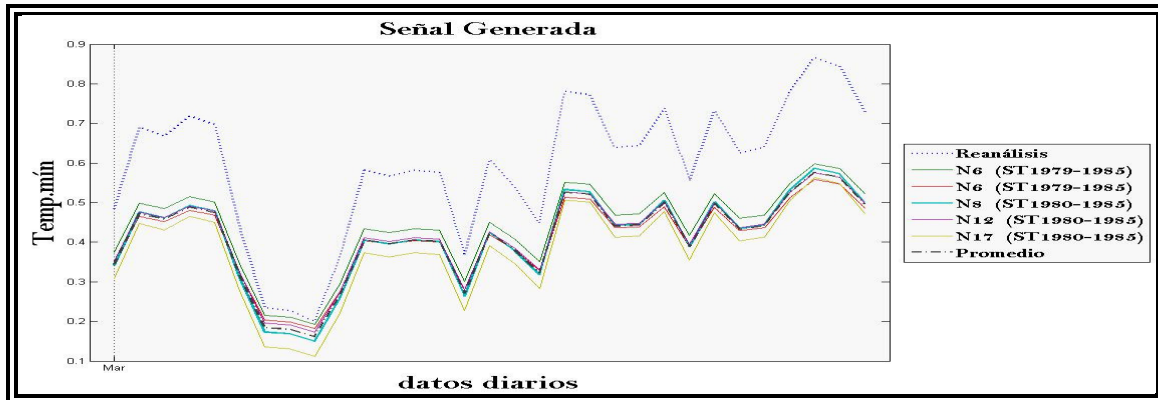


Figura 4.17 Datos Generados para el año de 1985.

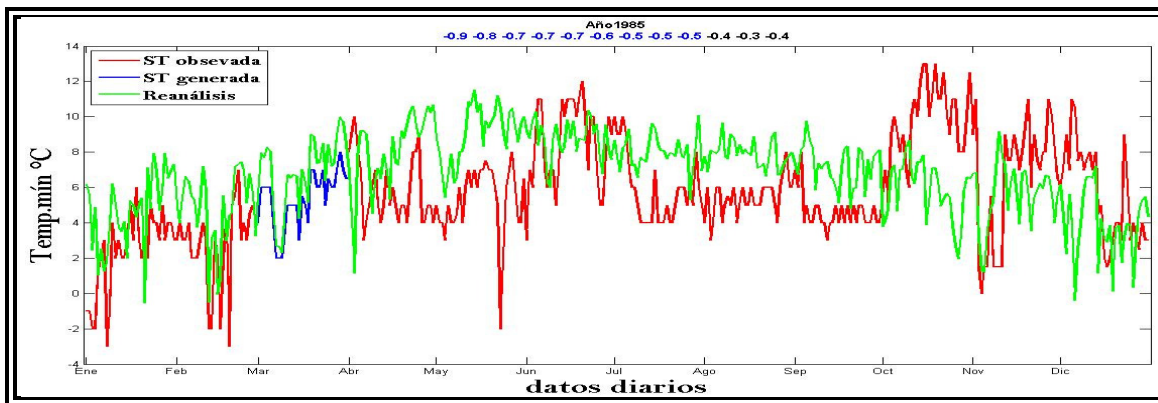


Figura 4.18 ST Reconstruida del año de 1985.

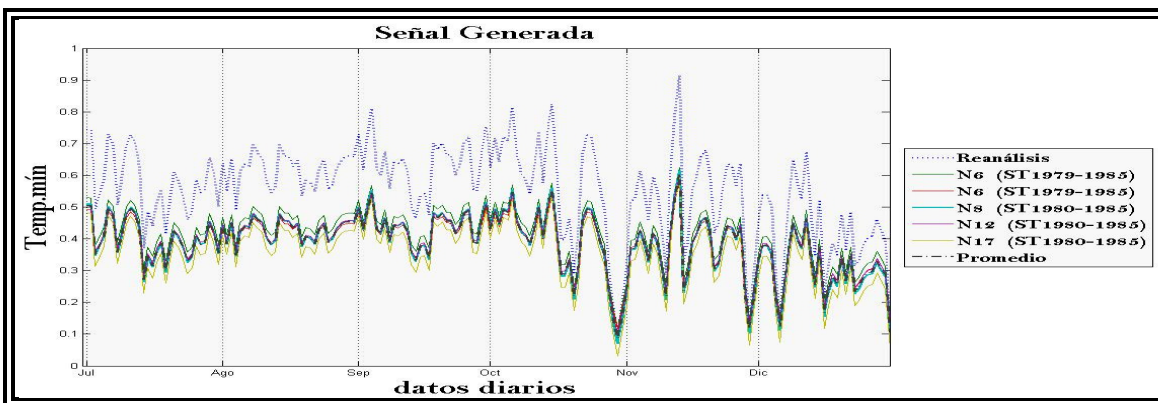


Figura 4.19 Datos Generados para el año de 1986.

CAPÍTULO 4.
EXPERIMENTOS NUMÉRICOS.

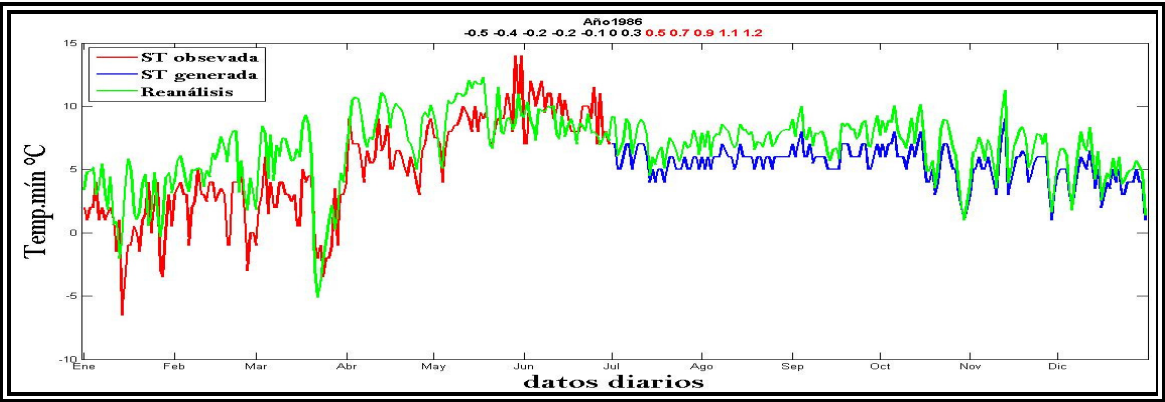


Figura 4.20 ST Reconstruida del año de 1986.

	Datos Observados.	Datos generados empíricamente.	Datos generados con RNA
1987	4 5.5 4 5 5.5 9 5.5 5.5 9 NaN		
	Fin del mes de Feb.		
	5.5 5 3 3 2 -6 -1 4 4	7	5.2793 - 5
1989	Inicio del mes de Mar.		
	4.5 5 4.5 4 4.5 3.5 3 5.5 8	4	4.4230 - 4
	Fin del mes de Mar.	5	4.8185 - 5
1991	2.5 NaN NaN NaN 7.5 4 2.5 2 4 6 4 5 7 6.5	6	5.3276 - 5
	Inicio del mes de Abr.		
	10.5 9.5 8.5 9 10 8 8 6 7 10.5		
1991	Fin del mes de Mayo.		
	NaN 6.5 7 6 6.5 5 7.5 9 9 8 10.5	8.5	6.5228 - 7
	Inicio de Jun.		

Tabla 4.15 Generación de los Datos Faltantes de 1987, 1989 y 1991 (temp.mín. Estación de Tlaxco).

4.4 ESTACIÓN DE ESPAÑITA.

Error Cuadrático Medio				
años	Temp.máx.		Temp.mín.	
	Normalización de la Estd.	Normalización de la ST	Normalización de la Estd.	Normalización de la ST
1979	0.6956	0.0279	0.9143	0.1595
1980	2.4192	0.0285	1.5162	0.1515
1981	1.0791	0.0293	0.3236	0.1149
1982	1.4208	0.0238	0.8084	0.1440

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

1983	1.0449	0.0363	0.7768	0.1651
1984	2.2611	0.0541	2.0511	0.3839
1985	2.7506	0.0296	2.1231	0.7646
1986	NaN	NaN	NaN	NaN
1987	NaN	NaN	NaN	NaN
1988	NaN	NaN	NaN	NaN
1989	NaN	NaN	NaN	NaN
1990	NaN	NaN	2.3589	1.4026
1991	NaN	NaN	NaN	NaN
1992	NaN	NaN	NaN	NaN
1993	3.4467	0.0718	NaN	NaN
1994	NaN	NaN	NaN	NaN
1995	1.4343	0.1263	1.1808	0.1543
1996	NaN	NaN	NaN	NaN
1997	2.0240	0.0440	3.6706	0.6378
1998	1.1388	0.0293	2.4670	0.5179
1999	1.0347	0.1013	1.9413	0.9647
2000	1.7732	0.0755	1.3925	0.3552
2001	2.2049	0.0239	2.2114	0.6094

Tabla 4.16 Análisis de errores de **temp.máx.** y **temp.mín.**

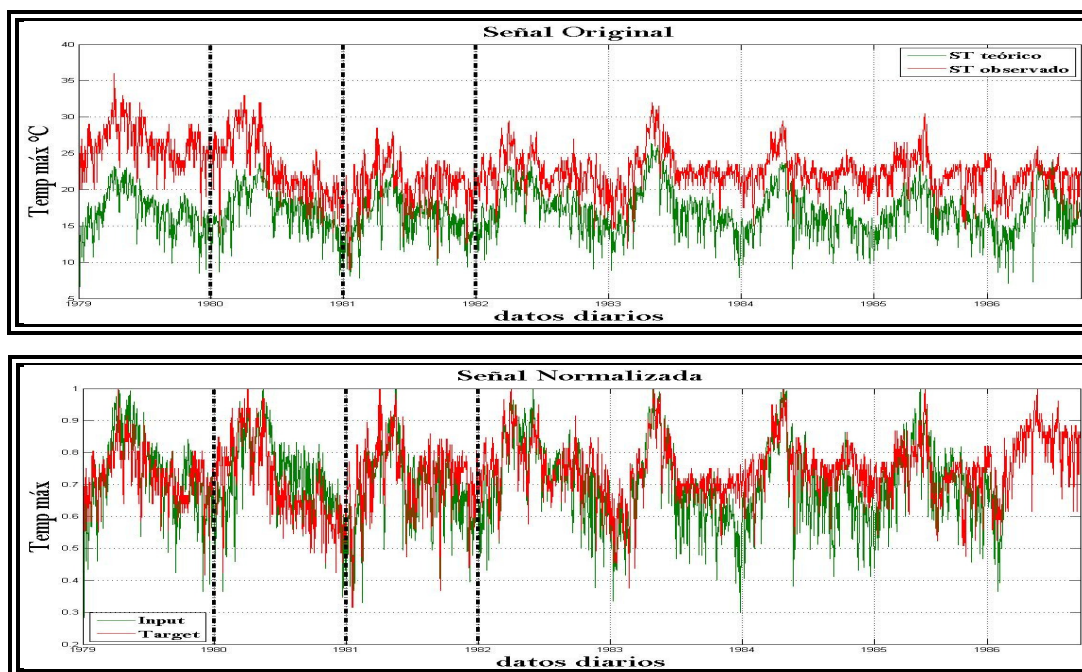


Figura 4.21 ST para Entrenar (**temp.máx.** estación de España).

Para esta estación meteorológica, se tiene una ST a entrenar de 1979 a Agosto de 1986 teniendo el siguiente registro para la **temp.máx.** mostradas en la tabla 4.17. El error de la ST de entrada respecto al de objetivo, ambas normalizadas, es de 0.2753. La ST completa para probar los pesos y umbrales óptimos obtenidos en la RNA es del año de 1993 (ver figura 4.22),

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

el error de esta ST de los datos observados respecto a los teóricos, ambas normalizadas, es de 0.0718.

ST Entrenada	# Neuronas en la C.O.	Tiempo de Entrenamiento (s)	E.P. 20% ST	E.A.	E.S.O.	E.S.C.
1979-1986	11	734.3480	0.0098	0.0194	0.0609	0.0104
	18	1006.1604	0.0103	0.0179	0.0636	0.0093
	23	584.4616	0.0098	0.0192	0.0609	0.0105
	24	540.5036	0.0105	0.0178	0.0637	0.0091
1980-1986	10	278.7656	0.0096	0.0185	0.0606	0.0091
	15	569.9898	0.0101	0.0179	0.0638	0.0087
	30	894.0254	0.0096	0.0181	0.0617	0.0092
1981-1986	11	405.3452	0.0082	0.0178	0.0581	0.0086
	16	674.8662	0.0082	0.0189	0.0583	0.0090
	19	593.5271	0.0096	0.0171	0.0634	0.0072
1982-1986	15	206.7163	0.0075	0.0147	0.0584	0.0074
	16	258.2546	0.0075	0.0146	0.0579	0.0076

Tabla 4.17 Análisis del Aprendizaje (**temp.máx.** Estación de España).

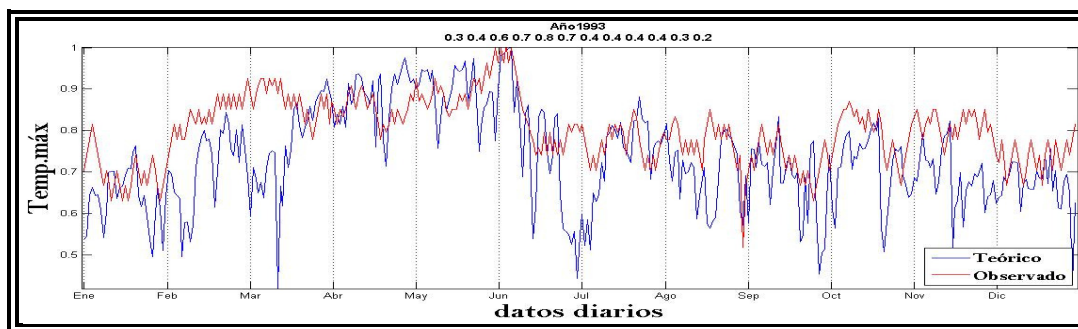


Figura 4.22 ST completa para probar los pesos y umbrales óptimos obtenidos en la RNA.

En 1986 hay ausencia de 153 datos a partir de Agosto a finales del año, generándolos y mostrándolos en la figura siguiente.

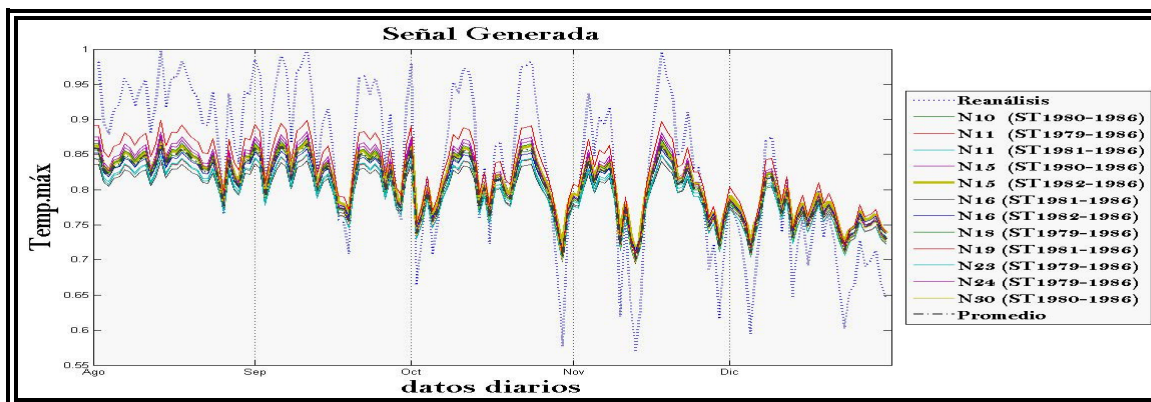


Figura 4.23 Datos generados de los faltantes del año de 1986.

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

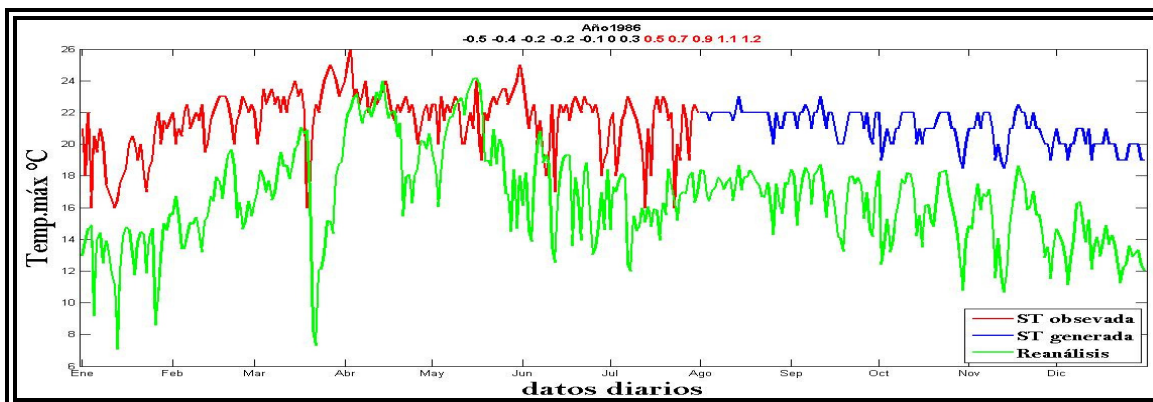


Figura 4.24 ST Reconstruida del año de 1986.

Para 1988 hay ausencia de 181 datos a inicios del mes de Julio a finales del año:

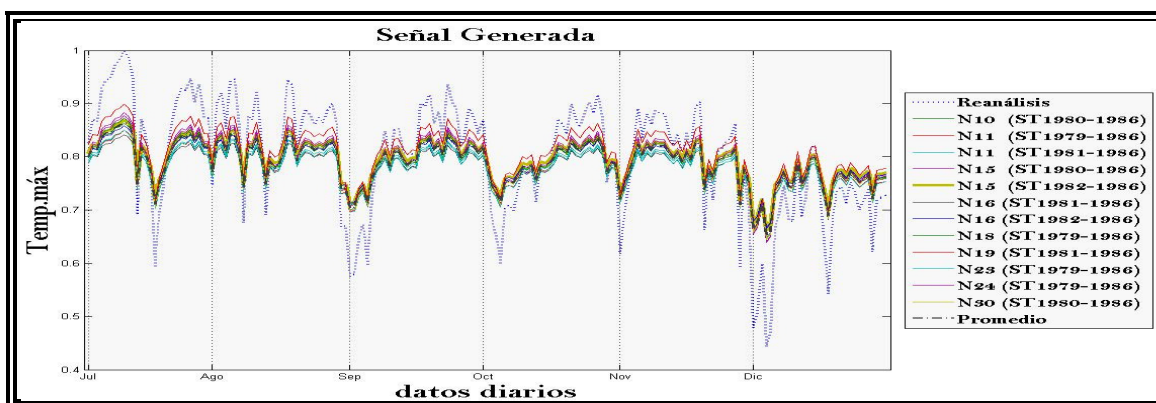


Figura 4.25 Datos generados de los faltantes del año de 1988.

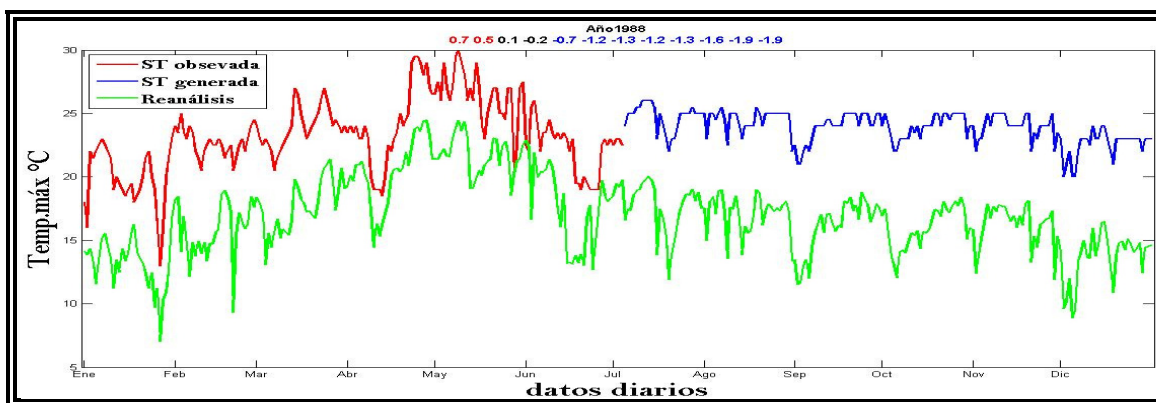


Figura 4.26 ST Reconstruida del año de 1988.

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

Generando 61 datos ausentes para inicios del año de 1989:

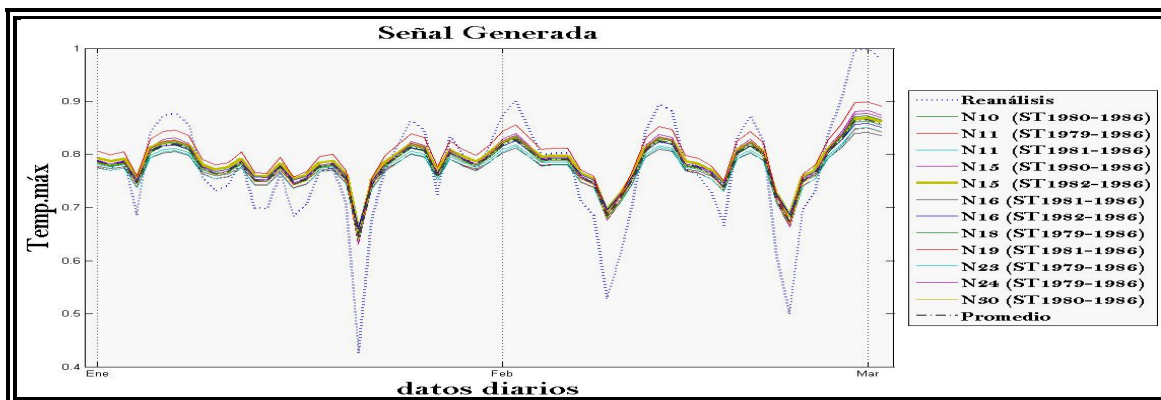


Figura 4.27 Datos generados de los faltantes del año de 1989.

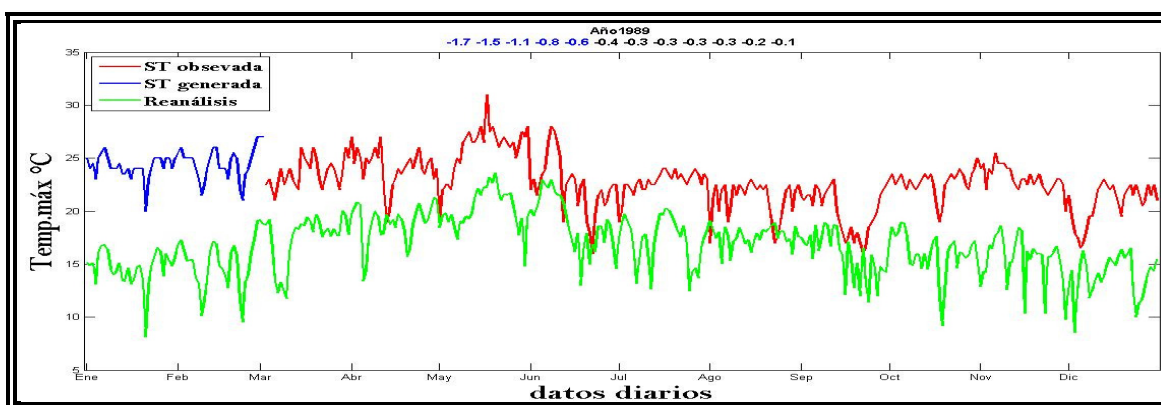


Figura 4.28 ST Reconstruida del año de 1989.

Siete datos seguidos hacen falta para 1992, a inicio del mes de Octubre:

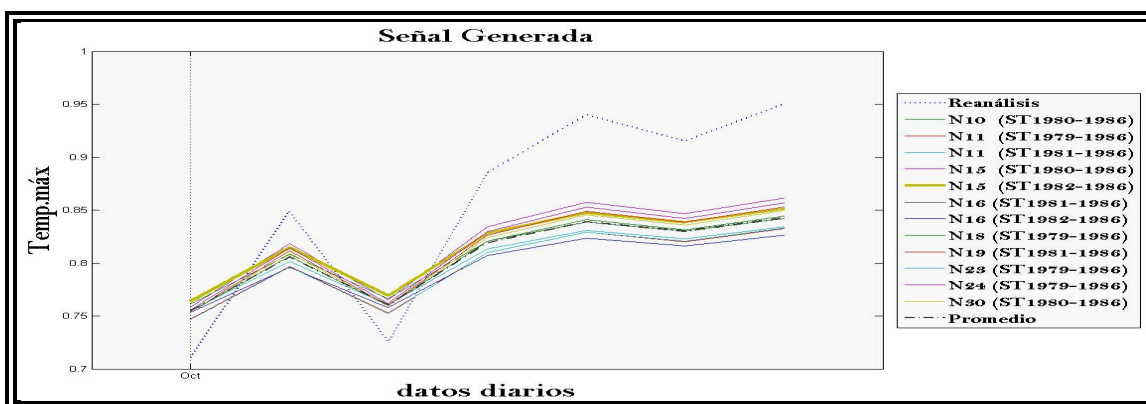


Figura 4.29 Datos generados de los faltantes del año de 1992.

CAPÍTULO 4.
EXPERIMENTOS NUMÉRICOS.

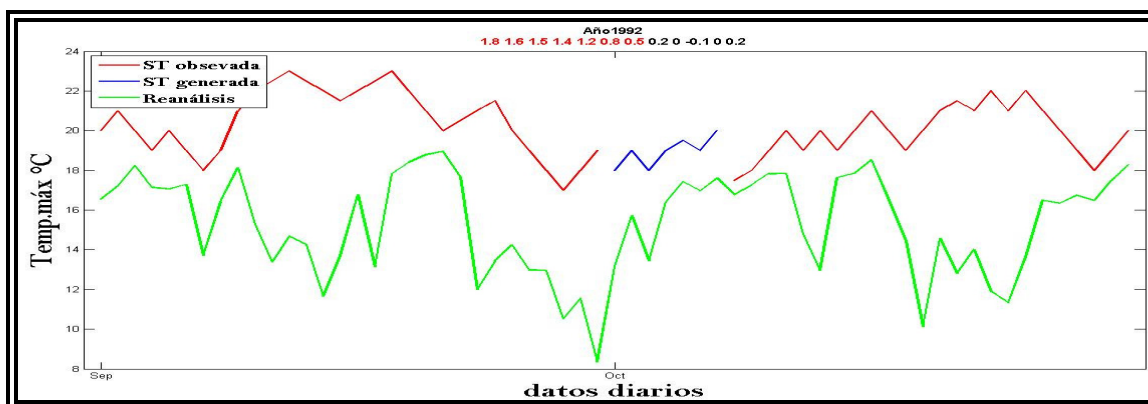


Figura 4.30 ST Reconstruida del año de 1992.

	Datos Observados.	Datos generados empíricamente.	Datos generados con RNA
1987	23 21.5 22.5 23 22.5 21.5 22.5 22.5 22.5 20	21	17.4509 - 17.5
	NaN NaN NaN NaN NaN 20 20.5 19.5 NaN	20	17.6004 - 18
	20.5 22 20.5 22 21.5 22	21	18.2733 - 18
	Fin del mes de Jul.	22	18.4420 - 18
	22.5 23 22 23.5 21.5 22 22.5 22.5 21.5 23 NaN	21	17.3291 - 17
	Fin del mes de Ago.	20	17.7489 - 18
1990	23.5 24 23.5 24 24.5 25 24.5 24 24.5 24	23	20.8336 - 21
	Inicio del mes de Sep.		
	22.5 22 22.5 23 23.5 23 23.5 23 23.5 24 NaN		
1991	Fin del mes de Sep.	24	20.5802 - 21
	23.5 23 23.5 23 23.5 22.5 22 22.5 23.5 22		
	Inicio del mes de Oct.		
1991	23.5 22.5 23.5 22.5 23 22.5 21.5 22 22.5 23		
	NaN 22.5		
	Fin del mes de Ago.	23	21.2137 - 21
1994	23 21.5 20 19 18 18.5 17 18 19		
	Inicio del mes de Sep.		
	22 23 23.5 24 24.5 23 23.5 24 23 22.5 NaN		
1994	22.5 23 23.5 24 24.5 25 24 24.5 25 25.5	22.5	21.0877 - 21
	Mes de Mayo.		
1996	22.5 23 23.5 23 23 23.5 23 23.5 22.5 23 NaN		
	23.5 23 22.5 23 23.5 23 23.5 23 23.5 23	23	19.0704 - 19
	Mes de Feb.		

Tabla 4.18 Generación de los Datos Faltantes de 1987, 1990, 1991, 1994 y 1996
(temp.máx. Estación de España).

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

En la **temp.mín.** se tiene una ST a entrenar de 1979 a finales de Julio de 1986.

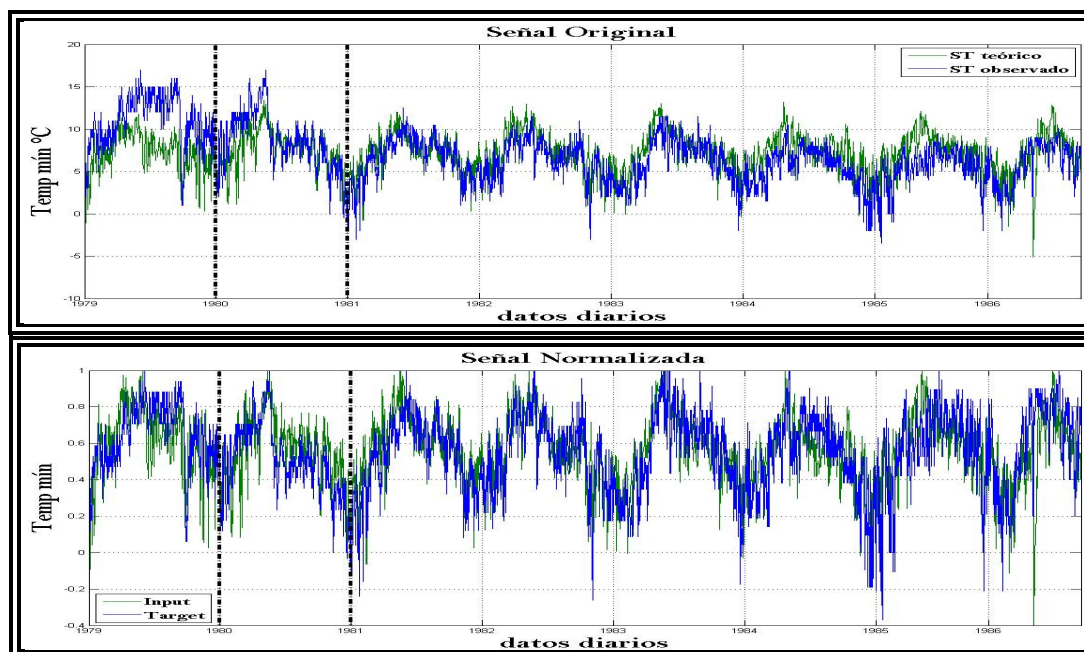


Figura 4.31 ST para entrenar (**temp.mín.** estación de España).

El error de la ST de entrada respecto al de objetivo, ambas normalizadas, es de 2.7598. La ST completa para probar los pesos y umbrales óptimos obtenidos en la RNA es del año de 1990, el error de esta ST de los datos observados respecto a los teóricos, ambas normalizadas, es de 1.4026.

ST Entrenada	# Neuronas en la C.O.	Tiempo de Entrenamiento (s)	E.P. 20% ST	E.A.	E.S.O.	E.S.C.
1979-1986	6	661.6941	1.1116	1.2473	1.9060	1.4038
	10	349.0997	1.2733	1.1389	1.9649	1.3299
	21	946.0172	0.5181	1.1888	1.8890	1.3857
	21	832.7382	1.3976	1.0981	2.0542	1.3473
1980-1986	7	245.5487	0.7935	1.3859	1.9383	1.4884
	12	294.6589	1.3395	1.1331	2.2031	1.2430
	16	351.8748	1.4821	1.1264	1.3245	1.2386
1981-1986	9	207.7414	1.2962	1.0620	2.1865	1.2420

Tabla 4.19 Análisis del Aprendizaje (**temp.mín.** Estación de España).

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

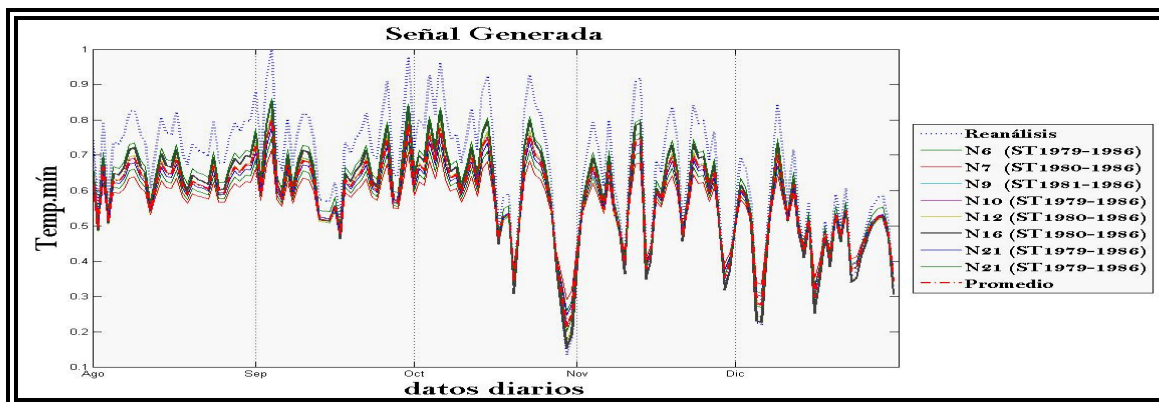


Figura 4.32 Datos Generados para el año de 1986.

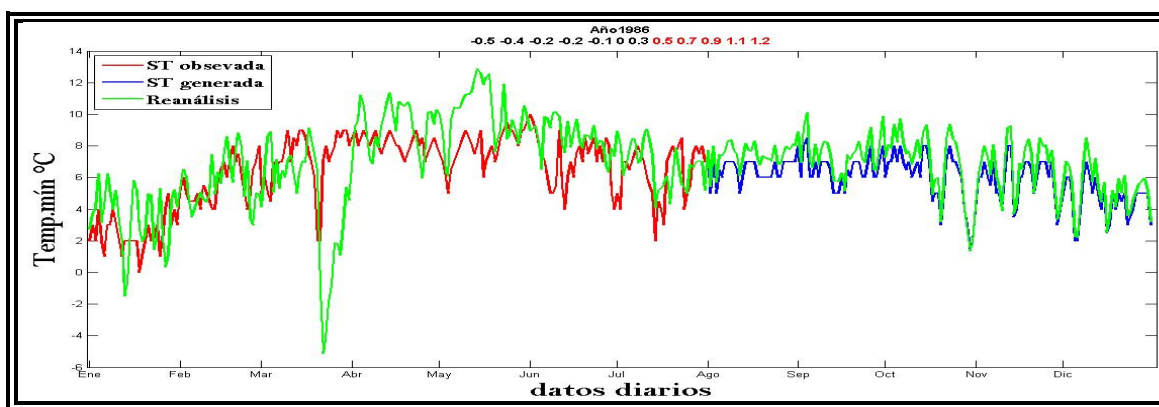


Figura 4.33 ST Reconstruida del año de 1986.

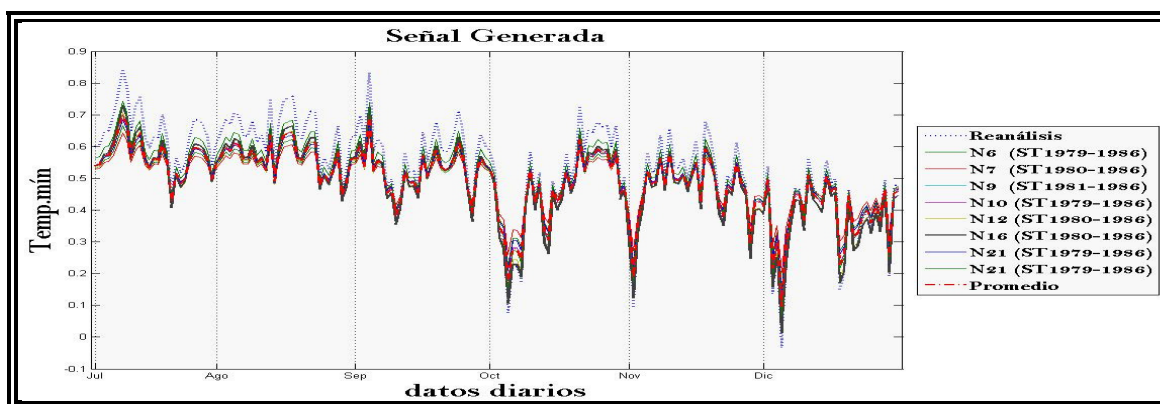


Figura 4.34 Datos Generados para el año de 1988.

CAPÍTULO 4.

EXPERIMENTOS NUMÉRICOS.

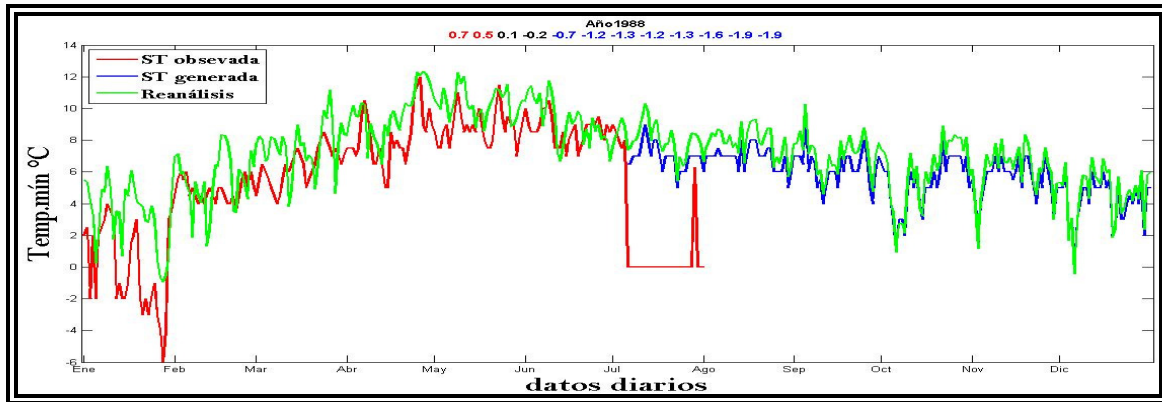


Figura 4.35 ST Reconstruida del año de 1988.

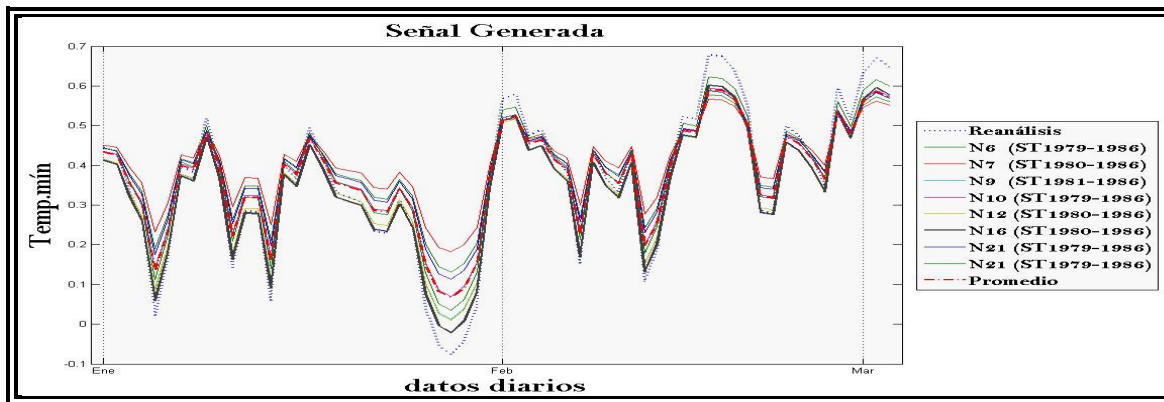


Figura 4.36 Datos Generados para el año de 1989.

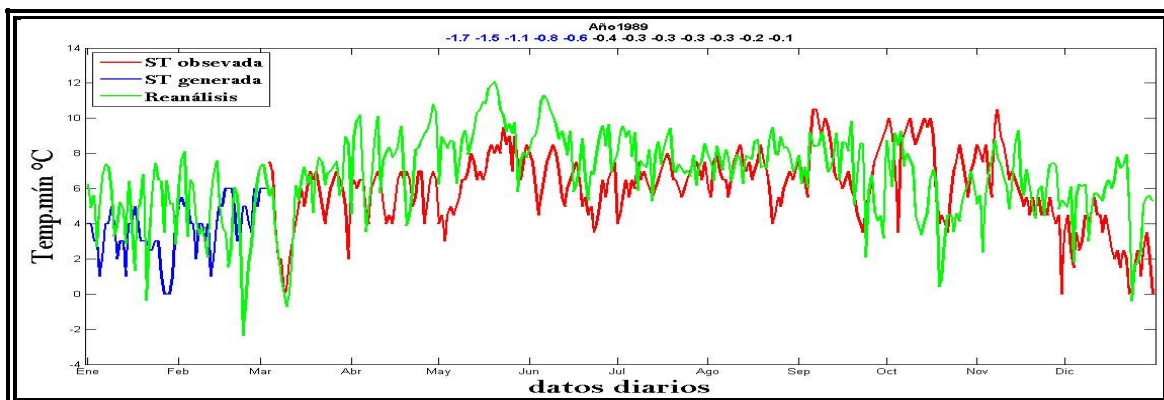


Figura 4.37 ST Reconstruida del año de 1989.

CAPÍTULO 4. EXPERIMENTOS NUMÉRICOS.

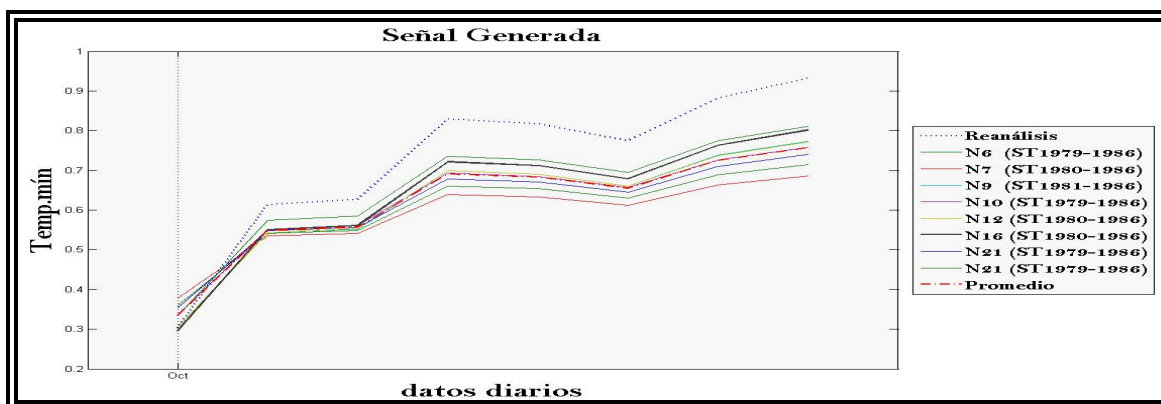


Figura 4.38 Datos Generados para el año de 1992.

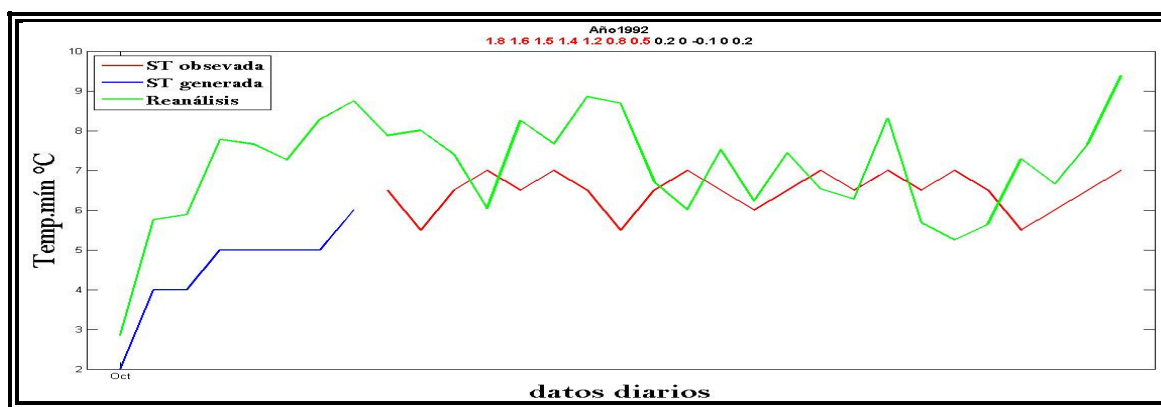


Figura 4.39 ST Reconstruida del año de 1992.

	Datos Observados.	Datos generados empíricamente.	Datos generados con RNA
1987	12 11 10 9 8.5 7.5 7 6.5 7 5 NaN NaN	5.5	8.2150 - 8
	NaN NaN NaN 9 8.5 8.5	6	8.5607 - 9
	Fin del mes de Jul.	7.5	8.3848 - 8
	7.5 8.5 10 9 8.5 8 7.5	8	10.2440 - 10
	Fin del mes de Ago.	8.5	7.4226 - 7
1991	7.5 6.5 7.5 6.5 7 6.5 7 5.5 6.5 7 NaN		
	Fin del mes de Ago.	6	6.9870 - 7
	6.5 6 7 5.5 5 5.5 5 5.5 5 5.5		
1993	Inicio del mes de Sep.		
	8 9 8 8.5 8 8.5 8 9 10 9		
	Fin del mes de Mar.	8	8.5962 - 9
	NaN 7.5 8.5 9 8.5 9 9.5 10 9.5 9 8.5		
	Inicio del mes de Abr.		

CAPÍTULO 4.
EXPERIMENTOS NUMÉRICOS.

1994	6 6.5 7.5 8 8.5 8 8.5 8 7.5 7 NaN 6.5 7 7.5 7 7.5 8.5 9 8.5 8 Mes de Mayo.	6.5	7.0825 - 7
1996	7 6.5 7 6.5 7.5 8 8.5 8 8.5 7.5 NaN NaN 7.5 8 7.5 8 7.5 8 7.5 8 7.5 8 Mes de Feb.	7.5	8.4418 - 8

Tabla 4.20 Generación de los Datos Faltantes de 1987, 1991, 1993, 1994 y 1996
(temp.mín. Estación de España).

A continuación, en el capítulo 5, se hará el análisis de los experimentos numéricos mostrados este capítulo, generando los datos faltantes con ayuda de los datos de reanálisis y la RNA.

CAPÍTULO 5

ANÁLISIS DE RESULTADOS.

En este trabajo, se descargaron ST de datos de reanálisis desde 1979 al 2001, 22 años seguidos, con la finalidad de experimentar si el uso de estos, son buenos para solucionar el problema de datos faltantes. Teniendo que, para futuras ST hidrometeorológicas observadas, que tengan ausencia de datos, se puedan generar con ayuda de una base mundial y con el método que ha sido empleado.

5.1 ANÁLISIS DE TEMPERATURAS MÁXIMAS

En los experimentos previos a este capítulo, se realizó un entrenamiento o aprendizaje de la RNA multicapa supervisada, llamada red de retropropagación, con el fin de encontrar un mínimo error y con ello la obtención de pesos y umbrales óptimos, para posteriormente generar los datos faltantes que se tienen para cada estación, de tal manera de que cada una de las estaciones y variables, tienen su propio entrenamiento en la RNA.

Los datos teóricos se utilizaron como serie de entrada, la cual se experimento al reducir la ST, por año, para tener diferentes registros en encontrar el mínimo error; observando que entre mayor número de neuronas en la capa oculta y una gran cantidad de datos de entrada, aumenta el tiempo de entrenamiento, también depende del número de iteraciones para repetir el proceso; al reducir el número de neuronas en la capa oculta y la cantidad de datos de entrada, disminuye el tiempo de entrenamiento.

5.1.1 Estación de Apizaco.

Se considera utilizar los pesos y umbrales óptimos, obtenidos con trece neuronas en la capa oculta, debido a los errores mínimos que se obtienen tanto en la prueba, como en una serie completa. El error de prueba se obtiene al aplicar la validación cruzada, al mismo tiempo que para obtener el error de aprendizaje al estar entrenando la RNA, sólo, que el error de

CAPÍTULO 5.

ANÁLISIS DE RESULTADOS.

prueba no pasa por el entrenamiento, son datos reservados para evitar el llamado sobreajuste; en la literatura, se apartan el 20% de los datos a entrenar (últimos datos de la ST de entrada); en los experimentos de esta estación, se consideró un número de 110 datos para probar la RNA, mismo número de datos de prueba para cada una de las reducciones en la ST de entrada, se aplicó de esta manera, para observar cómo afecta la cantidad de datos para entrenar, y comprobar que tan bueno fue el aprendizaje. La cantidad de 110 datos son el 20% de tres años, y sólo se consideró esta cantidad para esta estación y la variable de **temp.máx**, en las demás estaciones y variables, se respetó el 20% para el error de prueba, de cada tamaño de la ST a entrenar.

Se comprueban los pesos y umbrales óptimos que se obtuvieron en la RNA con una ST completa, la cual nos referimos a una ST de un año que no tenga ausencia de datos, teniendo tanto los datos teóricos como observados. Esta comprobación se realiza para cada estación y con ello poder elegir, con cual cantidad de neuronas en la capa oculta, se realizará la generación de los datos.

Se hace una reconstrucción de toda la ST, considerada desde 1979 hasta el año de 1986, este último es hasta donde se tiene registro, que es hasta el mes de Junio. Esta ST original, es la que se considera para ser entrenada, y se hace esta reconstrucción para saber que tan bueno o que tan semejante ha sido esta. Al reducir la ST para entrenar, podemos tener un mismo error al hacer su reconstrucción de la ST original, aunque se haya realizado un entrenamiento con cantidades diferentes de datos para entrenar, por lo que podemos obtener un error mínimo óptimo sin tener una gran tamaño en la ST de entrada.

La ausencia de uno a cinco datos seguidos, se pueden reemplazar empíricamente, y es que se tienen datos antes y después del ausente, por lo que podemos saber cómo se pudo comportar la temperatura. También se han generado con ayuda de la RNA, por lo que se presentan estas dos formas; comparando ambos métodos, teniendo como resultado una generación de datos idénticos, por lo que podemos decir que el número de neuronas escogida para la generación de datos, ha sido buena, por lo que podemos confiar, en una primera estancia, que son buenos los datos que se generaron con la ayuda de esta RNA, pues ambos métodos coinciden con el resultado que han dado, para esta variable y estación.

5.1.2 Estación de Tlaxco.

Para esta estación, se consideró que la ST de entrenamiento de 1981 a 1985, con un número de ocho neuronas en la capa oculta, es la mejor para generar los datos ausentes, debido a los errores mínimos obtenidos al probar con una ST completa, como en la reconstrucción de la ST original.

Hay datos generados empíricamente, hasta un máximo de tres datos faltantes seguidos. Generando los datos ausentes con la RNA, como empíricamente, se obtienen semejantes valores, algunos idénticos, o bien hay una diferencia de $\pm 3^\circ$ aproximadamente; por lo que se considera ajustable para la ST, en otras palabras, ambos métodos son buenos aunque tengan diferentes valores, pues se puede presentar los dos comportamientos de cada método.

Para esta estación de Tlaxco, al observar los errores mínimos obtenidos, se tiene que no son menores a diferencia de la estación de Apizaco, ya que esta última reduce el error a un número de milésimas para las de prueba y de una serie completa, mientras que los valores de error de Tlaxco llegan a centésimas; esto se puede explicar por los puntos de malla que se encuentran los datos teóricos, recordando que la base de reanálisis son de una escala regional y lo estamos direccionando a un punto local, que son los datos observados de las estaciones meteorológicas. Aunque el error no sea en milésimas como las obtenidas para la estación de Apizaco, el error obtenido para Tlaxco es bueno, ya que se reduce en comparación de las ST teóricas respecto a las observadas.

5.1.3 Estación de Españita.

Para esta estación, tenemos una buena disminución del error, pues se presentan valores de milésimas, tal como Apizaco, pero Españita reduce un poco más su error.

Solo fue necesario una ST de menos de cuatro años para entrenar y obtener un mínimo error, para poder generar los datos faltantes con los pesos y umbrales óptimos, obtenidos en la serie de entrada de 1982 a 1986, este último año se tiene registrado hasta el mes de Julio.

Analizando el tamaño de las ST para entrenar, en las tres estaciones meteorológicas, de la variable de **temp.máx.**, no es necesario tener una gran cantidad de datos, pero, si las suficientes para que la RNA aprenda adecuadamente, y es debido a que se ha utilizado una

CAPÍTULO 5.

ANÁLISIS DE RESULTADOS.

RNA supervisada, con ayuda de los datos de reanálisis para entrenar y los datos observados como targets (objetivos). Por lo que se podría tener una ST de 4 a 5 años seguidos para entrenar y obtener un error mínimo para la generación de datos; si esta ST fuese menor, por ejemplo, de tres años, el error no disminuiría y empezaría a oscilar en busca de un mínimo global; si fuese más amplia, como se tiene registrado en los experimentos numéricos, tampoco disminuiría y se quedaría estancado en un mínimo local.

Finalizando el análisis para la estación de Españita, la comparación de los datos generados con la RNA y empíricamente, se tiene una diferencia hasta de 4° , los valores generados por la RNA están por debajo a los obtenidos empíricamente, por lo que se tendría que observar el comportamiento que tengan ambos métodos en el resto de la ST (valores anteriores y posteriores al dato faltante), pero podemos decir que ambos métodos son válidos, aunque esta diferencia se presenta en tres datos, sabemos que se puede presentar una disminución o aumento de la temperatura súbitamente.

5.2 ANÁLISIS DE TEMPERATURAS MÍNIMAS

Para la **temp.mín.**, en todas las estaciones, se tiene un error más grande respecto a la **temp.máx.** Con ayuda del método de RNA, podemos disminuir el error del que se tiene de los datos teóricos respecto a los observados, tal como se consiguió en la de **temp.máx.**

5.2.1 Estación de Apizaco.

Las ST de entrada se tienen de 1979 a 1986, y de 1980 a 1986; sólo se vario una reducción de un año, puesto que si reducíamos mas la ST de entrada, oscilaba demasiado el error y no disminuía. Se tiene esta oscilación debido a que al normalizarla, se muestra la sobrestimación de los teóricos respecto a los observados, esto descrito en el capítulo 2, pero en unas partes de la ST se presenta lo contrario, una subestimación, por lo que la ST de entrada presenta ambos comportamientos respecto al objetivo (ver figura 4.6), causando las oscilaciones al ser entrenada y no tener una mayor disminución en el error como los obtenidos en la **temp.máx.**

Teniendo solo dos ST de entrada, mostradas en la tabla 4.8, y haciendo las comprobaciones de una serie completa como la reconstrucción de la serie original, se elige los pesos y

CAPÍTULO 5.

ANÁLISIS DE RESULTADOS.

umbrales óptimos obtenidos en la serie de entrada de 1980-1986 de nueve neuronas en la capa oculta, para la generación de datos ausentes de esta estación y de esta variable. También se presentan datos generados empíricamente, mostrando diferencias entre los resultados de ambos métodos, pero de la misma forma que la **temp.máx.**, ambos resultados se pueden presentar, por lo que ambos métodos son buenos y los datos generados por la RNA son confiables.

5.2.2 Estación de Tlaxco.

De igual manera que para Apizaco, se tiene solo una reducción de la ST a entrenar, por las mismas razones presentes en las ST. Escogiendo con la de ocho neuronas en la capa oculta, para la generación de datos, mostrando resultados muy similares en los reemplazados empíricamente, con una diferencia de 2°, teniendo una mayor seguridad en que son buenos los datos generados por la RNA.

5.2.3 Estación de Españita.

En la generación de datos para esta estación se tomo la de dieciséis neuronas en la capa oculta, ya que los errores obtenidos en una serie completa y en la reconstrucción de la serie original son las más bajas en comparación de los demás experimentos; los pesos y umbrales óptimos se obtuvieron en la ST de entrada de 1980 a 1986.

Las diferencias que hay en los datos generados empíricamente y con la RNA son de hasta 3°, teniendo gran similitud en los resultados obtenidos para ambos métodos, lo cual podemos encontrar ambos comportamientos para la temperatura y decir que los datos generados por la RNA son buenos y confiables.

Generando de esta manera los datos ausentes en las ST para las estaciones meteorológicas de las variables de temperatura, y esto, se pudo realizar con la ayuda de los datos de reanálisis, ya que con ellos se pudo generar las ST de hasta más de medio año, y que estos datos generados sean buenos y confiables.

CAPÍTULO 6

CONCLUSIONES Y TRABAJOS FUTUROS.

4.1 CONCLUSIÓN

Para este trabajo de tesis, el objetivo principal, o bien la problemática a resolver, fue la ausencia de datos en las ST de las estaciones meteorológicas, y con ayuda de una base mundial, se pudo lograr esta generación, debido a que se presentan una gran cantidad de datos faltantes seguidos, con aproximadamente medio año de ausencia, por lo que en principio se pensaría que para poder generar o rellenar las ST incompletas, se tendrían que tener una gran cantidad de datos; por ejemplo, para hacer pronóstico de tiempo-clima, se tiene considerar un periodo de 10 años para pronosticar de uno a dos datos futuros, por lo que al generar más de este número de datos, se ve algo imposible; pero la ventaja que tenemos, es la base de reanálisis.

¿Qué tan bueno es esta generación de datos faltantes? Como hemos visto en un inicio (capítulo 2), la tendencia o el comportamiento que tienen los datos teóricos con los observados, son muy similares, por lo que nos es muy favorable; en un primer plano, podríamos sustituir los datos ausentes con los datos de reanálisis, y en segundo, explotar hasta un máximo estos datos teóricos con una herramienta muy poderosa, que son las RNA, lo cual se ha reducido el error que se tiene en ambas bases, una respecto la otra, por lo que podemos hacer un gran uso de esta herramienta y resolver la problemática en cuestión.

Los diferentes trabajos que se han realizado mediante las RNA, su objetivo principal es el predecir, lo cual, comúnmente se ha de hacer para un solo dato futuro, trabajos como Eric Wan (1993) hacen este tipo de predicción. El trabajo de Tomás Casillas Lavín (2004), hace referencia sobre Wan, tomando en cuenta algo muy relevante llamado *límite de predictibilidad*, este concepto se es manejado en los trabajos de Sistemas Dinámicos No Lineales, Teoría del Caos y Geometría Fractal, todos estos en un conjunto, como lo hace Daniel Peña Maciel (2012). Este concepto de límite de predictibilidad, es hasta cuántos

CAPÍTULO 6.

CONCLUSIONES Y TRABAJOS FUTUROS.

puntos o datos puedo predecir, en otras palabras, es un número de datos confiables que se puede tener al hacer la predicción. Por lo que podemos caer en una contradicción para este trabajo de tesis, si se tiene un límite de predictibilidad, así que ¿podríamos confiar que los datos generados son buenos, si se tiene una gran cantidad de datos ausentes?; se ha mencionado que, la gran ventaja que se ha tenido son los datos de reanálisis, por el comportamiento y la tendencia semejante a los datos observados, además, podemos decir que son realmente buenos los datos generados, debido al aprendizaje que se obtuvo en la RNA, y esto se muestra en las gráficas de error del capítulo de experimentos numéricos, ya que se tienen tanto el error de aprendizaje y de prueba, observando que el error de prueba es menor al de aprendizaje, cuando en la teoría, descrita en el capítulo tres (ver figura 3.12), se tiene que el comportamiento de los errores debería ser lo contrario, y se muestran de esta manera ya que se está comprobando que tan bueno ha sido el entrenamiento de la RNA, aun con esta prueba de la validación cruza se comprueban nuevamente con una ST completa de un año, mostrando que el método empleado es bueno y confiable, no negando que aún exista un error, pero mostramos que con la RNA se pudo reducir el error de entrada obteniendo una salida aproximadamente exacta al objetivo.

El número de neuronas implementadas en la RNA, es de gran importancia y cuestionamiento; la capa oculta representa la No Linealidad del método, por tener una función de transferencia no lineal tal como su derivada. En este trabajo, se utilizó una función de transferencia sigmoidea en la capa oculta y una función lineal en la capa de salida, recordando que si no se hubiese implementado una capa oculta sería un perceptrón simple. La cantidad de neuronas y capas, ayuda a la RNA para la obtención de un error mínimo óptimo, que es el tiempo que tarda en su entrenamiento, es evidente que entre mayor número de capas y de neuronas, el proceso tarda más, pero la cantidad de estos parámetros es para encontrar una función de aproximación para la disminución del error de la salida de la RNA con respecto al objetivo. Se obtuvieron diferentes cantidades de neuronas en la capa oculta para el entrenamiento de cada estación porque cada zona tiene diferente comportamiento, por ejemplo, Españita llueve mucho más comparado con las otras dos estaciones, teniendo una mayor cantidad de neuronas; por lo que podríamos tener una influencia con el comportamiento que presentede la zona que se esté trabajando.

CAPÍTULO 6.

CONCLUSIONES Y TRABAJOS FUTUROS.

Saber la cantidad de datos que se requieren para entrenar y poder implementar el buen uso de la RNA para la predicción, en este caso, para la generación de datos, se tiene que el tener suficientes datos sirve para tener un buen aprendizaje, en este caso varió tanto para la temperatura mínima como la máxima, pero se encontró que el tener una cantidad mayor de datos, no reducía su error y se estancaba en un error mínimo local, pensando que se estuviera sesgando la ST de entrada o bien provocando una mayor incertidumbre, en cambio, el tener una ST más corta, oscilaba demasiado el error y no lograba encontrar un mínimo ni tener una disminución de éste, por lo que el estar experimentando y estar reduciendo las ST, ayudó a encontrar un mejor rendimiento para el entrenamiento de la RNA, por lo que no fue necesario sobrepasar la cantidad de más de cinco años de datos diarios, un punto muy favorable para el tiempo del procesamiento que realizó, teniendo algo evidente, que entre mayor número de datos, se invierte un mayor tiempo de procesamiento.

4.2 TRABAJOS FUTUROS

Para dar seguimiento a este modelo de RNA, es la optimización del algoritmo, la cual, es llegar a minimizar a un mas el error, y se lograría aplicando *algoritmos genéticos*, el uso que tendría éste es para mejorar el entrenamiento de la RNA.

El ambiente de MatLab es muy amigable al programar, además de contar con la interfaz gráfica, que es de gran utilidad; pero, teniendo en cuenta los tiempos de entrenamiento que son bastante amplios, sería más óptimo trabajar en un ambiente de programación como Fortran o C++, el cual se reduciría el tiempo de ejecución.

Actualmente, quienes trabajan con RNA, lo hacen a tiempo real, en otras palabras, mientras registran o adquieren los datos de las variables, éstas se procesan al momento dando así una predicción o ajustando la que ya se tenía. Ejemplos de estos trabajos se realiza en España, la cual hacen predicciones para el contenido de esporas en el aire; Argentina lo hace con la precipitación; México, realizará un proyecto de patrón de escurrimiento en las cuencas. Por lo que la utilización de las RNA, es de un gran interés aplicada en diferentes áreas de estudio, se han realizado trabajos en el área de la sísmica, registro de pozos, en la medicina, economía, robótica, física espacial, entre otras. En el área de meteorología se ha aplicado con

CAPÍTULO 6.

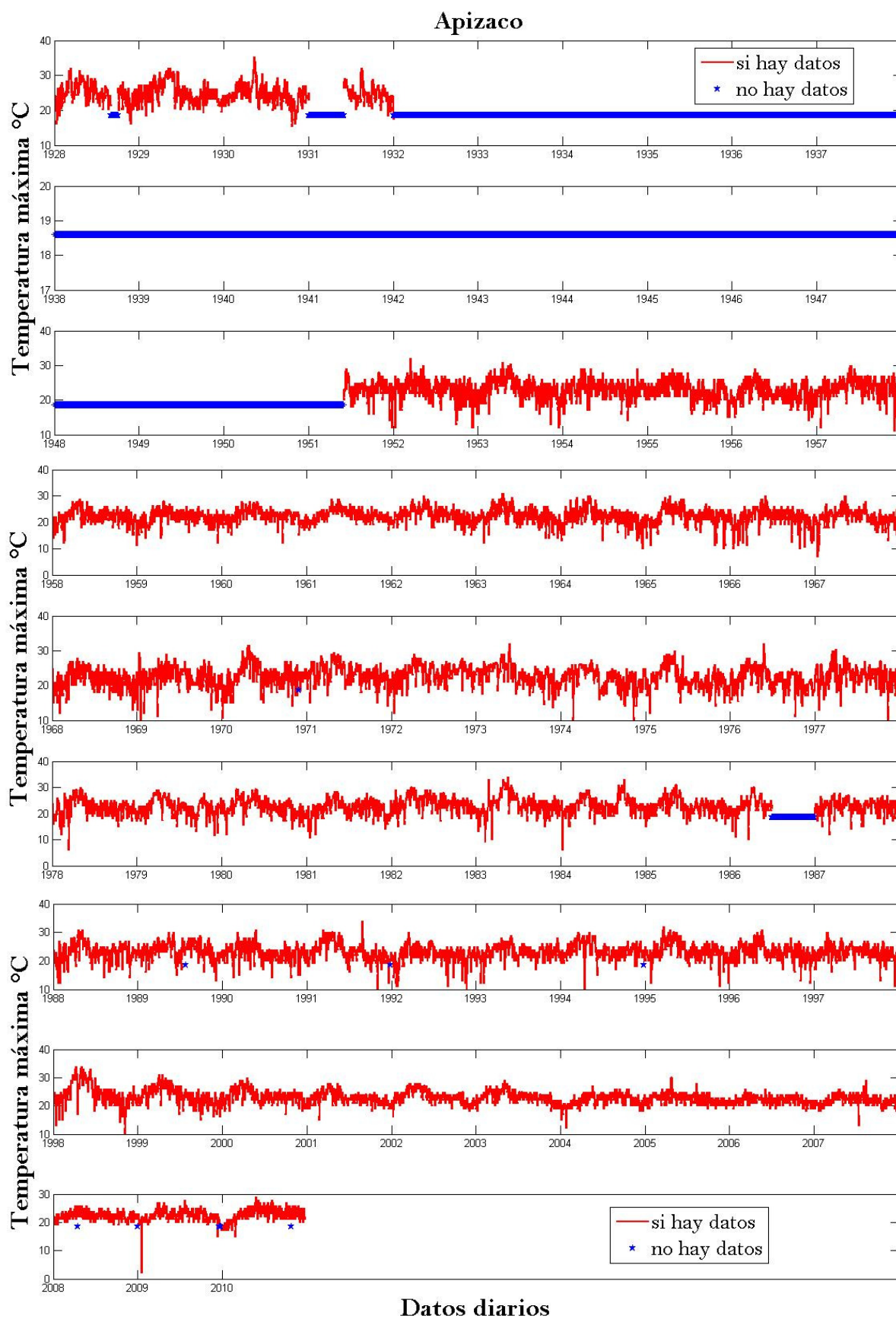
CONCLUSIONES Y TRABAJOS FUTUROS.

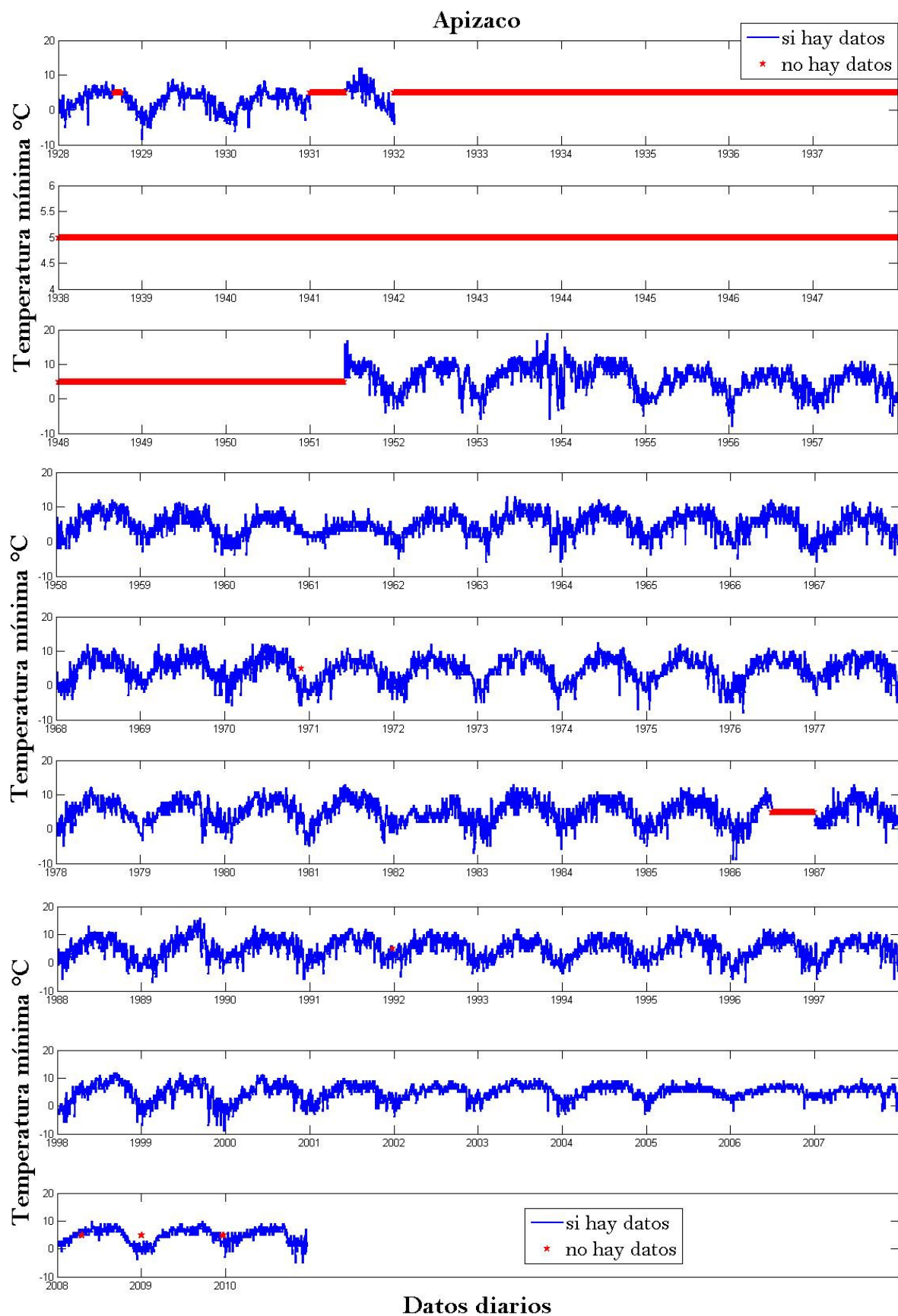
un mayor interés, en conjunción con algunos temas, metodologías y/o herramientas como: Sistemas Dinámicos, Teoría del Caos, Geometría Fractal, Wavelets, Análisis de Componentes Principales (PCA), Lógica Difusa (fuzzylogic), Algoritmos Genéticos, Inteligencia Artificial, etc. Lo cual hace falta explotar aun más todo el material en conjunción, aplicándolo a más de un área de estudio.

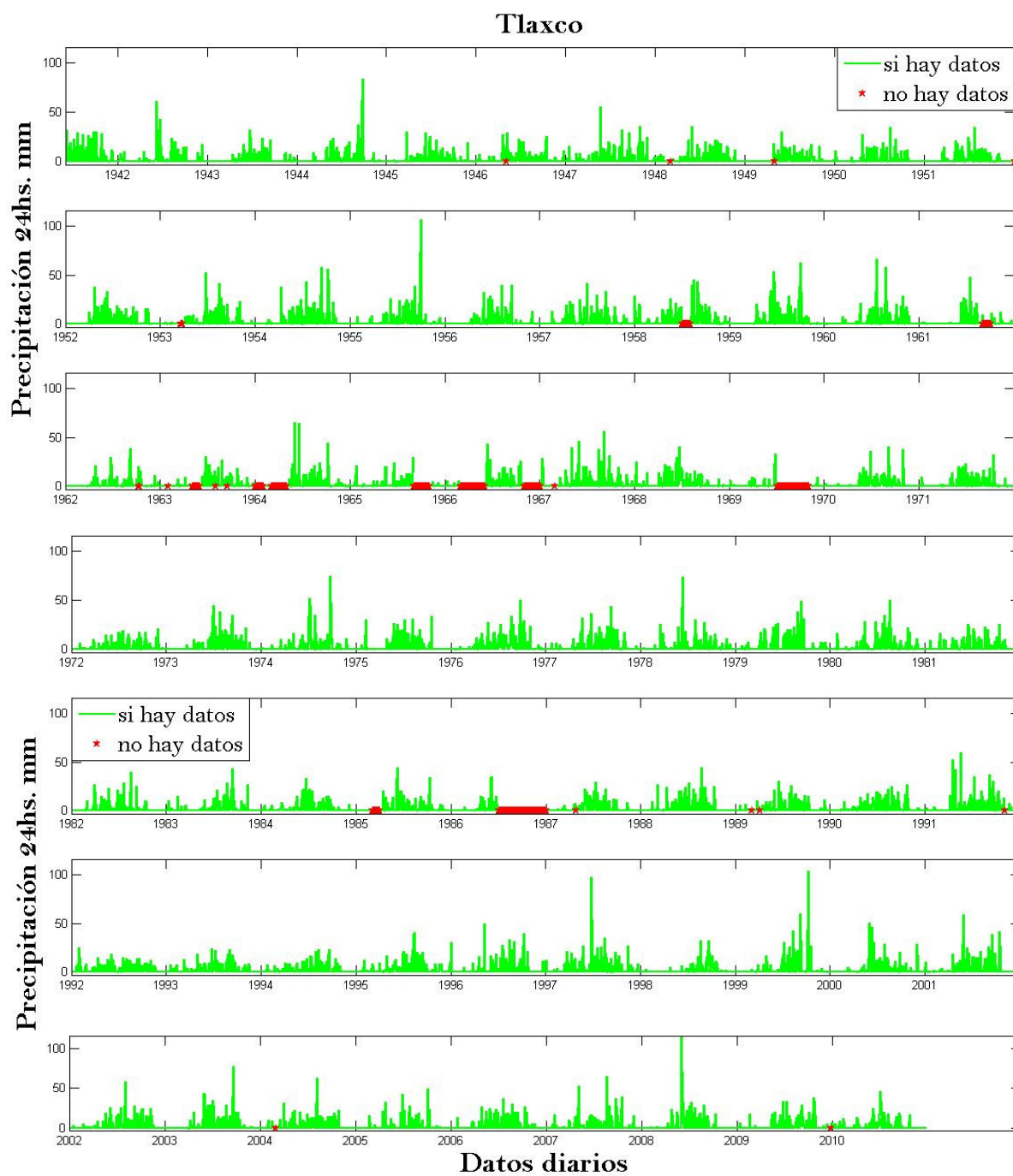
La calidad de los datos generados, es algo que se requiere y se tiene que tomar en cuenta, simplemente para tener una homogeneidad, por lo que haría falta esta implementación, recordando que se podrían aplicar las mencionadas en el capítulo uno, como la prueba de Sved-Eisenhart, prueba de igualdad de varianza, o bien hacer uso del control de calidad con el RClimDex.

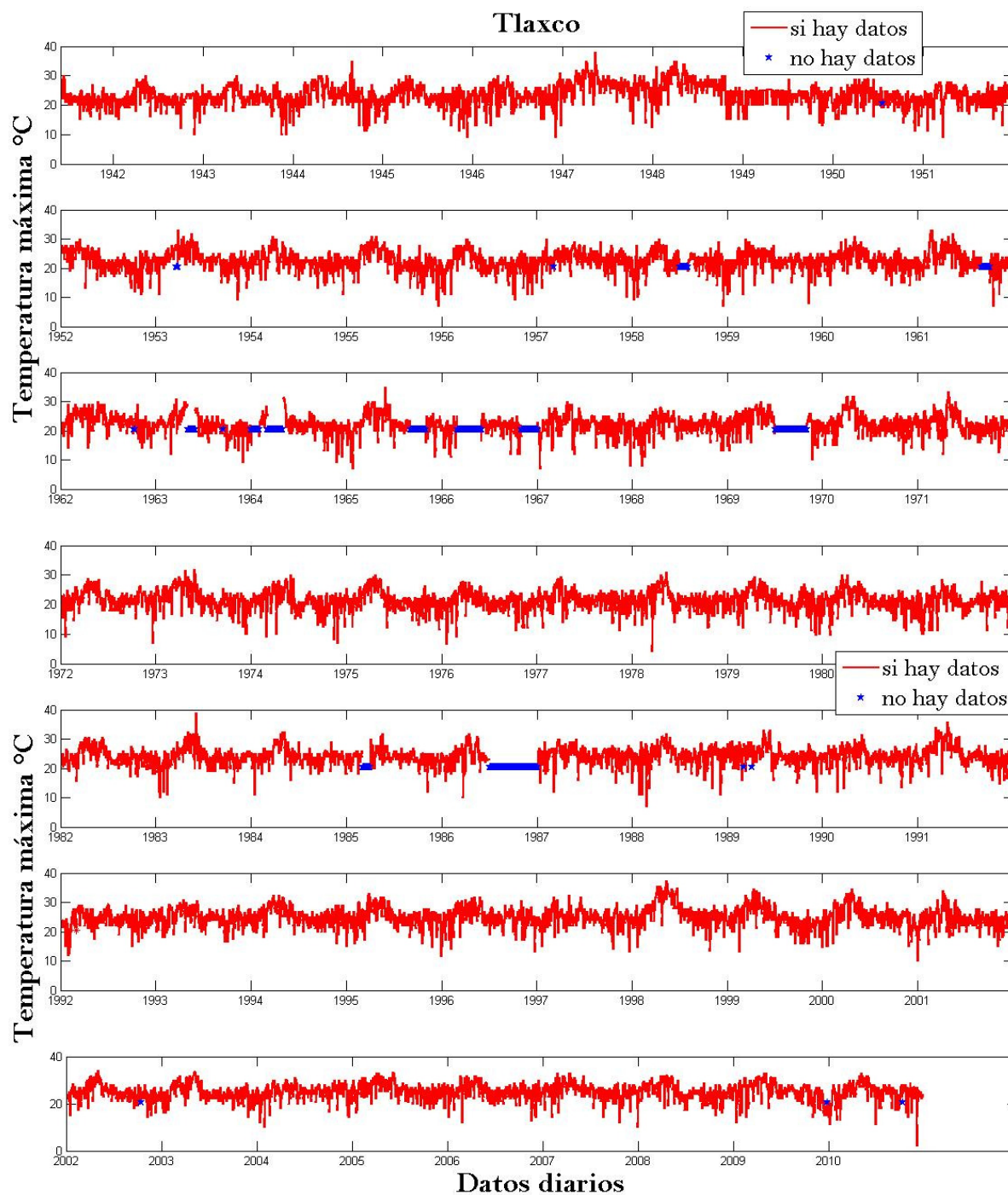
Al estar familiarizado con la base de datos, podemos ayudar a la RNA en su entrenamiento, y es posible con lo mencionó en el capítulo dos llamado *monitoring*, localizadas en la parte superior de cada gráfica anual, describiendo las temperaturas medias superficiales del océano en el Pacífico Ecuatorial, lo que sería si fue año NIÑO o NIÑA; entonces si sabemos la fase que tuvo el año que haya ausencia de datos, podríamos haber entrenado varias entradas de la misma fase en cuestión y así obtener una salida sin perder de vista lo que ocurre físicamente dentro de los eventos climáticos.

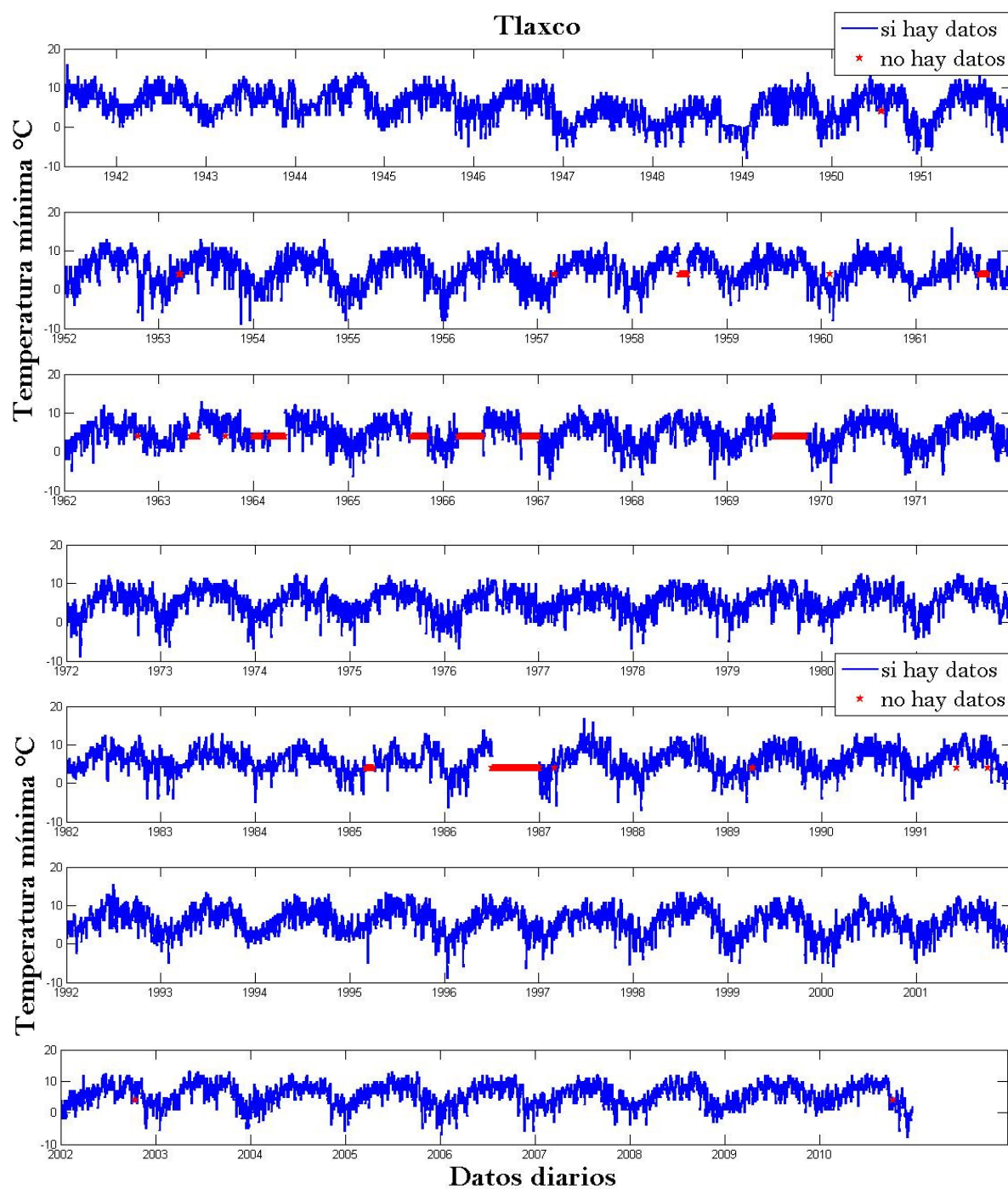
En este trabajo de tesis, la variable de precipitación, fue algo que no se trato y fue debido a que no se encontraron datos teóricos en la base mundial del NARR, pero se podría manejar de la siguiente forma: en principio se trabajaría ya no a una escala diaria, sino pentadal o mensualmente, esto para evitar ceros e intermitencia, ya que la precipitación puede actuar en que días consecutivos no hubo y en otro se registra una cantidad máxima, lo que es complicado tanto de pronosticarla como entrenarla en la RNA, a una escala diaria; la magnitud de la variable sería la acumulada de la escala que se esté manejando, teniendo límites en ella, por ejemplo, podríamos tener cantidades máximas de pcp acumulada para que sobrepasen la capacidad límite de una presa (inundación), o bien la escases (sequía), por lo que podríamos manejarlo como polos y así ayudar a la RNA para su entrenamiento, por lo que se le aplicaría un suavizado a la ST de entrada y así ubicar los polos que nos describan tales comportamientos; teniendo de esta manera su procesamiento para esta variable.

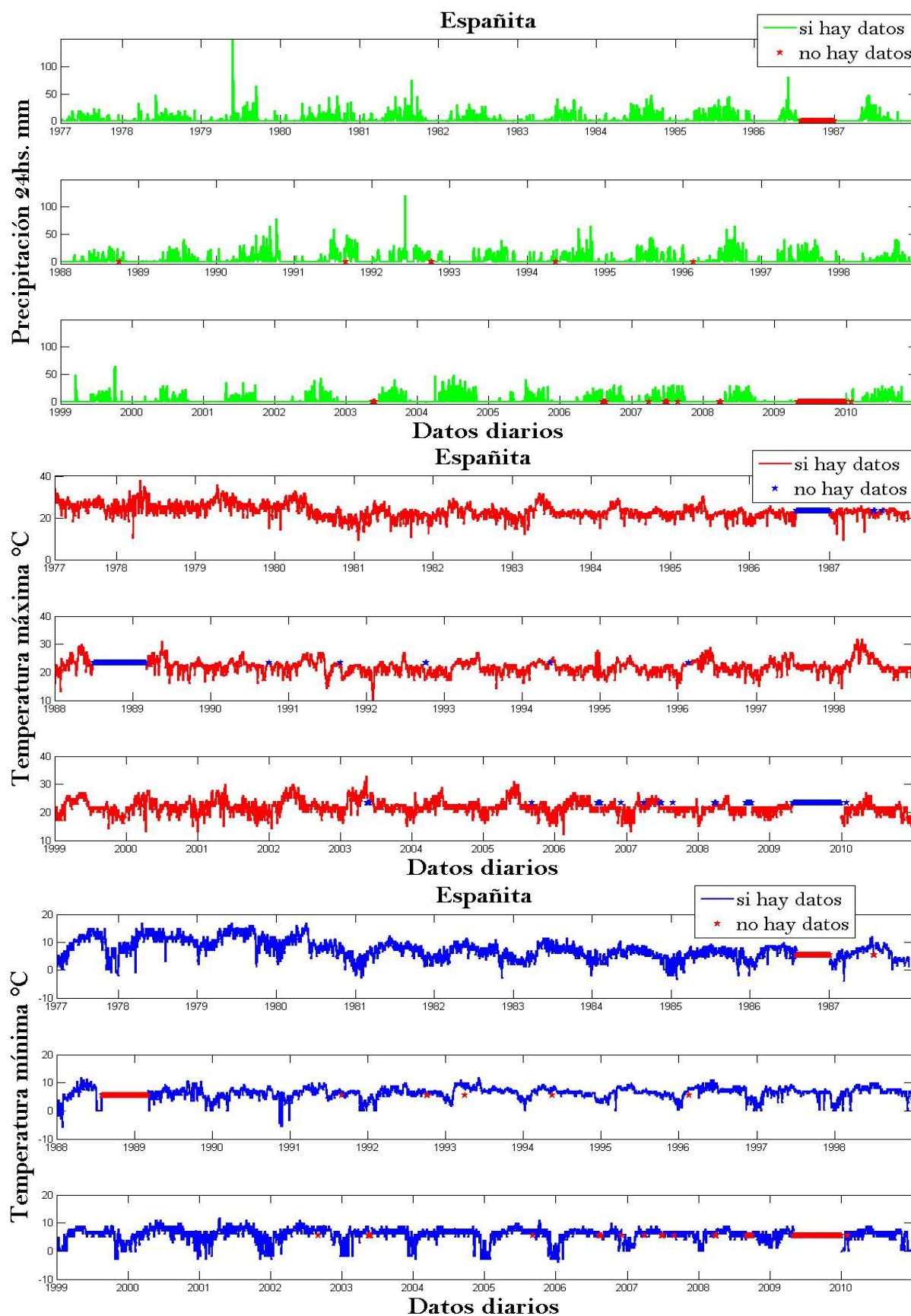












ANEXO B1

ALTURA GEOPOTENCIAL.

Se define el *geopotencial* Φ en un punto de la atmósfera como el trabajo que hay que realizar contra el campo gravitatorio terrestre para elevar una masa de 1 kg desde el nivel del mar hasta esa altura (J/kg).

$$d\Phi = g \cdot dz$$

$$\Phi = \int_0^z g \cdot dz$$

$$\Phi = 0 \text{ en } z = 0$$

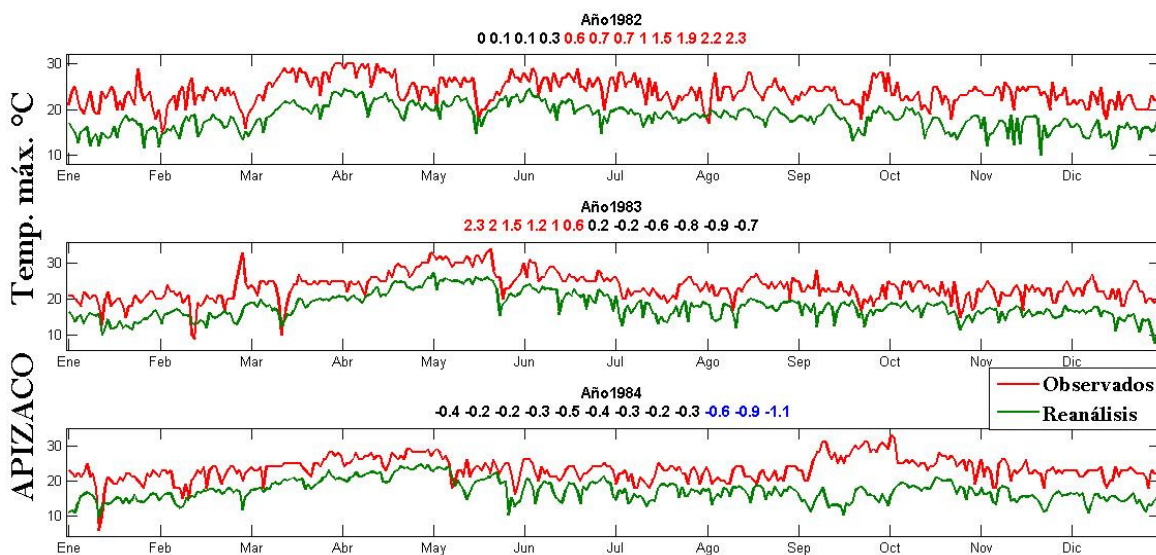
Se define *altura geopotencial* Z :

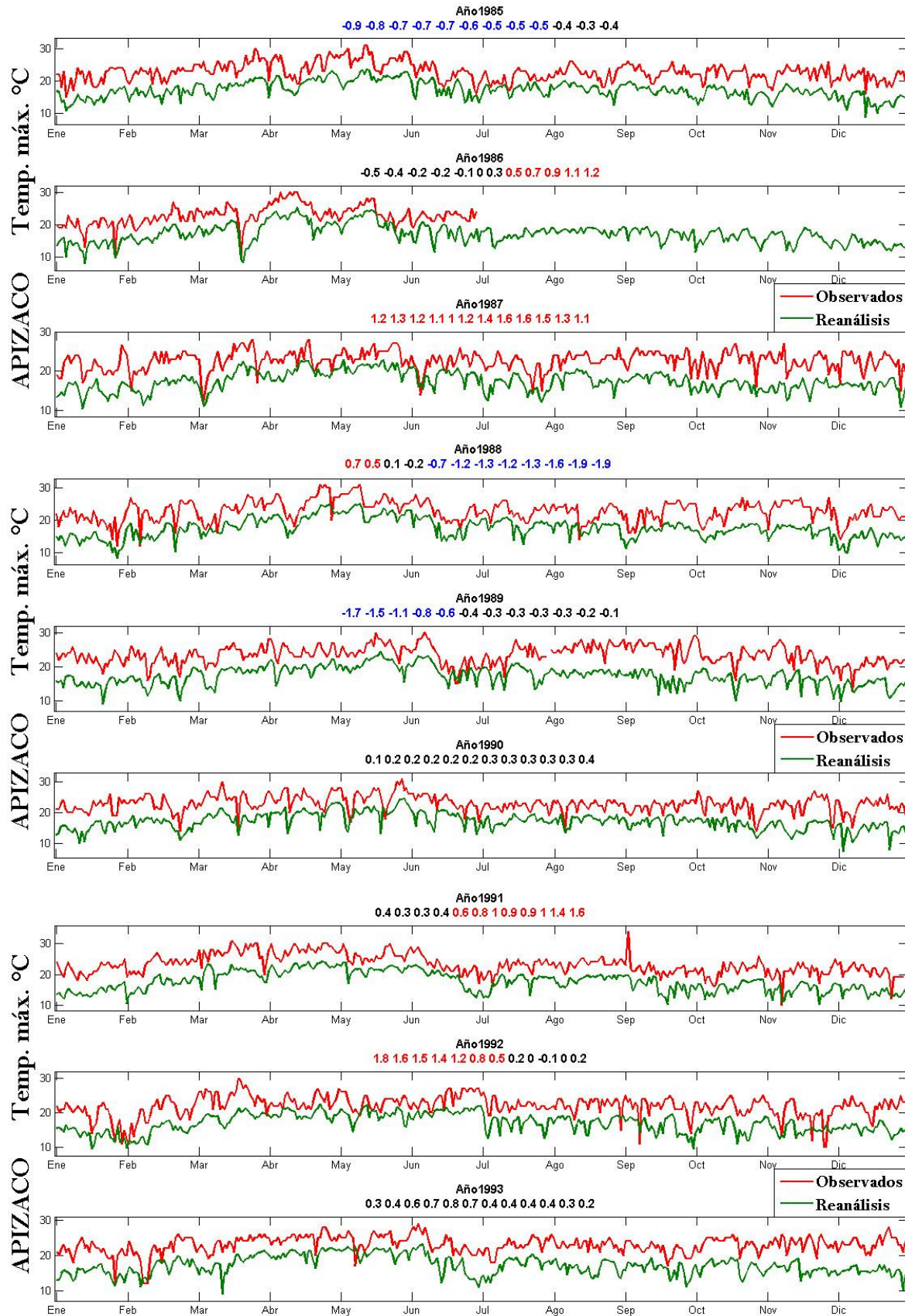
$$Z = \frac{\Phi}{g_0} = \frac{1}{g_0} \int_0^z g \cdot dz$$

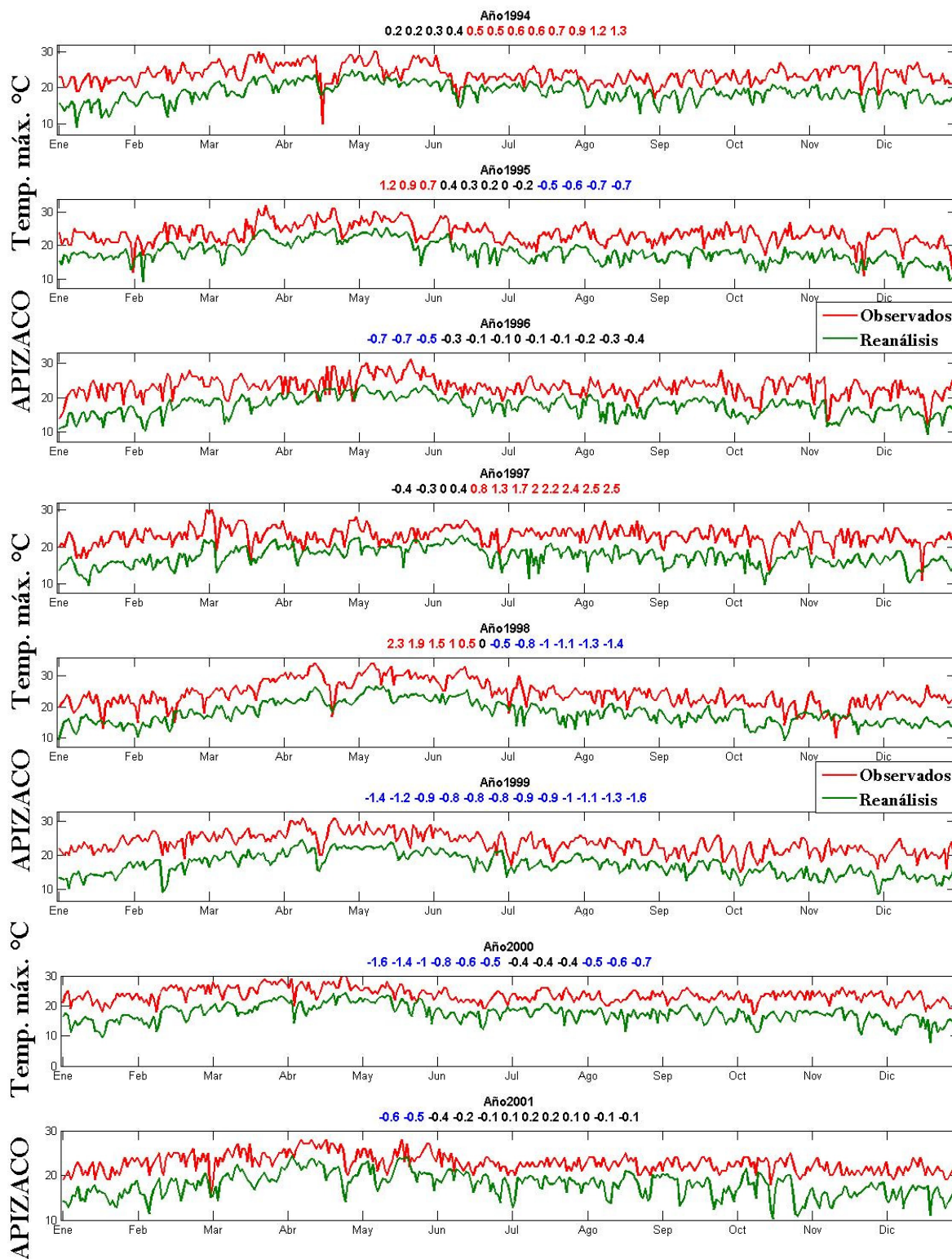
con $g_0 = 9.8 \text{ m/s}^2$ el valor de la gravedad en la superficie terrestre.

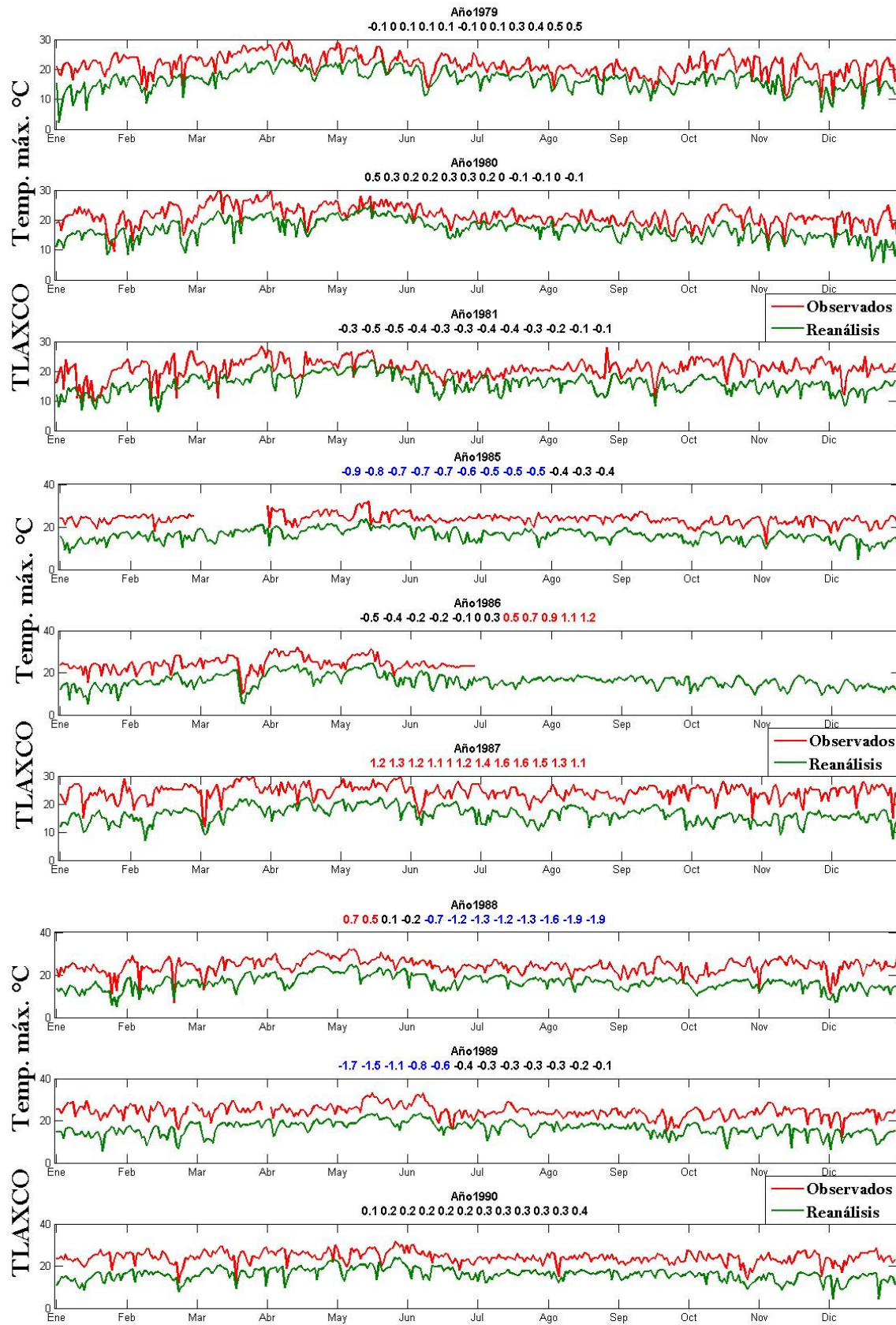
ANEXO B2

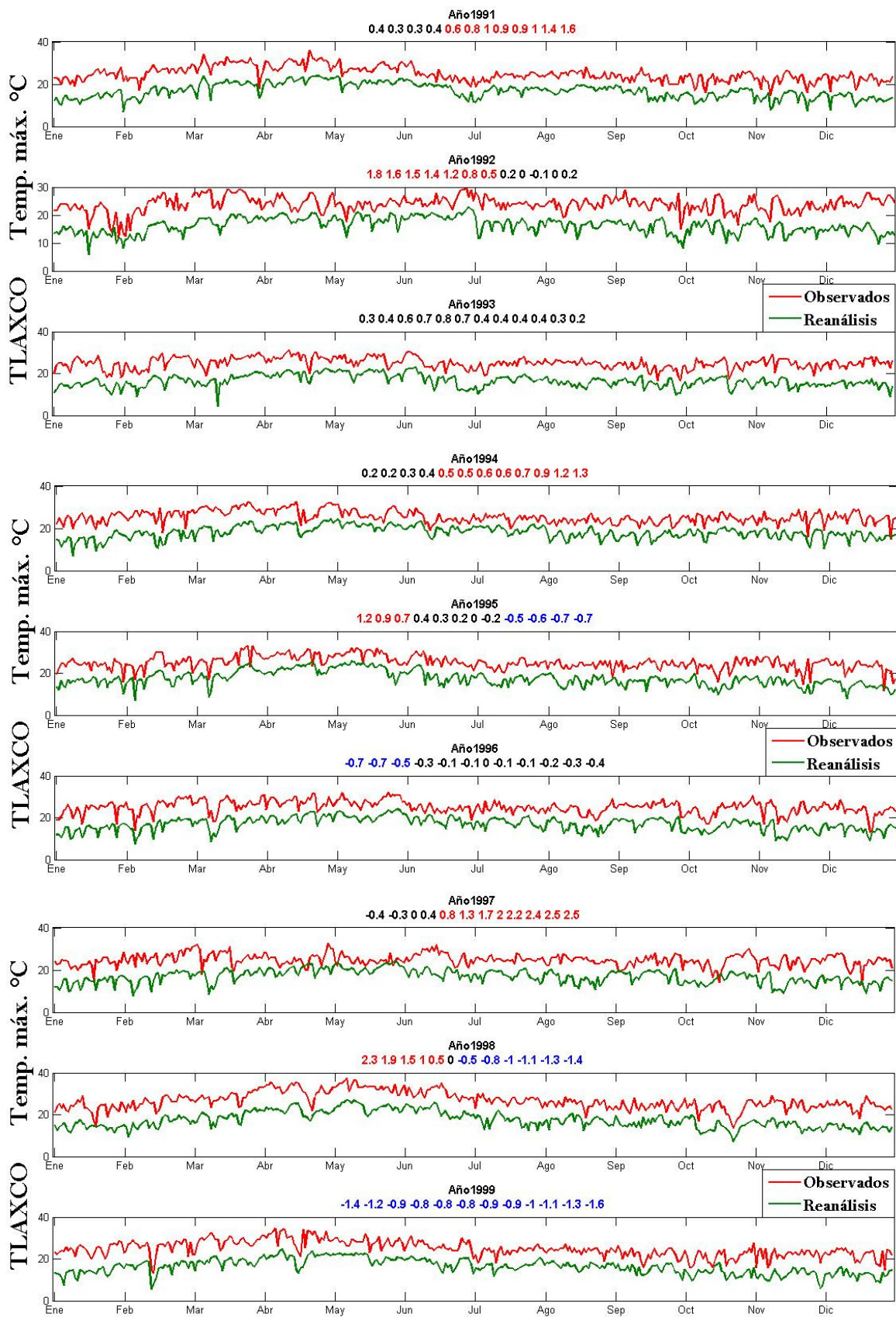
ST OBSERVADOS Y TEÓRICOS DE TEMPERATURAS MÁXIMA Y MÍNIMA A ESCALA DIARIA DE LAS ESTACIONES METEOROLÓGICAS EN ESTUDIO.

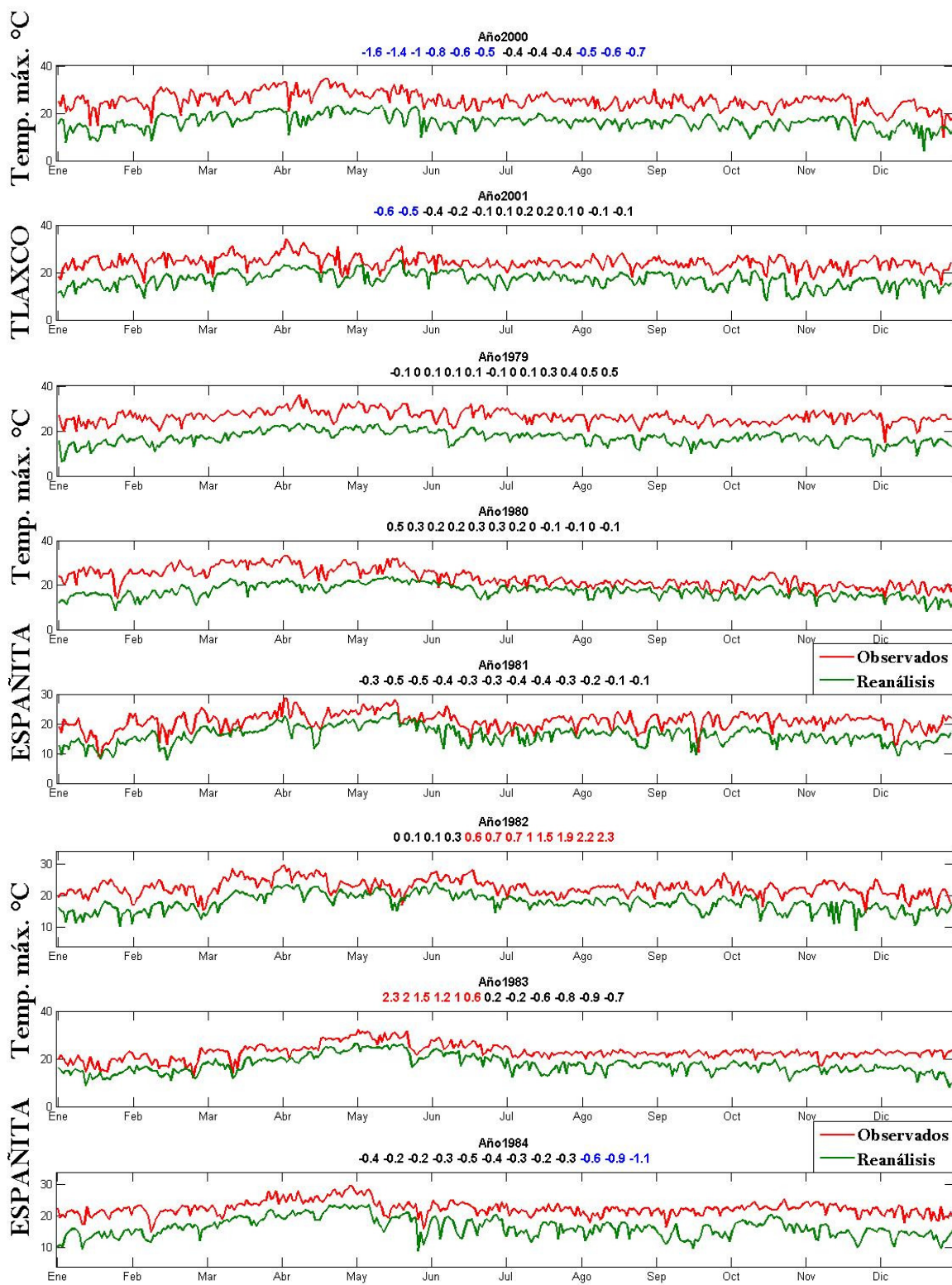


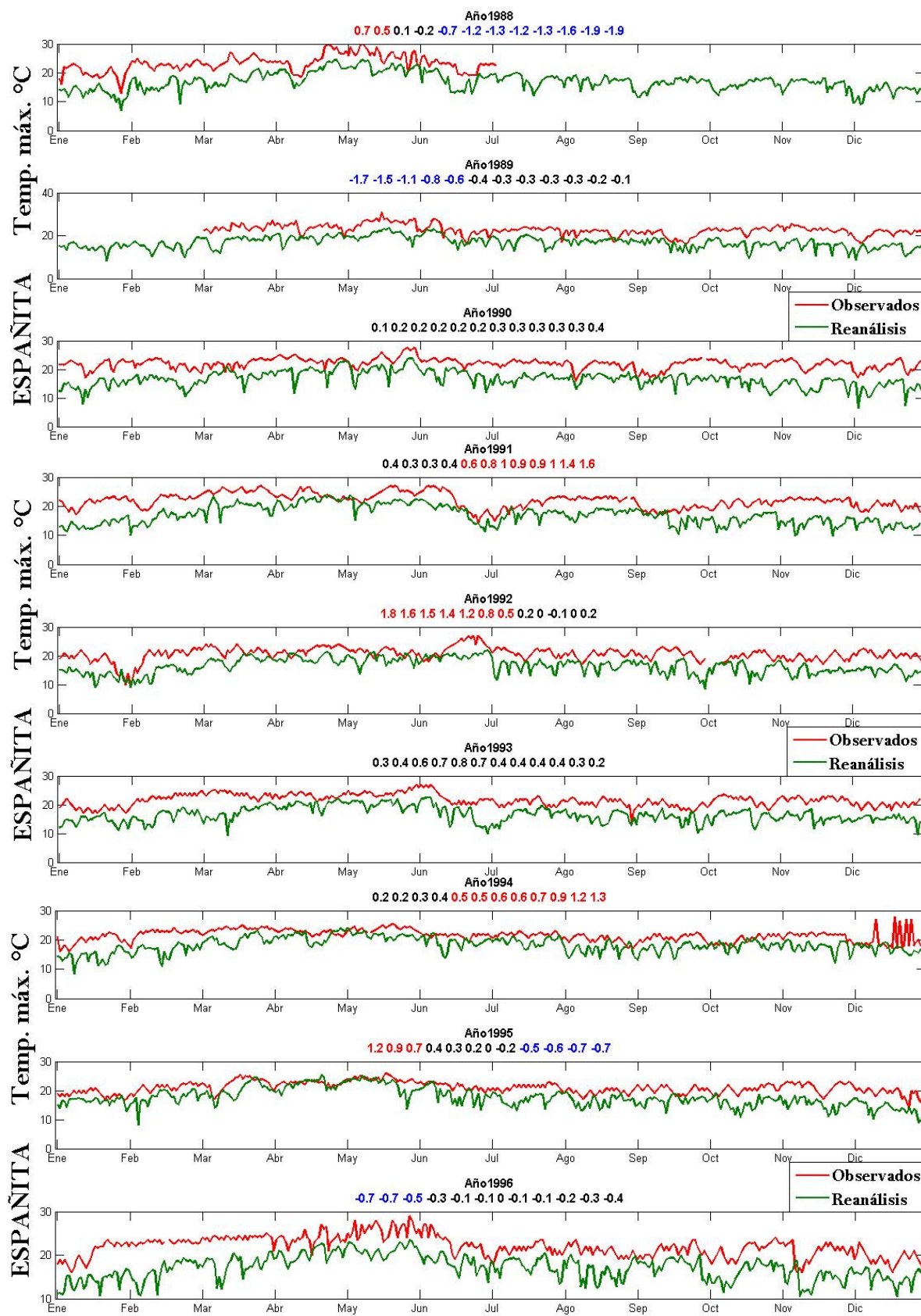


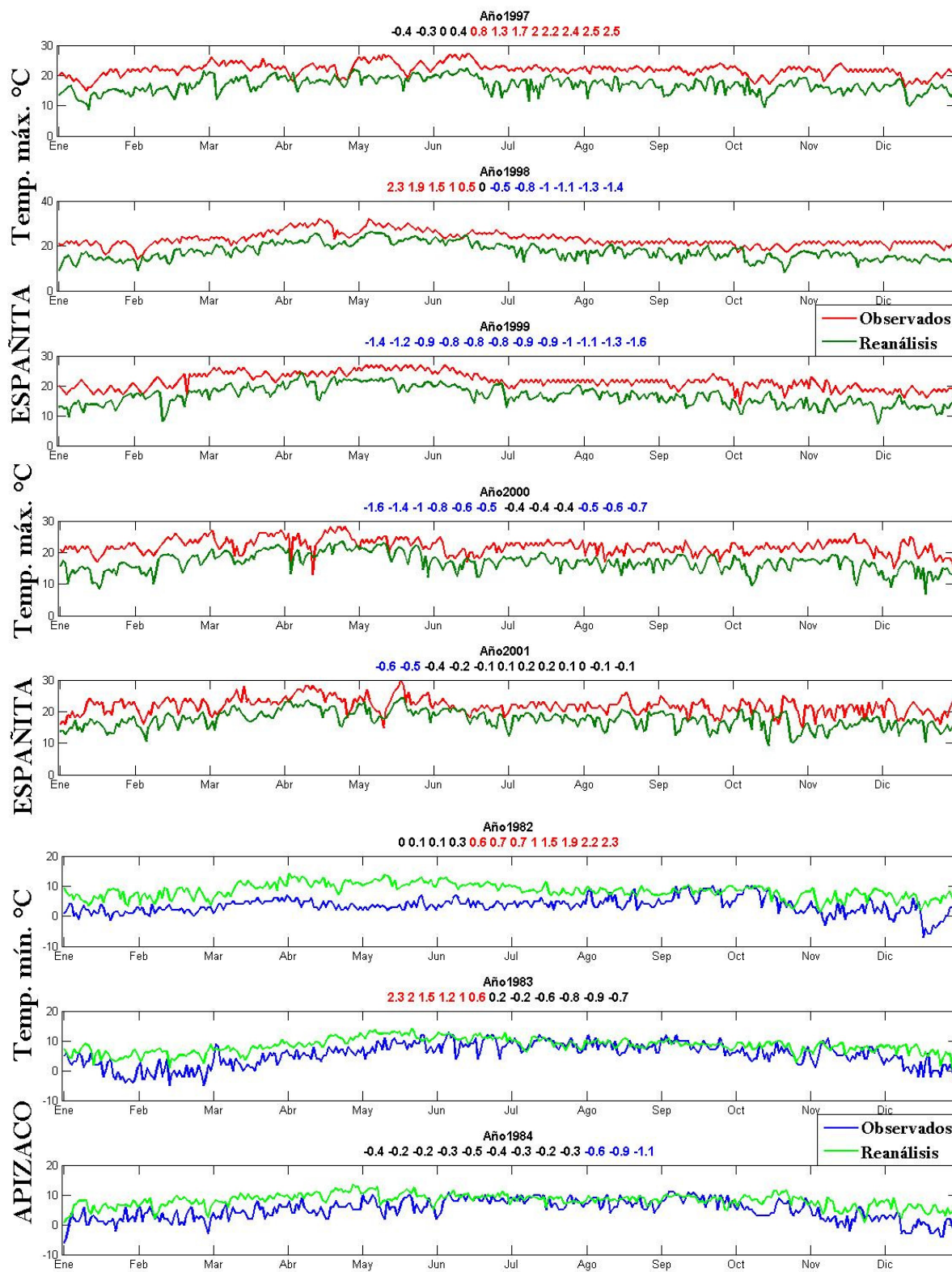


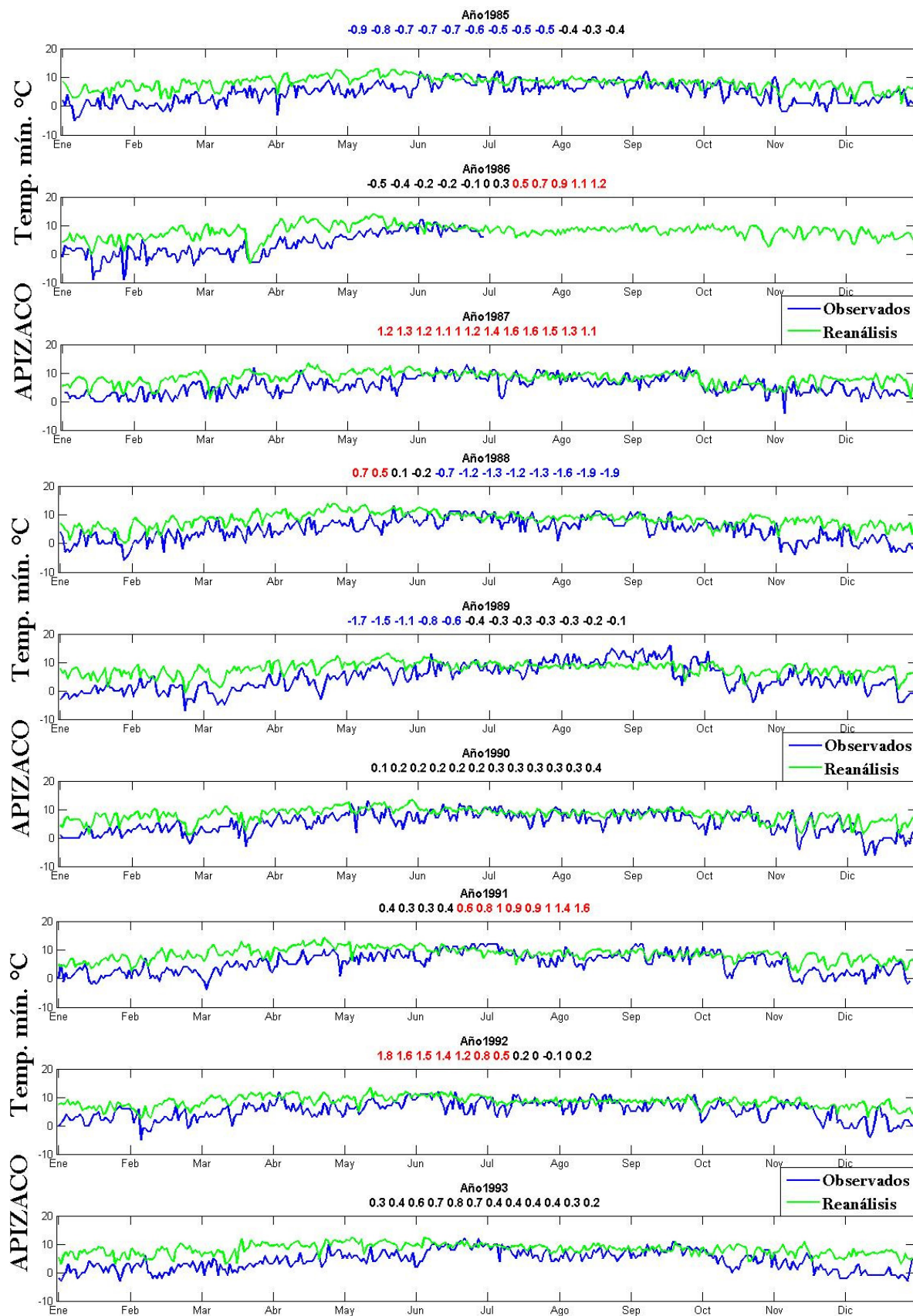


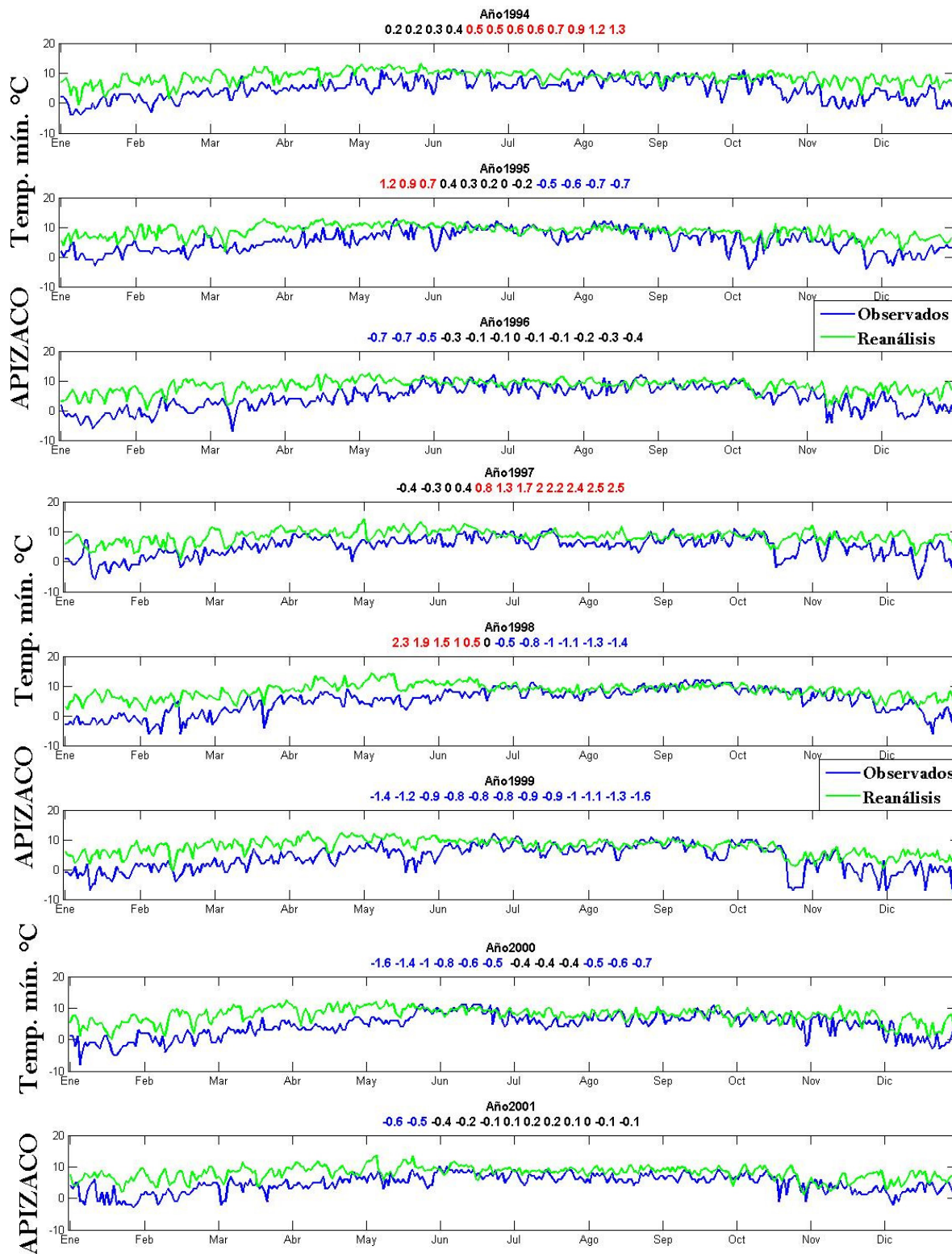


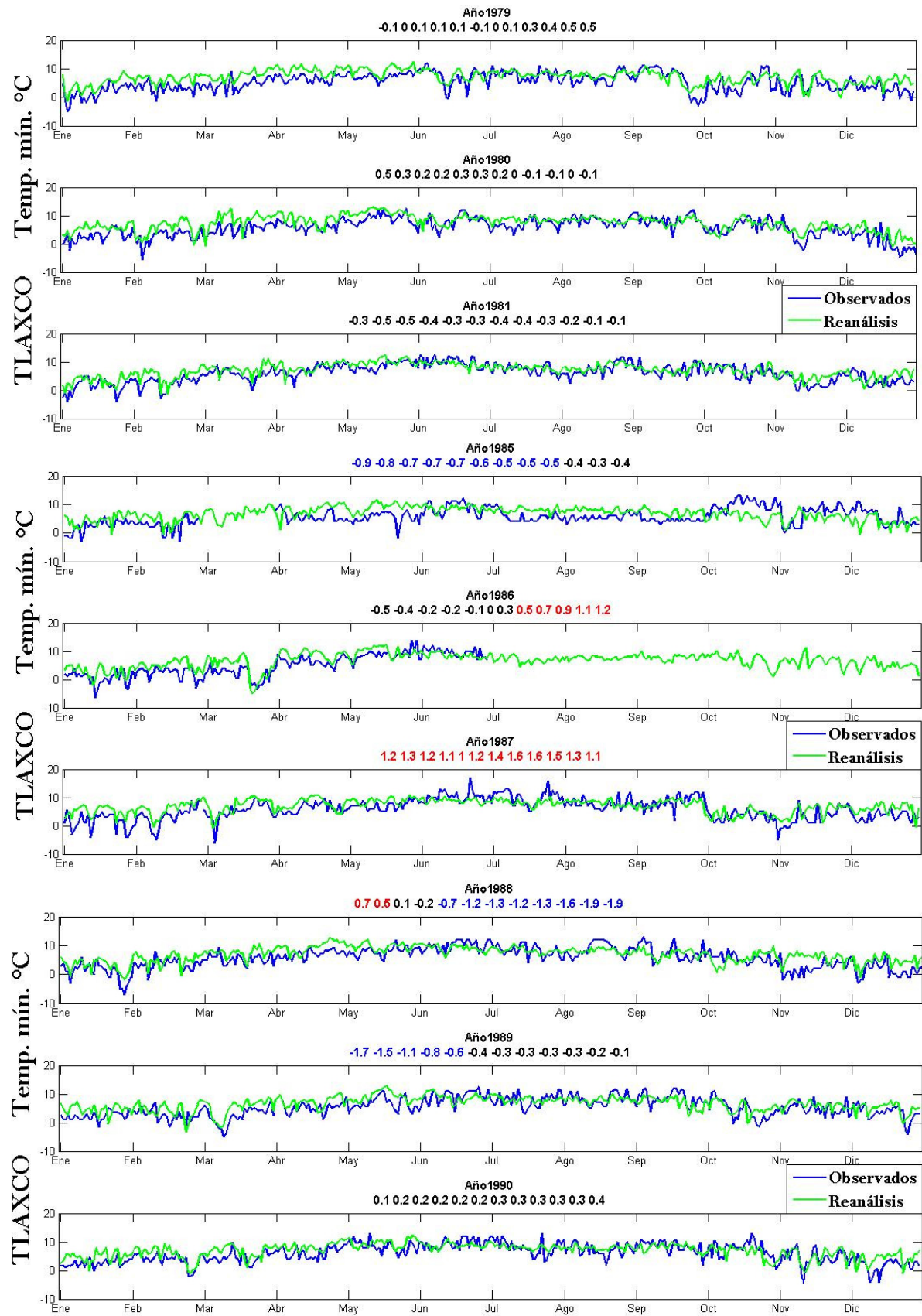


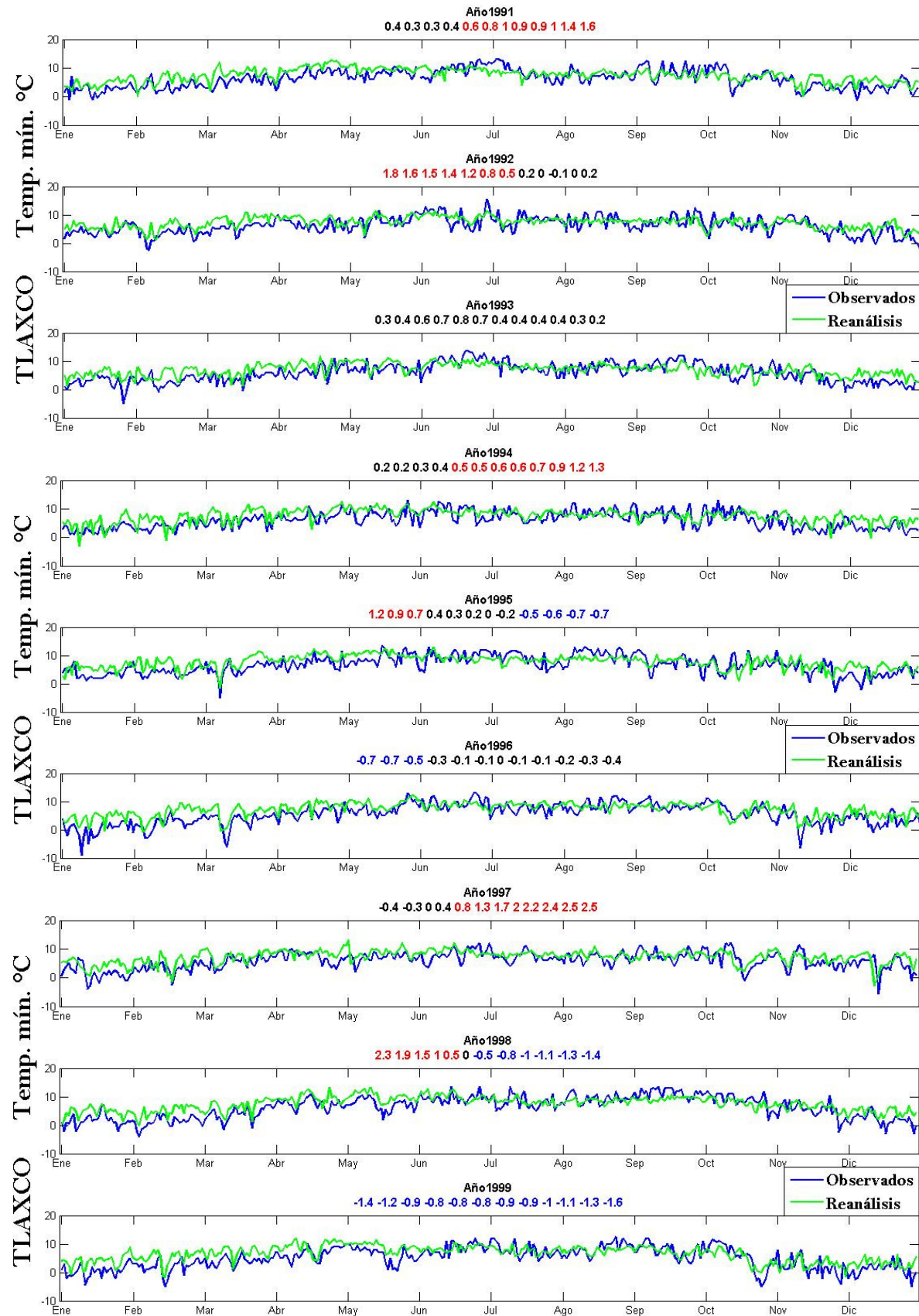


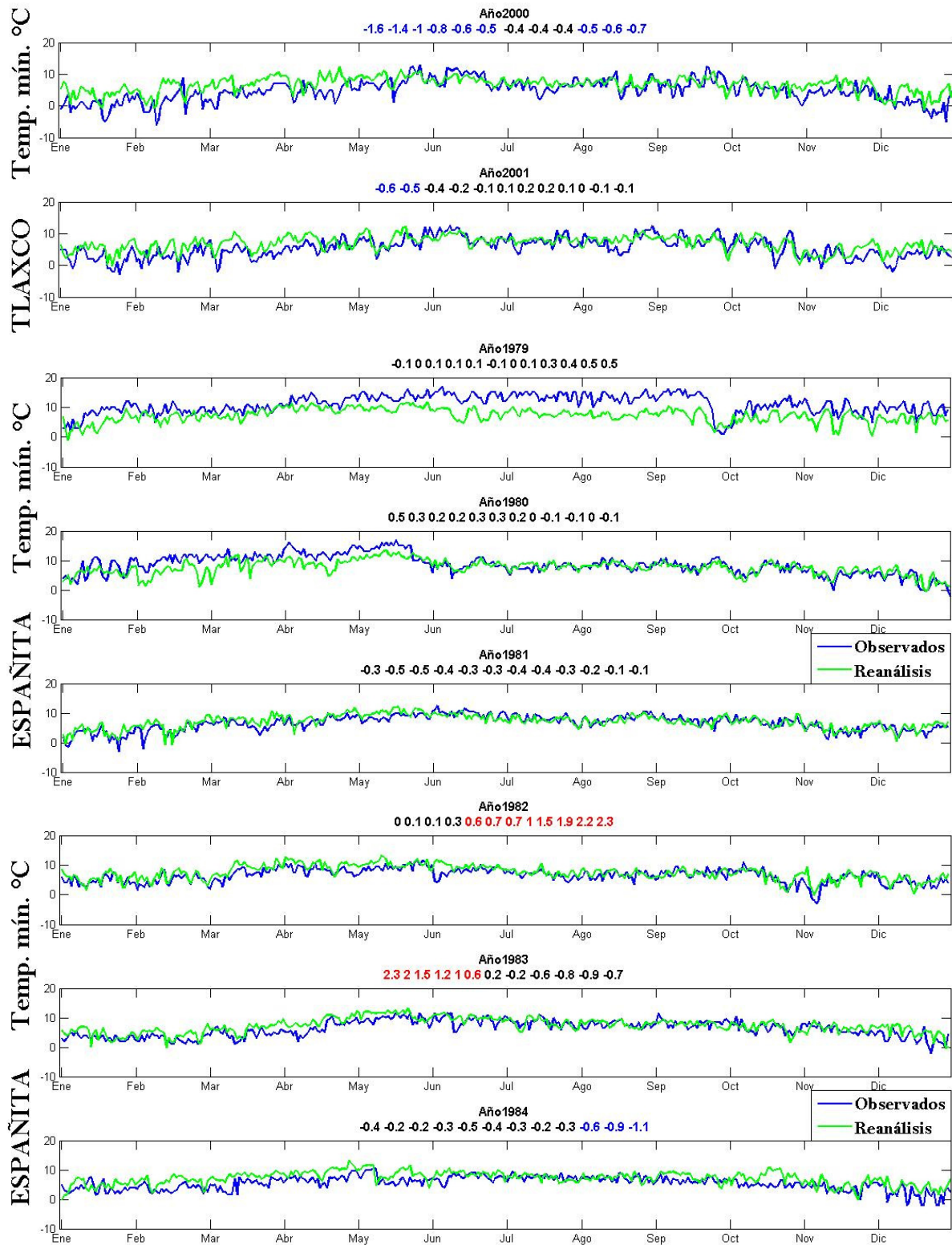


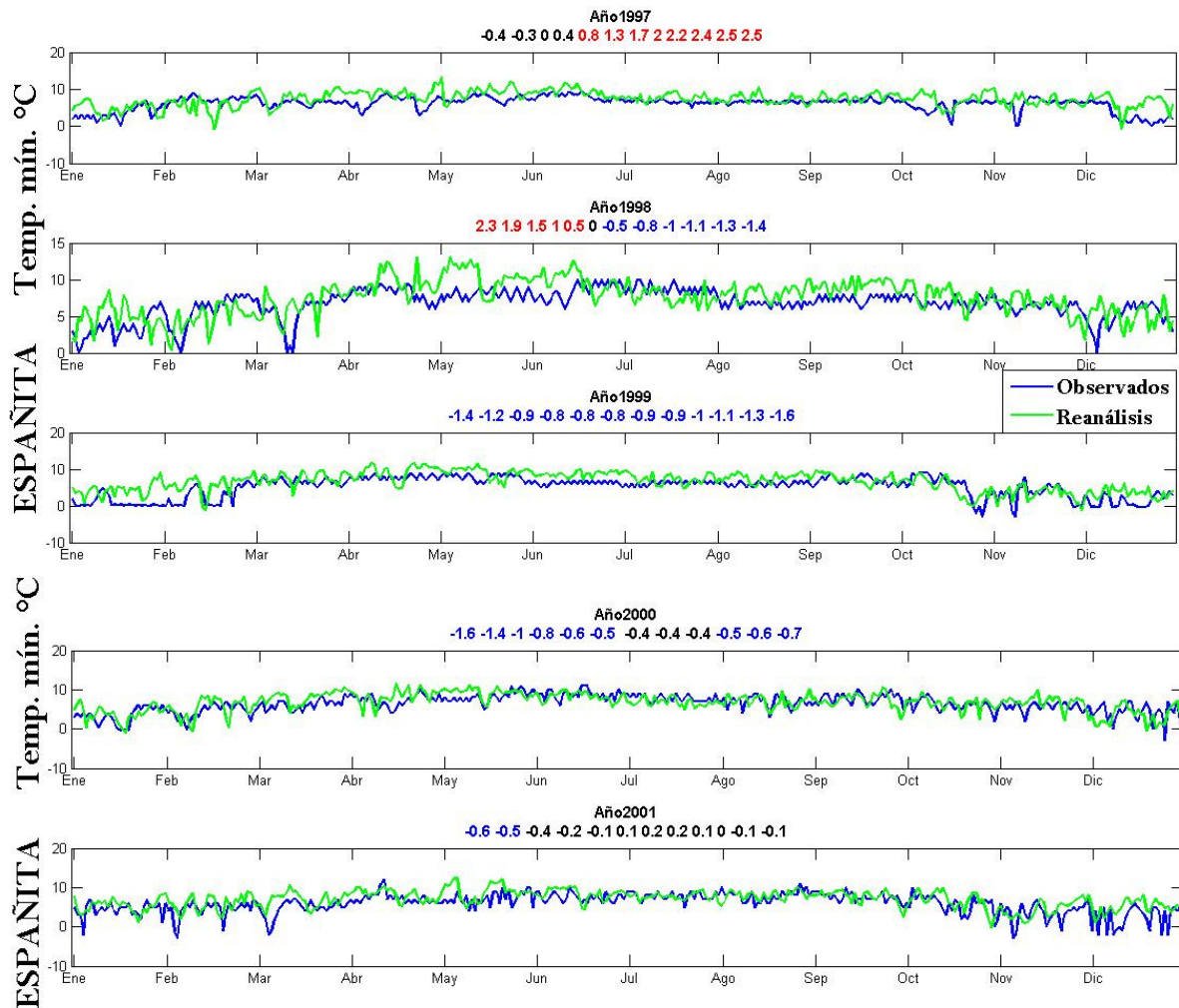












ANEXO B3

FUNDAMENTOS DE ESTADÍSTICA.

Algunas definiciones de conceptos estadísticos que son esenciales son:

Población.- colección de medidas u observaciones de todos los elementos de un universo acerca del que deseamos obtener conclusiones o tomar decisiones; la cual puede ser infinita o finita, en cuyo caso su tamaño se denota por n .

Muestra.- cualquier subconjunto de observaciones seleccionadas de la población.

Las variables aleatorias tales como $X_1, X_2, \dots, X_n$, forman una muestra aleatoria de tamaño n , si son independientes y tienen la misma distribución de probabilidad.

Un estadístico es cualquier función de las observaciones de una muestra aleatoria y que no depende de parámetros desconocidos, la media y la varianza calculadas a partir de los datos de una muestra aleatoria, son estadísticos.

Ejemplo de estadístico: $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

Si $\mathbf{X}$ es una variable aleatoria con función de densidad o de probabilidad $f_x(x; \theta)$, donde θ es un parámetro desconocido y si $X_1, X_2, \dots, X_n$ es una muestra aleatoria de tamaño n , entonces el estadístico $\hat{\theta} = h(X_1, X_2, \dots, X_n)$ recibe el nombre de estimador de θ . En términos sencillos, puede decirse que un estimador es un estadístico que tiene como propósito definido el “aproximar” un parámetro.

Una *estimación puntual* es cuando el estimador sólo proporciona un valor del parámetro que se desea estimar. El estimador de un parámetro cualquiera, por ejemplo θ , se denota utilizando un acento circunflejo; por lo que el estimador de θ , se simboliza: $\hat{\theta}$; y una estimación: $\hat{\theta}$.

Sea $\hat{\theta}$ un estimador puntual del parámetro θ . Entonces si $E(\hat{\theta}) = \theta$ se dice que $\hat{\theta}$ es un estimador **insesgado** de θ , de lo contrario es sesgado.

Sean $\hat{\theta}_1$ y $\hat{\theta}_2$ dos estimadores insesgados del parámetro θ , con variancias $Var(\hat{\theta}_1)$ y $Var(\hat{\theta}_2)$, respectivamente, entonces la eficiencia relativa de $\hat{\theta}_1$ con respecto a $\hat{\theta}_2$, ϑ , se define como: $\vartheta = \frac{Var(\hat{\theta}_2)}{Var(\hat{\theta}_1)}$

Cuando se desea comparar dos estimadores, de los cuales al menos uno no es insesgado, entonces la eficiencia relativa no se calcula como el cociente de las variancias, si no como el cociente de los **errores cuadráticos medios**.

El **error cuadrático medio** del estimador $\hat{\theta}$, del parámetro θ , se define como:

$$ECM(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$$

ANEXO C1

SISTEMAS Y MÁQUINAS INTELIGENTES (HISTORIA).

Norbert Wiener acuñó el término **cibernética** para designar el estudio unificado del control y de la comunicación en los animales y las máquinas, en 1948. De este modo, combinando las teorías sobre la realimentación, los avances de la electrónica de la posguerra y los conocimientos disponibles sobre los sistemas nerviosos de los seres vivos, se construyeron máquinas capaces de responder y de aprender como los animales.

El matemático de origen húngaro John von Neumann, quien en los años cuarenta concibe una computadora basada en lógica digital que opera ejecutando en serie (una tras otra) las instrucciones que componen un algoritmo que se codifica en forma de programa, el cual se encuentra almacenando en memoria. Debido a su eficacia, flexibilidad y versatilidad, y soportado por el impresionante desarrollo de las técnicas electrónicas, de modo que el binomio lógico booleana-máquina de Von Neumann es la base sobre la que se asientan la mayor parte de nuestros actuales computadores digitales.

Hacia 1960 John McCarthy acuña el término **Inteligencia Artificial (IA)**, para definir los métodos algorítmicos capaces de hacer pensar a los ordenadores.

El rápido progreso de la **IA** culminó en los años sesenta con la introducción de los **Sistemas Expertos (SE)**, complejos programas de computador en los que se codifica el conocimiento de expertos en una materia muy concreta (concesión de créditos, diagnóstico de enfermedades, etc.), en forma de reglas de decisión. No obstante, un cuarto de siglo más tarde nuestros ordenadores son miles de veces más potentes que los de la época de los pioneros de la **IA** y, sin embargo, en general, no resultan mucho más inteligentes. El problema radica en que al binomio lógica booleana-máquina Von Neumann sobre el que se asienta la **IA**, pese a su gran potencia, presenta problemas a la hora de abordar ciertas tareas, como aquellas denominadas del mundo real, donde la información que se presenta es masiva, imprecisa y distorsionada. Para abordar este tipo de tareas, desde la década de los ochenta (principalmente) se han vuelto a retomar, o bien han resurgido, una serie de paradigmas de cómputo alternativos, como **redes neuronales**, **los sistemas borrosos**,

algoritmos genéticos o la computación evolutiva y en consecuencia sistemas híbridos como **neuro-difuso**.

Así, el resurgimiento de las redes neuronales, también denominadas sistemas neuronales artificiales o ANS (Artificial Neuronal Systems)²⁶, se debe, por una parte, a la dificultad para resolver con la eficiencia deseada problemas como los de visión o aprendizaje y debido a los altos requerimientos computacionales de este tipo de tareas, la arquitectura de Von Neumann no es muy apropiada y la introducción de sistemas de cálculo paralelo resulta necesaria si se quiere lograr respuestas en tiempo real [Churchland P.M., 1990]

Inteligencia Artificial IA	Redes neuronales
Enfoque descendente	Enfoque ascendente
Basado en la psicología	Basado en la biología
Qué hace el cerebro	Cómo lo hace el cerebro
Reglas <i>Si/Entonces</i>	Generalización a partir de ejemplos
Sistemas programados	Sistemas entrenados
Lógica, conceptos, reglas	Reconocimiento de patrones, gestalt ²⁷
Arquitectura Von Neumann. Separación hardware/software	Arquitectura paralelas, distribuidas, adaptivas. Autoorganización

Características de la **IA** convencional y de las RNA [McCord N., 1991]

ANEXO C2

CEREBRO VS COMPUTADOR.

Los actuales computadores son máquinas de Von Neumann, en esencia una máquina de procesamiento (hardware) que actúa ejecutando en serie (una tras otra) una secuencia de instrucciones o programa (software), que almacena en su memoria.

²⁶ Existen otras denominaciones, como redes neuronales artificiales (Artificial Neuronal Networks), neurocomputación (Neuronal Computing) o sistemas conexionistas.

²⁷ Gestalt es un sustantivo de la lengua alemana por lo que siempre debe escribirse en mayúscula. Aunque ha sido traducido como forma o configuración, suele utilizarse sin traducción ya que no cuenta con un equivalente exacto en el idioma español. La psicología de la Gestalt es una corriente de la psicología moderna, que surgió a principios del siglo XX en Alemania. Esta escuela sostiene que la mente configura, a través de distintas leyes, los elementos que llegan a ella mediante la percepción o la memoria.

La máquina está compuesta por cuatro unidades básicas: unidad de entrada de información, unidad de salida, unidad de procesamiento (compuesta a su vez por la unidad lógico-aritmética y la unidad de control), y memoria. Este sistema constituye una máquina universal de cómputo en el sentido de Turing²⁸, por lo que puede llevar a cabo, en principio, cualquier tarea sólo con programarla de forma adecuada.

Para un computador, el verdadero corazón de éste, es el **microprocesador**, potente y complejísimo circuito electrónico, que en la actualidad puede integrar varios millones de componentes electrónicos en un espacio de alrededor de un centímetro cuadrado [Martín del Brío, 1999]. La mayor parte de los microprocesadores no se encuentran dentro de los ordenadores, como se podría pensar, si no contenidos en todo equipo que contenga cierta cantidad de electrónica, como puedan ser hornos de microondas, aparatos de televisión, videos, cámaras de video, lavadoras, etc²⁹.

La neurona es en realidad un pequeño procesador, sencillo, lento y poco fiable (a diferencia de nuestros potentes microprocesadores), sin embargo, en nuestro cerebro cohabitan unos cien mil millones de neuronas operando en paralelo. Es aquí donde reside el origen de su poder de cómputo. Aunque individualmente las neuronas sean capaces de realizar procesamiento muy simples, ampliamente interconectadas a través de las sinapsis (cada neurona puede conectarse con otras diez mil en promedio) y trabajando en paralelo pueden desarrollar una actividad global de procesamiento enorme.

²⁸ Alan Turing emprendió el estudio formal de la computación, para lo que introdujo una máquina ideal conocida como máquina de Turing [Conrad M., 1992; Hopcroft J.E., 1984].

²⁹ Estos microprocesadores que se dedican a tareas de control se denominan microcontroladores, integrando en una única pastilla todos los elementos del computador: CPU (Central Processing Unit), memoria, entradas y salidas, etc. [Martín del Brío, 1999].

GLOSARIO.

Canícula.- Corresponde a una disminución en las lluvias durante Julio y Agosto, provocando el aumento de la temperatura. Este periodo canicular, no se presenta en todas las regiones de México [Magaña Rueda et.al., 1999].

Cognitivo.- Es aquello perteneciente o relativo al conocimiento. Éste, a su vez, es el conjunto de información almacenada mediante la experiencia o el aprendizaje.

Determinista.- Del determinismo. El mundo físico (sistema) evoluciona en el tiempo según principios o reglas totalmente predeterminadas.

Empírico.- Del empirismo o relativo a él. Partidario del empirismo filosófico. Deriva del griego antiguo de experiencia, ἐμπειρία, que a su vez deriva de ἐν en y πείρα prueba, experimento.

Empirismo.- Conocimiento que se origina desde la experiencia. Sistema filosófico basado fundamentalmente en los datos de la experiencia [Real Academia Española, 22.^a edición].

Índices Climáticos.- Caracterización numérica de las variables del clima y de sus relaciones. (Índice de lluvia útil, índice de sequía de Palmer, índice de lluvias extremas).

Método Empírico.- Modelo de investigación científica, que se basa en la experimentación y la experiencia, así como la observación de sucesos y su análisis estadístico.

Micra.- Unidad de longitud equivalente a la millonésima parte del metro; la micra se representa con la letra griega μ .

Monótono, na.- Uniforme, que no cambia.

Neuronas (biológicas).- Componente básico del sistema nervioso, incluyendo al cerebro. Son células que presentan una forma especial; tienen tres partes principales: las dendritas, el cuerpo de la célula o soma y el axón. Las dendritas son el elemento receptor; son especies de fibras que cargan de señales eléctricas al cuerpo de la célula; el soma realiza la suma de las señales y el axón es una fibra larga que lleva la señal desde el cuerpo de la célula hacia otras.

El punto de contacto entre el axón de una célula y la dendrita de otra se denomina sinapsis [Stevens, 1979].

Percentil.- Mat. Valor que divide un conjunto ordenado de datos estadísticos de forma que un porcentaje de tales datos sea inferior a dicho valor. Así, un individuo en el percentil 80 está por encima del 80% del grupo a que pertenece [Real Academia Española, 22.^a edición]. Cada uno de los noventa y nueve segmentos que resultan de dividir algo en cien partes de igual frecuencia.

Retrospectivo.- Se refiere a un tiempo pasado.

Robusto.- Procede del latín *robustus*, hace referencia a aquello vigoroso, fuerte o firme. Por ejemplo, para un sistema informático, puede presumir de su robustez si puede ejecutar diversos procesos de manera simultánea sin generar fallos o bloquearse (colgarse). Un ejemplo de sistema robusto es aquel destinado a la creación de gráficos en 3D que posibilita trabajar con imágenes muy pesadas sin problemas. Si el programa “se cuelga” al cargar la primera imagen, se trata de una aplicación poco robusta.

Sistema.- Consiste de muchas partes individuales que interactúan cada una con las otras.

Sistema Experto (SE).- Es una aplicación informática capaz de solucionar un conjunto de problemas que exigen un gran conocimiento sobre un determinado tema. Un sistema experto es un conjunto de programas que, sobre una base de conocimientos, posee información de uno o más expertos en un área específica. Se puede entender como una rama de la inteligencia artificial, donde el poder de resolución de un problema en un programa de computadora viene del conocimiento de un dominio específico. Estos sistemas imitan las actividades de un humano para resolver problemas de distinta índole (no necesariamente tiene que ser de inteligencia artificial). También se dice que un **SE**, se basa en el conocimiento declarativo (hechos sobre objetos, situaciones) y el conocimiento de control (información sobre el seguimiento de una acción).

BIBLIOGRAFÍA

[Baidyk Tatiana, 2009]

Baidyk, Tatiana, **Redes neuronales, Visión Computacional y Micromecánica**, UNAM-Itaca, México, D.F., 2009.

[Casillas Lavín, 2004]

Casillas Lavín, Tomás, **Análisis de Series Temporales Caóticas para la Detección de Dinámica no Lineal: Sistemas Tiempo/Clima**, Tesis de Licenciatura, Facultad de Ciencias, UNAM, México, D.F., 2004.

[Carmona Rodríguez, 2009]

Carmona Rodríguez, María, A., **Minería de Datos como una Herramienta en las Ciencias Ambientales: Análisis de la Precipitación y Temperaturas Extremas en Tlaxco y Atlagatepec**, Tesis de Licenciatura, Facultad de Agrobiología, Campus Tlaxco, UATx, 2009.

[Chamizo Checa, 2008]

Chamizo-Checa, Silvia, **Caracterización Dinámica de Precipitación Extrema en el Estado de Tlaxcala: Caso de Estudio 04 de Octubre de 1999**, Tesis de Maestría, UATx, 2008

[Chatfield C., 1995]

Chatfield, C., **The Analysis of Time Series: an Introduction**, Chapman & Hall/crc, 1995.

[Churchland P.M., 1990]

Churchland, P. M., Smith Churchland, P., **¿Podría Pensar una Máquina?**, Investigación y Ciencia, pp. 18-24, marzo 1990.

[Conrad M., 1992]

Conrad, M., **Molecular Computing Paradigms**, IEEE Computer, pp. 6-9, diciembre 1992.

[Cooray T., 2008]

Cooray, T., M., J., A., **Applied Time Series: Analysis and Forecasting**, Alpha Science International Ltd, 2008.

[Corchado J.M., 2000]

Corchado, J.M., Díaz, F., Borrajo, L., Fernández, F., **Redes Neuronales Artificiales, un Enfoque Práctico**, servicio de publicaciones de Universidad de Vigo, España, 2000.

[Hassoon M.H., 1995]

Hassoon, M. H., **Fundamentals of Artificial Neural Network**, The MIT Press, London, England, 1995.

[Haykin S., 1994]

Haykin, S., **Neural Networks, a Comprehensive Foundation**, Macmillan College Publishing Company, New York, USA, 1994.

[Haykin S., 1999]

Haykin, S., **Neural Networks, a Comprehensive Foundation**, 2a edición, Prentice Hall, New Jersey, USA, 1999.

[Hernández Arriaga, 1996]

Hernández Arriaga, R., **Determinación de la Dimensión del Atractor y el Límite de Predictibilidad de Una Serie de Tiempo de Horas de Calor**, Tesis de licenciatura, Facultad de Ciencias, UNAM, 1996.

- [Hilera González, 1995]
Hilera González, J. R., Martínez, V. J., **Redes Neuronales Artificiales, Fundamentos, Modelos y Aplicaciones**, Addison-Wesley, Madrid, España, 1995.
- [Hopcroft J.E., 1984]
Hopcroft, J. E., **Máquinas de Turing**. Investigación y Ciencia, 1984.
- [Infante Saba et al., 2008]
Infante, Saba, Ortega J., Cedeño, Fernando, **Estimación de Datos Faltantes en Estaciones Meteorológicas de Venezuela vía un Modelo de Redes Neuronales**, Facyt, Universidad de Carabobo, Vol. 8 ISSN 1578-8768, pp. 51-70, Venezuela, 2008.
- [Kantz H., 1997]
Kantz H., Schreiber T., **Nonlinear Time Series Analysis**, University Press, Cambridge, 1997.
- [Magaña Rueda et al., 1998]
Magaña Rueda, V., Morales, T., Pérez, J. L., Orozco, S., **El Niño y La Niña en el estado de Tlaxcala**, Fundación Produce, A. C., SEMARNAP & CONACYT, Tríptico, 1998.
- [Magaña Rueda et al., 1999]
Magaña Rueda, V., J. A., Amador, S. Medina, **The mid-summer drought over Mexico and Central America**, J. Climate, 12, 1577-1588, 1999.
- [Magaña Rueda, 2004]
Magaña Rueda, Víctor, **Los Impactos del Niño en México - Capítulo2: El Niño y el Clima**, Centro de Ciencias de la Atmósfera, UNAM, Secretaría de Gobernación, México, pp. 23-68, 2004.
- [Mandelbrot B., 1982]
Mandelbrot, B., **The Fractal Geometry of Nature**, W. H. Freeman, New York, 1982.
- [Martin del Brio, 1999]
Martin del Brio, Bonifacio, **Sistemas Electrónicos basados en Microprocesadores y Microcontroladores**, Editorial Prensas Universitarias, Zaragoza, 1999.
- [Martin del Brio, 2002]
Martin del Brio, Bonifacio, Sanz Molina, Alfredo, **Redes Neuronales y Sistemas Difusos**, 2a edición ampliada y actualizada, Alfaomega, México, c2002.
- [Martin del Brio, 2007]
Martin del Brio, Bonifacio, Sanz Molina, Alfredo, **Redes Neuronales y Sistemas Difusos**, 3a edición revisada y ampliada, Alfaomega, México, D.F., 2007.
- [Martínez Robles, 2006]
Martínez Robles, Manuel, **Clasificación del Grado de Astrocitoma Cerebral Infantil: Segmentación de Imágenes, Morfología Matemática y Redes neuronales**, Tesis de Maestría, Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, 2006.

[Mendivil G., 2011]

Mendivil G., Fabián, Urquizo Rojas, Daniela, **Aplicación de Redes Neuronales Artificiales para el Análisis de la Inflación en Bolivia**, 4to. Encuentro de Economistas de Bolivia, Sucre, 2011.

[Meneses Popocatl, 1993]

Meneses Popocatl, A., **Aplicación de la Teoría del Caos para Identificar el Número de Variables con Capacidad de Predicción del Tiempo para el Centro-Norte de Tlaxcala**, Tesis de Licenciatura, Depto. de Agrobiología, UATx, 1993.

[McCord N., 1991]

McCord, N., Illingworth W. T., **A Practical Guide to Neural Nets**, Addison-Wesley Publishing, 1991.

[Mirzaee Hossein, 2009]

Mirzaee, Hossein, **Linear Combination Rule in Genetic Algorithm for Optimization of Finite Impulse Response Neural Network to Predict Natural Chaotic Time Series**, Khorasan research institute for Food Science and Tecnology, pp. 2681-2689, 2009.

[Montesdeoca P., 2005]

Montesdeoca P., Pablo, **Cómo Resolver Problemas: Longitud y Área de Curvas Fractales, Dimensión Fractal**, Universidad de Palmas de Gran Canaria, 2005.

[Morales Acoltzi et al., 2011]

Morales Acoltzi, Tomás, Alva Pacheco, J. C., Peña Maciel, D., Bustamante García, A. R., Monroy Martínez, J. D., Bernal Morales, R., **Experimentos Numéricos de Predicción del Tiempo/Clima: un Enfoque No Lineal**, Memorias Congreso UGM 2011, Puerto Vallarta, Jal., México.

[Morales Acoltzi et al., 2012]

Morales Acoltzi, T., Bernal Morales, R., Suárez Sánchez, J., Hernández Rodríguez, M. de L., Magaña-Rueda, V. O., Herrera Cortés, S., Gómez, L., Ordoñez Díaz B. J., Barrera Bassols, N., Cuevas Sánchez A., **Variabilidad y cambio climático de la lluvia en tres sitios en el estado de Tlaxcala**, Cap. del libro Foro del agua para enfrentar el cambio climático, p. 23, 2012.

[Orr G.B., 1998]

Orr, G.B., Müller, K.R., **Neural Networks: Tricks of the Trade**, Springer-verlag, 1998.

[Prasad Girijesh, 2004]

Prasad, Girijesh, Leng, Gang, Martin McGinnity, Thomas, **An on-line Algorithm for Creating Self-Organizing Fuzzy Neural Networks**, University of Ulster at Magee, Derry, Northern Ireland, pp. 1-17, 2004.

[Peña Maciel, 2012]

Peña Maciel, Daniel, **Análisis de Series de Tiempo de Variables Atmosféricas Aplicando la Teoría de Sistemas Dinámicos No Lineales y un Enfoque de Minería de Datos: Estado de Tlaxcala**, Tesis de Licenciatura, Facultad de Ingeniería, UNAM, 2012.

- [Prechelt L., 1998]
Prechelt, L., **Early Stopping-But When?**, EN [Orr 98], pp. 55-70, 1998.
- [Reyes Oliver, 2006]
Reyes Oliver, Irma, **Desarrollo de un Sistema de Pronóstico de Series de Tiempo Intermitente: Lluvia Diaria**, Tesis de Maestría, Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, 2006.
- [Rojas Raul, 1996]
Rojas, Raul, **Neural Networks: a Systematic Introduction**, Springer-Verlag, Berlin, 1996.
- [Rumelhart D.E., 1986]
Rumelhart, D.E., Hinton G.E., Williams R.J., **Learning Representations by Backpropagating Errors**, Nature, 323, pp. 533-6, 1986.
- [Sánchez Camperos, 2006]
Sánchez Camperos, Edgar Nelson., **Redes Neuronales: Conceptos Fundamentales y Aplicaciones a Control Automático**, Pearson Educación, Madrid-México, 2006.
- [Sánchez Pérez, 2006]
Sánchez Pérez, Francisco Javier., **Análisis No Lineal de Series de Tiempo de Viento y Temperatura de Mesoescala: Laguna Verde, Veracruz**, Tesis de Maestría, Instituto Tecnológico de Apizaco, Tlaxcala, 2006.
- [Santos J.L., 2004]
Santos, J. L., **RClimDex 1.0, Manual del Usuario**, Centro Internacional para la Investigación del Fenómeno de El Niño CIIFEN, 2004.
- [Schreiber T., 1999]
Schreiber T., **Interdisciplinary Application of Nonlinear Time Series Methods**, Physics Reports 308-1, 1999.
- [Simpson P.K., 1990]
Simpson, P.K., **Artificial Neural Systems: Foundation, Paradigms, Applications and Implementations**, Pergamon Press, New York, USA, 1990.
- [Smith S., 1999]
Smith S., **The Scientist and Engineer's Guide to Digital Signal Processing**, 2a Edition, Technical Publishing, California, 1999.
- [Stevens C.F., 1979]
Stevens, C. F., **The Neuron**, Scientific American, 241(3), pp. 54-65, 1979.
- [Tlaxcala Castañeda, 2005]
Tlaxcala Castañeda, Alma Nelly, **Clasificación de No Linealidad de Modelos Determinísticos**, Tesis de Maestría, Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, 2005.

[Villavicencio Sánchez, 2009]

Villavicencio Sánchez, Rodrigo, **Reconstrucción de Forzamiento en Sistemas Dinámicos**, Tesis de Licenciatura, Facultad de Ciencias, UNAM, México, D.F., 2009.

[Wan Eric, 1993]

Wan, Eric A., **Finite Impulse Response Neural Networks with Applications in Time Series Prediction**, Stanford University, PhD Dissertation, 1993.

[Werbos P.J., 1974]

Werbos, P. J., **Beyond Regression: New Tools for Prediction and Analysis in Behavioral Sciences**, Doctoral Dissertation, Applied Mathematics, Harvard University, November, 1974.

[Zhang X., 2004]

Zhang, X., Yang, F., **RClimDex 1.0 User's manual**, Climate Research Branch Environment Canada Downsview, Ontario Canada, 2004.

[Zhang G., 2001]

Zhang G. Peter, Patuwo, B. Eddy, Y. Hu, Michael, **A simulation Study of Artificial Neural Networks for Nonlinear Time-Series Forecasting**, Computers & Operations Research, Georgia State University, Atlanta, GA 30303-3083, USA, pp. 381-396, 2001.