

Información Importante

La Universidad de La Sabana informa que el(los) autor(es) ha(n) autorizado a usuarios internos y externos de la institución a consultar el contenido de este documento a través del Catálogo en línea de la Biblioteca y el Repositorio Institucional en la página Web de la Biblioteca, así como en las redes de información del país y del exterior, con las cuales tenga convenio la Universidad de La Sabana.

Se permite la consulta a los usuarios interesados en el contenido de este documento, para todos los usos que tengan finalidad académica, nunca para usos comerciales, siempre y cuando mediante la correspondiente cita bibliográfica se le dé crédito al trabajo de grado y a su autor.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, La Universidad de La Sabana informa que los derechos sobre los documentos son propiedad de los autores y tienen sobre su obra, entre otros, los derechos morales a que hacen referencia los mencionados artículos.

BIBLIOTECA OCTAVIO ARIZMENDI POSADA
UNIVERSIDAD DE LA SABANA
Chía - Cundinamarca



El presente formulario debe ser diligenciado en su totalidad como constancia de entrega del documento para ingreso al Repositorio Digital (Dspace).

TITULO	Desarrollo de un modelo de red neuronal artificial para la reducción de escala (downscaling) de datos de temperatura del modelo Climático Global Canadiense 3.1 a estaciones meteorológicas colombianas.		
SUBTITULO			
AUTOR(ES) Apellidos, Nombres (Completo) del autor(es) del trabajo	Cardozo Vásquez, Andrés		
PALABRAS CLAVE (Mínimo 3 y máximo 6)	Redes neuronales artificiales		
	temperatura media diaria		
	Colombia		
RESUMEN DEL CONTENIDO (Mínimo 80 máximo 120 palabras)	<p>Se desarrolló un modelo basado en redes neuronales artificiales (RNA) para el pronóstico de la temperatura media diaria a escala local en 5 zonas climáticas de Colombia. Se probaron perceptrones multicapa (MLP), redes recurrentes (RN), Generalized Feedforward (GFF), Time Lagged Recurrent Networks (TLRN), Time Delayed Neural Networks (TDNN) y Radial Basis Function (RBF). Se encontraron modelos RNA que superaron métodos lineales y que simularon mejor los datos de anomalías de la temperatura media diaria que el reanálisis NCEP/NCAR. Posteriormente se hizo una proyección de la temperatura media diaria en el periodo del 1 de enero de 2001 al 31 de diciembre de 2100 bajo los escenarios A2 y A1B descritos por el Panel Intergubernamental sobre el Cambio Climático.</p>		

Autorizo (amos) a la Biblioteca Octavio Arizmendi Posada de la Universidad de La Sabana, para que con fines académicos, los usuarios puedan consultar el contenido de este documento en las plataformas virtuales de la Biblioteca, así como en las redes de información del país y del exterior, con las cuales tenga convenio la Universidad.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, "Los derechos morales sobre el trabajo son propiedad de los autores", los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.

DESARROLLO DE UN MODELO DE RED NEURONAL
ARTIFICIAL PARA LA REDUCCIÓN DE ESCALA
(DOWNSCALING) DE DATOS DE TEMPERATURA DEL MODELO
CLIMÁTICO GLOBAL CANADIENSE 3.1 A ESTACIONES
METEOROLÓGICAS COLOMBIANAS

Disertación
Énfasis Biosistemas
Maestría en Diseño y Gestión de Proceso

Por: Andrés Cardozo Vásquez
Proyecto de Grado

Facultad de Ingeniería
Universidad de la Sabana
Bogotá D.C., Colombia
Noviembre, 2.012

Director: M.Sc. Ing. Carlos Felipe Gaitán
Director: M.Sc. Ing. Luis Mauricio Agudelo

“Los verdaderos cambios medioambientales los lograremos nosotros. No podemos depender de nuestros líderes políticos. Tenemos que ver con claridad cuáles son nuestras responsabilidades y cómo podemos lograr que se produzca el cambio.”
Severn Cullis-Suzuki, del panel asesor de la Cumbre Mundial, afirma que si queremos que cambie la situación, todos tenemos que poner de nuestra parte.

PREFACIO

La Tierra, entendida como un biosistema, es un complejo sistema en el que propiedades emergentes afectan su totalidad y dan lugar a causalidades circulares en las que el hombre juega un papel importante. La alteración generada por el ser humano al equilibrio del biosistema está provocando cambios en la biosfera de la tierra: cambios en la temperatura y las condiciones atmosféricas y climáticas. Resulta muy difícil percibir las consecuencias medioambientales globales de la actividad humana, porque los efectos de cada actividad local se diluyen por toda la biosfera y lo que amenaza el equilibrio del biosistema es la combinación y acumulación de todos los efectos. Por tanto, aparentemente ninguna manifestación concreta de desequilibrio es motivo de alarma. Pero esperar a tener la evidencia de que estamos destruyendo el equilibrio actual del biosistema supone esperar a que nosotros mismos seamos víctimas de dicha destrucción. Por esta razón debería ser de suma relevancia estudiar y entender cómo nuestra existencia está afectando el planeta.

El clima de la Tierra siempre ha sufrido cambios, la diferencia con la situación actual es que con la presencia del hombre y su desarrollo industrial, existe el riesgo de cambios rápidos en el clima debido a las emisiones crecientes de CO₂ y otros gases de efecto invernadero que harán que en los próximos cien años veamos el mayor calentamiento de la Tierra desde el final de la Era Glacial. Cuanto más producimos y consumimos, más influimos en el ambiente que nos rodea. Durante los últimos 50 años, por primera vez en la historia, hemos visto claros signos de la influencia humana en el medio ambiente del planeta; estamos creando problemas medioambientales que no sólo son locales, sino también globales. Uno de estos problemas medioambientales globales es el riesgo de un cambio climático producido por el hombre.

El estudio del cambio climático y en general el clima, es por ende una de las áreas de investigación más relevantes del ámbito científico. El clima está determinado por elementos como la temperatura, humedad, vientos, precipitaciones y la presión atmosférica que son modificados a nivel global por los factores como son la latitud, altitud, bosques y masas de agua. A nivel regional, la orografía de la zona es un factor determinante que adicionalmente afecta la nubosidad, vientos y lluvias.

En las últimas décadas los modelos climáticos globales (GCM por sus siglas en inglés) han demostrado la capacidad para proveer información de gran valor sobre los cambios climáticos potenciales futuros en la tierra. Actualmente para describir adecuadamente los fenómenos climáticos regionales o locales se requiere contar con información espacial y temporal de alta resolución. Aunque los modelos climáticos globales pueden proveer esta información satisfactoriamente, la resolución espacial a la cual pueden ser simulados los procesos atmosféricos está limitada computacionalmente a 2.75° lat. x 2.75° long. Una necesidad clara para estrechar la brecha entre los GCM y la resolución espacial de modelos de superficie (modelos climáticos que representan la interacción entre la atmósfera y la superficie terrestre) se hace evidente. La aplicación de técnicas de reducción de escala es particularmente necesaria en terrenos complejos donde la resolución espacial de las simulaciones de los RCM no describe completamente la variabilidad climática. En el presente trabajo un modelo de interface ha sido desarrollado para mejorar desajustes de la escala regional permitiendo la investigación del cambio de temperatura a escala local en 5 estaciones meteorológicas colombianas, un trabajo de relevancia para Colombia por sus condiciones climáticas y orográficas características.

La relevancia de este trabajo para el país consiste en su capacidad de mejorar la confiabilidad en los datos de pronóstico de temperatura media diaria, respecto a las

resoluciones y correlaciones que ofrecen los modelos climáticos globales o regionales. El conocimiento de tendencias futuras permite tomar decisiones sobre conservación, monitoreo e impacto sobre la población, los cuerpos de agua e influencia agroalimentaria. Inclusive se ha convertido en relevante para el estudio de efectos en cultivos como por ejemplo el café arábico donde aún hay un amplio campo de estudio según reporta Davis y otros (Davis, Gole , Baena, & Moat, 2012). En Colombia se utiliza el modelo climático regional (RCM por sus siglas en inglés) PRECIS para la generación de información climática, que es alimentado por la información del reanálisis NCEP/NCAR. Este RCM es de alta demanda computacional comparado a técnicas de reducción de escala estadística y su resolución de superficie es de máximo 50km. Los modelos presentados en este trabajo permiten la reducción de escala a puntos locales superficiales mejorando la resolución de los RCM con una demanda computacional baja, presentando beneficios sobre las herramientas actuales empleadas por el Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM). Hasta donde es del conocimiento del autor no existen trabajos de reducción de escala de variables climáticas basados en redes neuronales artificiales para Colombia.

Junto con el desarrollo del modelo de reducción de escala, se determinan las variables de mayor peso para explicar la variabilidad de la temperatura en las diferentes estaciones meteorológicas estudiadas lo que se espera contribuya al mayor entendimiento de los complejos sistemas superficie-atmósfera. Se describe ampliamente el desarrollo de los modelos, la selección de variables y las zonas climáticas estudiadas para una mayor comprensión de las diferencias entre localidades. Los modelos simulan las anomalías de la temperatura media diaria determinadas por un proceso de eliminación de la tendencia y la estacionalidad de la variable de estudio y utilizan como variables de entrada las variables de

salida del reanálisis NCEP/NCAR. Posteriormente se correlacionan las variables de entrada con las variables del modelo acoplado de circulación global canadiense versión 3.1 (CGCM 3.1) determinándose las limitaciones de los modelos para el pronóstico de la temperatura.

En este punto quiero agradecer a todas las personas e instituciones que colaboraron directa o indirectamente con el buen término de este proyecto. Primero, quiero agradecer al director y supervisor de proyecto Prof. M.Sc. Carlos F. Gaitán, quien no solo me dio la oportunidad sino que tuvo confianza en mí para participar en un estudio de este tipo. Su conocimiento en profundidad del tema aportó a mi aprendizaje técnico para el desarrollo y realización de ideas propias. Más allá de su apoyo en el proceso de esta tesis, tuve la oportunidad de ganar experiencia invaluable aplicable a muchas otras áreas de conocimiento. También agradezco al co-director del proyecto Prof. Ing. Luis Mauricio Agudelo por su interés en el proyecto y la segunda revisión del trabajo para el buen término del proyecto.

Adicionalmente quiero agradecer a Environment Canadá por la información de los modelos climáticos CGCM 3.1 y del reanálisis NCEP/NCAR. Así mismo, agradecer al Centro de Investigaciones Oceanográficas e Hidrográficas del Caribe (CIOH) por los datos de temperaturas medias diarias en las diferentes estaciones meteorológicas estudiadas.

Finalmente, gracias a Dios por darme la sabiduría y la paciencia para llevar a feliz término este trabajo, a mis padres por ser lo que son y a todos mis amigos que constantemente se interesaron y me animaron para no desfallecer por el camino. Por último, y no por ser menos importante, agradezco a mi novia Deisy Castillo por su apoyo incondicional, su paciencia y su comprensión. Es a ella a quien dedico este trabajo por el tiempo nuestro que se sacrificó durante mis estudios.

TABLA DE CONTENIDO

1.	Introducción.....	1
1.1.	Objetivo general.....	2
1.2.	Objetivos específicos.....	3
2.	Marco Conceptual.....	3
2.1.	Características generales de los modelos climáticos.....	7
2.1.1.	Modelos Climáticos Globales y Regionales.....	10
2.1.2.	Escenarios IPCC.....	13
2.1.3.	Efectos de la resolución espacial.....	16
2.1.4.	Técnicas de reducción de escala.....	17
2.2.	CLIMA COLOMBIANO.....	21
2.2.1.	Generalidades De La Región Andina Colombiana.....	21
2.2.2.	Clasificación Climática de Köppen-Geiger.....	27
2.2.3.	Medición del Clima en Colombia.....	30
3.	Regresión Lineal.....	31
3.1.	Definición de Regresión Lineal.....	31
3.2.	Regresión Lineal Simple.....	33
3.3.	Regresión lineal múltiple.....	33
4.	Redes Neuronales Artificiales en la Predicción de Series de Tiempo.....	34
4.1.	Definición de Red Neuronal Artificial.....	34
4.2.	Taxonomía.....	38
4.3.	Topología de las redes neuronales.....	39
4.3.1.	Perceptrón.....	40
4.3.2.	Perceptrón Multicapa (MLP).....	42
4.3.3.	Redes de Función de Base Radial (RBF):.....	44
4.3.4.	Redes Análisis de Componentes Principales (ACP).....	46
4.3.5.	Redes Dinámicas.....	48
4.4.	Paradigmas de aprendizaje.....	54
4.5.	Parámetros del Aprendizaje.....	56
5.	Aplicación de las RNA al Pronóstico de Temperaturas.....	60
5.1.	Selección Estaciones meteorológicas IDEAM.....	60
5.2.	Pre-procesamiento de datos.....	66
5.2.1.	Reconstrucción datos.....	66
5.2.2.	Separación anomalías.....	68
5.2.3.	Análisis exploratorio de información.....	71
5.2.4.	Organización de la información.....	72
5.3.	Selección de Arquitecturas.....	75
5.3.1.	Taxonomías y topologías evaluadas.....	76

5.3.2.	Desarrollo del experimento.....	77
5.3.3.	Resultados mejores arquitecturas	80
5.4.	Desarrollo del modelo compuesto	107
5.4.1.	Arquitecturas evaluadas.....	107
5.4.2.	Resultados Modelo Compuesto	109
5.5.	Comparación Modelos RNA contra Reanálisis NCEP/NCAR	111
5.6.	Pronóstico de la Temperatura media diaria	112
5.6.1.	Evaluación de variables de entrada para el Pronóstico.....	112
5.6.2.	Generación del pronóstico las temperaturas medias diarias para las estaciones de estudio	114
6.	Conclusiones y Perspectivas.....	119
	Bibliografía.....	132

LISTA DE TABLAS

Tabla 1: Predictores	73
Tabla 2: Características de las arquitecturas.....	80
Tabla 3: Comparativo desempeños modelos individuales y compuesto	110
Tabla 4: Mejores modelos seleccionados	111
Tabla 5: Datos de desempeño modelo compuesto con reducción de variables.....	115
Tabla 6: Datos estaciones meteorológicas de estudio	138
Tabla 7: Última línea de entrada/salida de la regresión stepwise.....	152
Tabla 8: Mejores modelos estación E1.....	154
Tabla 9: Mejores modelos estación E2.....	154
Tabla 10: Mejores modelos estación E3.....	155
Tabla 11: Mejores modelos estación E4.....	155
Tabla 12: Mejores modelos estación E5.....	155
Tabla 13: Modelos del modelo compuesto.....	159
Tabla 14: Correlación variables NCEP/NCAR vs. CGCM entradas E1	161
Tabla 15: Correlación variables NCEP/NCAR vs. CGCM entradas E2	161
Tabla 16: Correlación variables NCEP/NCAR vs. CGCM entradas E3	162
Tabla 17: Correlación variables NCEP/NCAR vs. CGCM entradas E5	163
Tabla 18: Correlación variables NCEP/NCAR vs. CGCM entradas modelo compuesto ..	163

LISTA DE FIGURAS

Figura 2-1: Esquema celda global tridimensional (Research 2012).....	9
Figura 2-2: Ejemplo de dominios del RCM (Rummukainen 2010).....	12
Figura 2-3: Clasificación climática Köpen-Geiger, mapa mundial.....	29
Figura 4-1: Bloques constitutivos de las RNA.....	35
Figura 4-2: RNA totalmente conectada y su matriz de pesos.....	39
Figura 4-3: Perceptron.....	41
Figura 4-4: Red de función de base radial (RBF).....	46
Figura 4-5: Red análisis de componente principales (ACP).....	48
Figura 4-6: Memoria gamma.....	50
Figura 4-7: Estructuras de memoria conexionista y la localización del dominio de frecuencia del polo.....	52
Figura 4-8: Uso de gamma kernels en una arquitectura MLP.....	53
Figura 4-9: Comportamiento del MSE para los conjuntos de entrenamiento y prueba.....	60
Figura 5-1: Reconstrucción series de tiempo.....	67
Figura 5-2: Ubicación celdas y estaciones de estudio.....	74
Figura 5-3: Posición estación E1 y numeración de variables.....	81
Figura 5-4: Mejores resultados modelos estación E1.....	82
Figura 5-5: Posición estación E2 y numeración de variables.....	87
Figura 5-6: Mejores resultados modelos estación E2.....	88
Figura 5-7: Posición estación E3 y numeración de variables.....	94
Figura 5-8: Mejores resultados modelos estación E3.....	94
Figura 5-9: Posición estación E4 y numeración de variables.....	99
Figura 5-10: Mejores Resultados modelos estación E4.....	100
Figura 5-11: Posición estación E5 y numeración de variables.....	103
Figura 5-12: Mejores resultados modelos estación E5.....	104
Figura 5-13: Comparación desempeño modelos individuales y compuesto.....	113
Figura 5-14: Correlación entre NCEP/NCAR y CGCM3.1 para las variables entrada.....	114
Figura 5-15: Pronóstico de temperatura media diaria estación E1.....	116
Figura 5-16: Pronóstico de temperatura media diaria estación E2.....	117
Figura 5-17: Pronóstico de temperatura media diaria estación E3.....	117
Figura 5-18: Pronóstico de temperatura media diaria estación E4.....	118
Figura 5-19: Pronóstico de temperatura media diaria estación E5.....	118

LISTA DE ANEXOS

Anexo 1: Datos estaciones meteorológicas de estudio.....	138
Anexo 2: Pruebas estadísticas alteracion series reconstruidas	139
Anexo 3: Gráficas sepración tendencia, estacionalidad y anomalías estaciones meteorológicas.....	144
Anexo 4: Analisis descriptivo de las anomalías de la estacione meteorológicas	145
Anexo 5: Análisis de componentes principales estaciones	150
Anexo 6: Regresion lineal stepwise.....	152
Anexo 7: Arquitecturas de modelos evaluados en el proyecto.....	153
Anexo 8: 10 mejores modelos por estación.....	154
Anexo 9: Series de tiempo modelos E3.....	156
Anexo 10: Modelos para modelo compuesto	159
Anexo 11: Analisis de sensibilidad modelo compuesto	160
Anexo 12: Correlación de variables NCEP/NCAR y CGCM3.1 para las variables de entrada de los modelos RNA	161

RESUMEN

Se desarrolló un modelo basado en redes neuronales artificiales (RNA) para el pronóstico de la temperatura media diaria a escala local en 5 zonas climáticas de Colombia, según la clasificación de Köppen-Geiger. Para el desarrollo del modelo se partió de los datos históricos de temperatura media diaria en 5 estaciones meteorológicas seleccionadas, reportados por el Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia IDEAM. Debido a la discontinuidad de los datos (periodos de tiempo sin reporte de datos de temperatura) inicialmente se hizo una reconstrucción de las series de tiempo basada en promedios móviles, se verificó y validó estadísticamente las series reconstruidas para no alterar la distribución poblacional de la muestra. A partir de las series reconstruidas se ajustó un modelo de serie de tiempo mixto para separar las anomalías de la temperatura y utilizar estos datos como variables de salida (valores deseados).

Los modelos RNA fueron desarrollados inicialmente para cada estación meteorológica (una variable de salida) y posteriormente se implementó un modelo compuesto con cinco variables de salida (una para cada estación meteorológica).

Se usaron regresiones lineales, simples, múltiples y stepwise para comparar contra los modelos RNA y determinar que efectivamente hubo una mejora en la simulación (no linealidad y complejidad del sistema) de las anomalías de temperatura media diaria. Se probaron 95 arquitecturas de redes neuronales con taxonomías recurrentes y feedforward y diversas topologías: perceptrones multicapa (MLP), redes recurrentes (RN), Generalized Feedforward (GFF), Time Lagged Recurrent Networks (TLRN), Time Delayed Neural Networks (TDNN) y Radial Basis Function (RBF). Se evaluaron dos algoritmos de

aprendizaje: Momentum y LevenbergMarquardt y dos métodos de presentación de los datos de aprendizaje: On line y Batch. Se hicieron grupos de entrenamiento, validación y prueba con las variables de entrada y salida de cada uno de los modelos RNA. El desempeño de los modelos se determinó mediante el coeficiente de correlación (r), el error medio absoluto (MAE) y el error medio cuadrático (MSE).

Se encontró que en todos los casos existieron modelos RNA que superaron las regresiones lineales y que simularon mejor los datos de anomalías de la temperatura media diaria que la variable No. 25 del reanálisis NCEP/NCAR correspondiente a la Temperatura superficial a 2 metros en la celda de ubicación de la estación meteorológica.

Posteriormente se hizo un pronóstico de la temperatura media diaria en el periodo comprendido entre el 1 de enero de 2001 y el 31 de diciembre de 2100. Para la selección del modelo de pronóstico se verificó la correlación de las variables de entrada de los mejores modelos individuales y compuesto, provenientes del reanálisis, con su respectiva variable en el modelo de circulación global canadiense versión 3.1 (CGCM 3.1). Se encontró una mejor capacidad del reanálisis para simular, a partir de los datos del CGCM 3.1, las variables de entrada del modelo compuesto. Se utilizaron las salidas requeridas del CGCM 3.1 pronosticadas para los escenarios futuros supuestos A2 y A1B descritos por el Panel Intergubernamental sobre el Cambio Climático (IPCC por sus siglas en inglés) en el periodo 1 de enero de 2001 a 31 de diciembre de 2100. Finalmente se presentan las gráficas de las temperaturas medias diarias en la ubicación de las estaciones meteorológicas estudiadas para los 100 años de pronóstico bajo los 2 supuestos escenarios.

1. INTRODUCCIÓN

En el siguiente trabajo se desarrollan una serie de modelos de redes neuronales artificiales como propuesta de reducción de escala (downscaling) para la variable de temperatura media diaria, a 2 metros sobre el nivel de la superficie, simulados a una escala regional con resolución de $3,75^\circ$ Lat. x $3,75^\circ$ Long. por el reanálisis NCEP/NCAR; para llevarlos a una escala local sobre 5 estaciones meteorológicas colombianas ubicadas en diferentes zonas climáticas según la interpretación de Köppen y Geiger. Con el proyecto se pretendió determinar si un modelo de redes neuronales artificiales acoplado al reanálisis NCEP/NCAR es capaz de mejorar la resolución de datos de temperaturas medias, estrechando la brecha entre la resolución computacional del reanálisis y la realidad local en las estaciones meteorológicas seleccionadas para el estudio.

Se decidió crear modelos de reducción de escala para cada una de las cinco estaciones y posteriormente un modelo con cinco variables de salida, una para cada estación.

Para corroborar que los modelos de redes neuronales artificiales son capaces de captar la no linealidad y complejidad de los datos de anomalías de la temperatura, estos se compararon contra modelos de regresión lineal simple, múltiple y Stepwise. La combinación de topologías y taxonomías de redes neuronales, algoritmos y métodos de presentación de datos de aprendizaje junto con las diferentes regresiones lineales generaron la evaluación de 100 modelos por estación meteorológica.

Estos modelos fueron evaluados bajo criterios de coeficiente de correlación, error absoluto y cuadrático medio en los grupos de validación y prueba para determinar el modelo que mejor simula la realidad de las anomalías de la temperatura. Basado en los mejores resultados de las estaciones individuales en cuanto a tipo de arquitectura, algoritmo

de aprendizaje y método de presentación de datos se propusieron nuevos modelos de redes neuronales, pero esta vez con cinco variables de salida, es decir, el modelo no simula independientemente las anomalías de temperatura para cada estación sino que simultáneamente simula las 5 estaciones.

Las variables de entrada a los modelos corresponden a las variables de salida del reanálisis para la celda en la cual se encuentra ubicada la estación de estudio y las celdas circundantes. Técnicas de eliminación/introducción progresiva (Stepwise) de variables, análisis de componentes principales y/o análisis de sensibilidad fueron aplicados en diferentes puntos del desarrollo de los modelos pues técnicamente un menor número de variables puede llevar a mejores convergencias en los modelos de redes neuronales.

Como capítulo final del trabajo y para cierre al objetivo general, basado en las variables del modelo acoplado de circulación global canadiense versión 3.1 que mejor fueron simuladas por el reanálisis NCEP/NCAR, se seleccionó un modelo de red neuronal en el que se hicieron pronósticos de temperatura media diaria en las 5 estaciones meteorológicas de estudio para el periodo de enero 1 de 2001 a diciembre 31 de 2100, con lo que se corroboró que es posible desarrollar una metodología que permite proporcionar escenarios futuros de temperatura media a escala local estrechando la brecha de la resolución escala global a modelos de superficie.

1.1. OBJETIVO GENERAL

Desarrollar un modelo de red neuronal artificial para 5 estaciones meteorológicas seleccionadas en diferentes zonas climáticas (según Koppen-Geiger) de Colombia que permitiera hacer pronósticos confiables, de temperaturas medias en el periodo de tiempo 2000 a 2100 bajo el escenario de clima A1B o A2

1.2. OBJETIVOS ESPECÍFICOS

- Evaluar la correlación y el error entre los datos históricos de temperatura media (observaciones) y el modelo para diferentes arquitecturas de RNA en cada una de las 5 estaciones meteorológicas seleccionadas.
- Determinar las 10 mejores arquitecturas de RNA que se ajusten con mayor correlación (r) y menor error promedio (MAE) en el grupo de validación a los datos históricos de temperaturas medias de cada estación meteorológica seleccionada.
- Determinar si existe una mejora en el ajuste de los datos históricos del modelo propuesto frente a los datos de temperatura obtenidos por reanálisis NCEP/NCAR.
- Comparar la arquitectura de la red que mejor se ajuste a los datos históricos de las 5 estaciones frente a una red neuronal artificial con 5 salidas (una para cada estación) entrenada y validada para ajustar con el menor error promedio (MAE) a los datos históricos evaluados.
- Proporcionar escenarios futuros coherentes, consistentes y posibles, de temperatura máxima y mínima para determinadas estaciones meteorológicas en el periodo 2010-2100 y comparar como las proyecciones para cada estación difieren de las del modelo GCM.

2. MARCO CONCEPTUAL

La población se ha multiplicado y la tecnología ha alcanzado una enorme y sofisticada producción de forma que se está presionando muchas partes del medio ambiente

terrestre siendo la atmósfera la zona más vulnerable. La atmósfera y sus procesos relacionados el estado del tiempo y el clima son de altísima importancia para la supervivencia del ser humano en la tierra. Aunque ha existido un entendimiento del efecto de los gases efecto invernadero desde el año 1896 con Svante Arrhenius, el riesgo de un cambio climático antropogénico no fue reconocido científicamente hasta la década de 1980 (Weart 2004). Hoy en día se reconoce que dado el reducido espesor atmosférico la alteración de algunos componentes moleculares básicos, que a su vez se encuentran en pequeña proporción, supone un cambio significativo y por tanto se explica el papel básico que estos gases tienen como reguladores de la temperatura del Planeta. Sin la presencia de los gases efecto invernadero la temperatura media de la tierra sería de -18°C en vez de 15°C .

La acumulación antropogénica de gases de efecto invernadero, además del aumento observado en la temperatura media atmosférica cerca de la superficie terrestre, ha estimulado actividades internacionales de investigación científica para analizar y comprender en qué medida y de qué forma, a diferentes escalas espaciales y temporales, los posibles impactos afectarán el sistema social y natural. Paralelamente, durante las últimas décadas, se ha suscitado una creciente necesidad de herramientas adecuadas para investigar los cambios climáticos futuros y sus impactos. Esto ha implicado investigación en el campo, así como una notable evolución de metodologías científicas, incluyendo modelos que permiten y apoyan la investigación (Hewitson and Crane 1996, Bronstert, Niehoff et al. 2002, Huth, Kyselý et al. 2003, Hartell and Pearman 2010). Un importante ejemplo de este progreso es la cantidad de variables físicas y fisiología que ha sido incorporada a los modelos de clima.

Los modelos numéricos representan la herramienta principal para la investigación de los impactos futuros sobre el cambio global (Hewitson and Crane 1996). Para desarrollarlos los expertos se basan en complejos cálculos matemáticos que realizan predicciones fiables de cómo será el ciclo climático en los próximos años, el aumento que tendrán las temperaturas a nivel oceánico, el deshielo de la capa de los polos, etc. Estos son algunos de los cálculos que realiza el Panel Intergubernamental sobre el Cambio Climático (IPCC por sus siglas en inglés). Para efectuar estas proyecciones, se basan en los modelos climáticos, que permiten predecir en un plazo determinado de tiempo resultados para el clima mundial y local sobre qué concentraciones de gases de efecto invernadero (GEI) absorberán la atmósfera, el suelo, la flora y los océanos. Por esto, resultan primordiales para la toma de decisiones a nivel mundial en materia energética, infraestructura, económica, de población y de ordenamiento territorial para los próximos años.

Los parámetros y requerimientos para proyectar condiciones climáticas futuras han sido altamente estudiados y claramente formulados por la comunidad científica. El progreso en las velocidades de computación en un factor de un millón desde la década de 1970 a la fecha, ofrece la posibilidad de incluir cada vez más variables y elementos de procesamiento en los modelos numéricos mejorando la resolución espacial y los periodos de tiempo de la simulación. Pero aun así, aún existen deficiencias y limitaciones en estos modelos. La resolución espacial a la que los procesos atmosféricos pueden describirse aún tiene limitaciones computacionales. Inclusive, hay deficiencias que hacen que los niveles de confianza de los modelos climáticos sean más altos en algunas variables (Ej.: temperatura) que en otros (Ej.: precipitación)(Randall, Wood et al. 2007).

La predicción de la temperatura es un factor crítico en modelos ambientales y ha sido usada en un amplio rango de aplicaciones debido a que tiene diferentes aplicaciones en

áreas industriales, agricultura y energía. Algunas de estas aplicaciones son: dimensionamiento de sistemas de aire acondicionado en edificios, diseño de sistemas de energía solar, predicción de desastres naturales como incendios forestales, etc. El trabajo de algunos autores basados en esta variable demuestra su importancia de estudio en diversos campos. Por ejemplo: evapotranspiración de cultivos (de Bruin, Trigo et al. 2010), balance energético (Savage, Everson et al. 2009), predicción de precipitaciones (Funk, Michaelsen et al. 2003), balance de agua (Vallet-Coulomb, Legesse et al. 2001), calentamiento global (Joubert and Tyson 1996), predicción de producción de cultivos (Balaghi, Tychon et al. 2008), detección de humedad en suelos áridos (Zhao, Yang et al. 2010), vectores de idoneidad del hábitat (Estrada-Pena, Horak et al. 2008, DeVisser, Messina et al. 2010, Moore and Messina 2010) y fenología (Piper, Smit et al. 1996).

Las técnicas computacionales en Informática, y en particular en los Sistemas Inteligentes, durante los últimos años han adquirido gran aplicabilidad en diversas áreas del conocimiento como lo son los problemas de predicción de temperatura y precisamente este tipo de modelos empíricos, como las Redes Neuronales Artificiales (RNA) se han utilizado como una alternativa apropiada de aproximación a los modelos físicos (Singh, Basant et al. 2009). La mayoría de estos modelos están basados en técnicas de computación neuronal, las cuales además de ser rápidas, dan resultados precisos. Por ejemplo, existen varios trabajos recientes que implementan Perceptrones Multicapa (MLP) a problemas de predicción de temperatura en diferentes escenarios y países (Ustaoglu, Cigizoglu et al. 2008, Smith, Hoogenboom et al. 2009). También, un tipo diferente de red neuronal, llamado red abductiva, fue aplicado con éxito a un problema de predicción de temperaturas con una frecuencia de tiempo horaria en datos de la ciudad de Seattle, USA (Abdel-Aal 2004). Otros tipos de redes neuronales como las Funciones de Base Radial también se les

ha encontrado aplicación y han sido objeto de estudio en los últimos años (Ruano, Crispim et al. 2006, Ustaoglu, Cigizoglu et al. 2008, Aminian 2010). Con la mejora de las capacidades y velocidades computacionales logrados a la fecha, los trabajos más recientes están basados en modelos de máquinas de aprendizaje (Goyal, Burn et al. 2012) y support vector regression (Paniagua-Tineo, Salcedo-Sanz et al. 2011).

2.1. CARACTERÍSTICAS GENERALES DE LOS MODELOS CLIMÁTICOS

Los modelos climáticos se han desarrollado a partir de versiones iniciales de modelos de predicción del tiempo. El Panel Intergubernamental sobre Cambio Climático (IPCC) define un modelo climático como una representación numérica del sistema climático basada en las propiedades físicas, químicas y biológicas de sus componentes, sus interacciones y procesos (Randall, Wood et al. 2007). Aunque ambos modelos aun comparten muchas características, aquellos usados para la predicción numérica del estado del tiempo típicamente están diseñados para correr a niveles de resolución espacial mayores de lo que es posible con las simulaciones climáticas. Aunque la primera generación de los modelos de circulación general (GCM) sobre simplificaba en extremo el complejo sistema climático, Manabe y Wetherald (1967) fueron capaces de simular el efecto del calentamiento global como una reacción a la duplicación de las concentraciones de CO₂ en la atmósfera usando un GCM. Las actuales proyecciones de clima hacen uso de modelos combinados atmosfera-océano con componentes químicos o bioquímicos interactivos. Estos modelos de circulación general atmosfera-océano (AOGCM) son los que mejor se ajustan la representación del sistema climático y su dinámica, desafortunadamente a un alto costo computacional (Randall, Wood et al. 2007).

Los modelos climáticos simulan matemáticamente los elementos climatológicos pasados, presentes o futuros de un territorio en intervalos de tiempo. No se trata de

pronosticar el clima, sino de indicar las tendencias a largo plazo. Por ejemplo, los modelos climáticos actuales calculan incrementos de temperatura de hasta 4°C en Europa para el 2100 (NOAA Research Team 2005). Los modelos climáticos correlacionan diversas y complicadas variables. Los procesos físicos se describen por separado en las diferentes esferas del sistema natural (atmósfera, océano, superficie terrestre, criosfera y biosfera), incluyendo interacciones complejas entre las esferas. Esto demanda herramientas de procesamiento altamente eficientes que consideren toda la corteza superficial terrestre. Además, se necesitan equipos de personas en todos los campos de la ciencia y estrategias de creación, exploración y mantenimiento de bases de datos a muy largo plazo. La descripción del proceso se basa en leyes fundamentales de la física (Ej.: Leyes de movimiento de Newton) representadas por ecuaciones matemáticas que son resueltas usando grillas espaciales de tres dimensiones sobre todo el globo (ver Figura 1). Las diferentes esferas se caracterizan por un tamaño de malla diferente y el número de capas verticales. Como salida del modelo se calculan vientos, transferencia de calor, radiación, humedad relativa e hidrología de superficie dentro de cada celda y evalúan las interacciones con los puntos vecinos (NOAA Research Team 2005).

El perfeccionamiento de los modelos climáticos refleja los avances experimentados por la comunidad científica en el conocimiento del comportamiento atmosférico. De esta forma, los resultados de las predicciones cada vez están más de acuerdo con los efectos climáticos reales del CO₂. Los niveles atmosféricos de CO₂ han subido su nivel de 280 partes por millón por unidad de volumen, en los tiempos preindustriales, a las actuales más de 370 partes por millón, este incremento sigue en marcha. Que los niveles actuales se doblen, significará sin duda un incremento de las temperaturas globales. Pero los modelos

no siempre están de acuerdo en el impacto producido por otros fenómenos, como las nubes, el hielo marino y otras secciones del sistema climático.

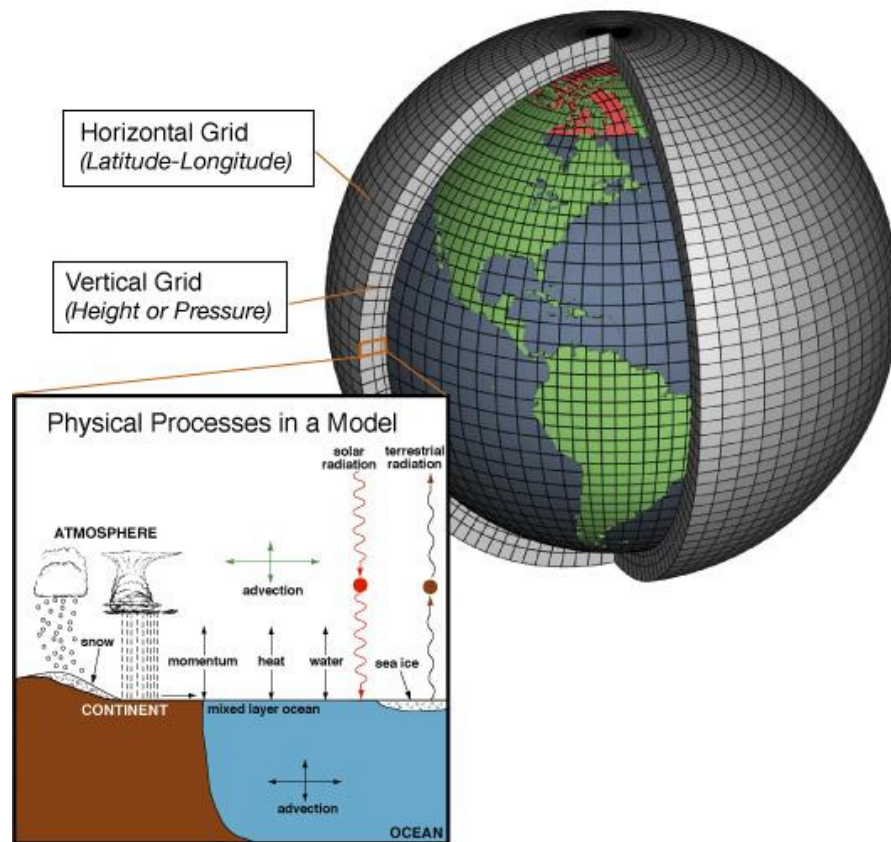


FIGURA 2-1: ESQUEMA CELDA GLOBAL TRIDIMENSIONAL (NOAA RESEARCH 2012)

Según el IPCC, dentro de los modelos climáticos, los modelos numéricos llamados GCM actualmente son las herramientas más avanzadas disponibles para simular la respuesta del clima global al incremento de las concentraciones de gases efecto invernadero. Sólo los GCM, posiblemente en conjunto con modelos regionales (Capítulo 2.1.1.2), tienen el potencial para proveer estimados regionales de cambio climático consistentes geográfica y físicamente, requisito fundamental para hacer análisis de impacto ambiental (Randall, Wood et al. 2007).

2.1.1. Modelos Climáticos Globales y Regionales

2.1.1.1. Modelo Climático Global Canadiense CGCM3.1

El Centro Canadiense para la Modelación y Análisis del Clima, CCCma (Canadian Centre for Climate Modelling and Analysis) ha desarrollado a través del tiempo un número importante de modelos climáticos (AGCM, AGCM2, AGCM3, AGCM4/CanAM4, CGCM1, CGCM2, CGCM3 y CGCM4/CanCM4). Estos modelos se utilizan para el estudio de la variación y cambio del clima, y para entender los diferentes procesos que gobiernan los sistemas climáticos. También se utilizan para hacer proyecciones cuantitativas de cambios futuros en el clima en periodos de tiempo prolongados (suponiendo diversos escenarios de emisión de gases efecto invernadero) en escalas de tiempo de inclusive décadas.

El Modelo Climático Global Canadiense CGCM3 es la tercera generación de estos modelos y actualizó la versión anterior (Second Generation Coupled Global Climate Model) con un nuevo componente atmosférico, sin alterar el componente oceánico (Flato and Boer 2001). Ha sido llevado a dos resoluciones, la versión T47 con un tamaño de malla de 3,75 grados para la resolución atmosférica horizontal y 1,85 grados para la resolución del componente oceánico y la versión T63 con una resolución atmosférica de 2,8 grados subdividido en 6 grillas oceánicas.

Los resultados preliminares que ha obtenido indican que las temperaturas globales podrían crecer más de lo indicado por la anterior versión del programa si la sociedad continúa emitiendo grandes cantidades de dióxido de carbono hacia la atmósfera. El CGCM3 revela que la temperatura media global podría aumentar en unos 2,6°C, en el hipotético escenario de que los niveles atmosféricos de CO₂ se dupliquen respecto a los actuales. Esta cifra es superior a los 2°C pronosticados por la versión anterior.

El CGCM3.1 es uno de los principales modelos climáticos mundiales. Incorpora fenómenos tales como los efectos de las erupciones volcánicas en los patrones de temperaturas o el impacto de la movilidad del hielo marino en la cantidad de luz solar absorbida por los océanos. El modelo funciona resolviendo fórmulas matemáticas que representan a los procesos químicos y físicos que conducen el clima terrestre, para cientos de puntos en la atmósfera, los océanos, el hielo marino y la superficie terrestre. El CGCM3.1 es tan complejo que precisa de 3 mil millones de cálculos para simular un solo día de clima global. Por eso requiere grandes superordenadores. Con el CGCM3.1 se han añadido cuatro veces más puntos para la tierra y la atmósfera que en la anterior versión del sistema. Por tanto, produce mucha más información acerca de las variaciones regionales en el clima. También ha aumentado la precisión en cuestiones tales como las temperaturas continentales y la alta atmósfera.

2.1.1.2. Modelos Climáticos Regionales

Los modelos climáticos regionales (RCM) son modelos de clima de alta resolución que cubren un área limitada del globo, típicamente 5000 Km. X 5000 Km., con una resolución horizontal típica de 50 kilómetros. Estos son aplicados a ciertos períodos de interés, gobernados por las condiciones de temperatura superficial del mar (TSM), hielo de mar y también condiciones laterales de frontera (por ejemplo: vientos y temperatura) (Jones, Noguer et al. 2004). Son modelos complementarios a los Modelos Climáticos Globales GCM. Su propósito principal es dar mayor resolución a los análisis o simulaciones climáticas globales, o estudiar procesos climáticos más detallados en regiones específicas del planeta. En los últimos veinte años, el desarrollo de estos modelos regionales dio lugar a una cada vez mayor resolución, corridas del modelo más largas y ha marcado el camino hacia el modelamiento regional de sistemas (Rummukainen 2010). Los

RCM son una herramienta fundamental para comprender el clima como método de investigación complementaria a los GCM, permitiendo un proceso más detallado de las condiciones regionales e incluso locales. De esta manera, representan una contribución esencial al estudio y simulación del cambio climático y estudios de impacto, así como a la planificación de adaptaciones por posibles daños u oportunidades relacionadas a la variabilidad y cambio climático.

Algunos ejemplos de los dominios de los RCM aparecen en la Figura 2-2; **Error!** **No se encuentra el origen de la referencia.** De estos, algunos están en uso como, NARCCAP, CLARIS, ENSEMBLES y RCMIP. Otros dominios todavía se encuentran en desarrollo, tal como en el Ártico y la Antártida.

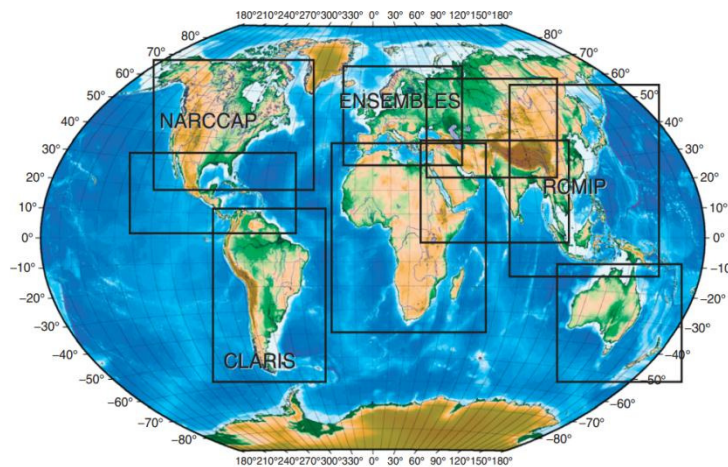


FIGURA 2-2: EJEMPLO DE DOMINIOS DEL RCM (RUMMUKAINEN 2010)

Los modelos regionales, son también conocidos como integrales, coherentes e interpoladores de base física o, en su término más popular, lupas. Su descripción de los procesos climáticos es tan compleja como en los modelos de circulación general integral. Pueden en principio ser aplicados a afinar los datos, independientemente de su resolución. Su desventaja es que ellos heredan los errores de los modelos de gran escala del que se

derivan y requieren grandes cantidades de datos de frontera previamente logrados en experimentos relevantes con los GCM (Jones, Noguer et al. 2004).

2.1.1.3. Reanálisis NCEP/NCAR

Durante el periodo comprendido entre 1991 y 1993, tres de los más grandes centros mundiales de investigación en predicción iniciaron (y aun hoy continúan) los proyectos denominados “reanálisis”. En estos proyectos, todas las posibles observaciones de diferentes tipos compiladas en archivos por décadas son procesadas en un sistema de asimilación de información de cuatro dimensiones. El sistema de asimilación produce un amplio número de campos en ambos sentidos, variables de entrada y sus derivaciones, que pueden ser usadas en diferentes aplicaciones como: investigación en cambio climático a diferentes escalas de tiempo, estudios de interacción atmósfera-océano, análisis de ciclos hidrológicos, etc. (Rubinstein and Sterin 2002). En particular el reanálisis iniciado entre dos entidades NCEP y NCAR tiene como objeto recolectar información de análisis globales en temas atmosféricos para soportar las necesidades de investigadores y comunidades de monitoreo del clima. Este esfuerzo involucra la recolección de datos de superficie terrestre, barcos, aviones, radiosondas, satélites, entre otros. Lo importante es que se controla estrictamente la simulación de estos datos con un sistema de asimilación que se mantiene sin alteraciones durante el periodo de reanálisis, eliminando los saltos climáticos observados asociados a cambios en el sistema de datos (Kalnay, Kanamitsu et al. 1996).

2.1.2. Escenarios IPCC

Los modelos climáticos están diseñados para modelar y cuantificar la respuesta del clima a las actividades actuales o futuras de la humanidad. Para lograr la confiabilidad en las simulaciones de los modelos climáticos, estos tienen que probar su habilidad para reproducir condiciones de clima pasadas y presentes. Dado un desempeño satisfactorio del

modelo, estas simulaciones sirven como línea base para comparar posibles resultados de condiciones atmosféricas futuras.

Para la generación del posible comportamiento del clima en el futuro, los modelos climáticos son forzados bajo diferentes escenarios de aerosoles y efectos gas invernadero. Los escenarios proponen diferentes perfiles del comportamiento futuro basándose en diferentes suposiciones sobre las emisiones de gases a la atmósfera que son relevantes al cambio climático. Por supuesto la estimación de los escenarios futuros viene con un número importante de incertidumbres incluyendo tanto el rango de las emisiones y la concentración de gases en el futuro, como también la reacción del clima a estas condiciones de frontera. Para lograr una coordinación de los esfuerzos en investigación de cambio climático y para que los resultados obtenidos por los diferentes grupos de investigación sean comparativos, se han desarrollado supuestos sobre estándares internacionales de la concentración de gases efecto invernadero. En el año 2000 el IPCC, que representa a los científicos líderes en el campo de cambio climático, presentó un conjunto de escenarios de emisiones en su Reporte Especial sobre Escenarios de Emisión (SRES por sus siglas en inglés) (IPCC 2000).

Los escenarios están basados en diferentes historias del desarrollo socioeconómico y demográfico cubriendo un amplio rango de futuros consistentes y factibles. Dependiendo de varias suposiciones relacionadas al crecimiento económico, intensidad y eficiencia energética así como también en el crecimiento de la población mundial; cuatro familias de escenarios se han propuesto (A1, A2, B1, B2).

La familia A se caracteriza por el direccionamiento preferencialmente por factores económicos, mientras que la B asume factores ambientales como su principal conductor. Otra diferencia está dada por el número que acompaña la familia. Mientras que A1 y B1

son familias orientadas globalmente, los escenarios A2 y B2 siguen políticas más bien regionales.

La siguiente breve descripción da una idea de los escenarios y sus principales características presentadas por el IPCC (2000):

A1: La línea evolutiva A1 y el escenario asociado se basan en una expansión económica de prosperidad, junto con una rápida introducción de tecnologías nuevas y eficientes. La población mundial alcanza su máximo a mediados de siglo seguido por un descenso más adelante, hasta el año 2100. Después de su orientación, se supone un aumento de las interacciones culturales y sociales, así como una reducción sustancial de las diferencias regionales en el ingreso per cápita. La familia de escenarios se distingue entre tres: direcciones del cambio tecnológico en el sistema energético representado por tres diferentes grupos de escenarios. Mientras que para el grupo de escenarios A1T se pone énfasis tecnológico en fuentes no fósiles de energía, el grupo de escenarios A1FI supone un uso intensivo de combustibles fósiles como fuente de energía. No dependiendo de una fuente de energía, el grupo de escenarios A1B asume un trabajo equilibrado y un mayor desarrollo de todas las fuentes de energía disponible.

A2: El argumento A2 y el relato de su escenario proponen un mundo muy heterogéneo caracterizado por una sociedad dispuesta a preservar las identidades locales. La población está en continuo crecimiento debido a una convergencia retrasada de patrones de fecundidad en todas las regiones. El desarrollo económico se lleva a cabo a nivel regional con crecimiento económico per cápita y cambios en tecnología tomando mucho más tiempo en comparación con otros escenarios.

B1: El escenario B1 y su historia muestran un mundo convergente con crecimiento poblacional similar al del escenario A1. Las estructuras económicas se desarrollan

paralelamente con una economía de información y servicio, junto a la introducción de tecnologías limpias y eficientes. Soluciones sociales, económicas y medioambientales son metas globales.

B2: Este escenario muestra un mundo en donde las metas en sostenibilidad económica, social y ambiental se persiguen a nivel local y regional. La población mundial está en constante crecimiento, aunque no tan rápido como el escenario A2. El desarrollo económico es menos distintivo y el cambio técnico toma más tiempo y es menos diverso comparado con los escenarios B1 y A1B. Los escenarios describen posibles evoluciones futuras de las fuerzas conductoras de las emisiones de gases efecto invernadero. Todos los escenarios asumen que iniciativas climáticas adicionales no se implementarán y que ninguno de los actuales propósitos (Ej.: Los acordados en el Protocolo de Kioto) se alcanzará. El IPCC hace especial énfasis en que todos los escenarios son igualmente probables.

2.1.3. Efectos de la resolución espacial

Actualmente existen limitaciones en la resolución espacial a la que algunos fenómenos climáticos pueden ser descritos en los modelos climáticos. Esto conlleva a que se tiene incertidumbre en los resultados de los modelos. Una razón de este problema es que los fenómenos meteorológicos solo se pueden resolver si su extensión espacial excede el tamaño matemático de la celda en un factor de cuatro (Von Storch, Zorita et al. 1993). Como consecuencia muchos fenómenos físicos que ocurren a escalas menores, como por ejemplo los relacionados con nubes, no pueden ser modelados apropiadamente y sus propiedades conocidas deben ser promediadas sobre una escala mayor, una técnica conocida como parametrización.

Otro efecto de una resolución espacial extensa es que las heterogeneidades de la superficie terrestre no pueden ser adecuadamente representadas en los modelos climáticos. En la actualidad, la superficie de la tierra en los modelos climáticos es, a menudo, representada por el tipo de la superficie terrestre más común en el área cubierta por la celda.

También es un problema que la baja resolución implica una representación burda de la topografía. El resultado es una discrepancia local en la altura entre la orografía de la simulación y la real (Fruh, Schipper et al. 2006). Como consecuencia las simulaciones de clima (Ej.: precipitación) difieren de las observaciones (Daly, Neilson et al. 1994).

Es así que los cambios climáticos estimados en escalas espaciales pequeñas (regionales o locales) están impregnados con una cantidad de incertidumbre considerable, derivada de diferentes fuentes. En estimaciones basadas directamente de las salidas de modelos climáticos globales, se encuentran diferentes niveles de incertidumbre asociados a: (i) Los escenarios (variabilidad interescenario), (ii) el uso de diferentes GCM (variabilidad intermodelo) (iii) diferentes corridas en un GCM (variabilidad intrínseca) (Giorgi and Mearns 2003).

El efecto neto de la resolución espacial actual es que las consecuencias locales de los cambios futuros de clima, son difíciles de investigar con las limitaciones actuales (Lozán, Grassl et al. 2001). Mientras que los modelos globales se vuelven prácticos computacionalmente hablando a medio plazo, las técnicas de reducción de escala actualmente son las más usadas para lograr mejor resolución.

2.1.4. Técnicas de reducción de escala

La reducción de escala es un método para obtener información de alta resolución acerca del clima o del cambio climático derivada de información climática global con

resolución más amplia. Estos métodos para mejorar la resolución de campos climáticos son importantes tanto para evaluar los impactos regionales del cambio climático global como para permitir una simulación realista de variables hidrológicas derivadas de la unión del clima y modelos numéricos hidrológicos. Generalmente, dos enfoques básicos son implementados para estimar el campo a nivel regional o a escala de cuenca: Reducción de escala dinámica (DD por sus siglas en inglés) y Reducción de escala Estadística (EDS). La DD hace uso de modelos climáticos regionales usando como condiciones de frontera los campos del GCM (la celda más cercana al área de estudio) (Giorgi and Mearns 1991), mientras que la EDS emplea técnicas estadísticas para establecer relaciones empíricas entre variables simuladas en baja resolución (predictores) y variables a nivel regional a escala local (predictandos) (Bronstert, Niehoff et al. 2002).

La reducción de escala dinámica ajusta las salidas de los GCM en modelos meteorológicos regionales. Más que usar ecuaciones para bajar proyecciones de escala global a nivel regional se usa modelación meteorológica para reflejar cómo los patrones globales afectan las condiciones del tiempo. El nivel de detalle depende de las capacidades computacionales. Sin embargo, los climatólogos consideran generalmente tres décadas como el mínimo para deducir las condiciones climáticas del comportamiento del clima. La cantidad de cálculos que involucra la reducción de escala dinámica hace muy dispendiosa la tarea de producir simulaciones de décadas con diferentes GCM o escenarios múltiples (Maurer and Hidalgo 2008). Como resultado, la mayoría de la investigación para producir proyecciones regionales está enfocada en la reducción estadística.

El método alternativo de reducción de escala estadístico (EDS) es una aproximación práctica a las necesidades actuales en investigación del cambio climático y es eficiente en términos computacionales comparado con las técnicas DD (Hewitson and Crane 1996). La

aproximación se hace a través de relaciones cuantitativas entre las observaciones de circulación a alta escala y el clima de escala local. Estas técnicas empíricas pueden dividirse en tres subcategorías: generadores climáticos, modelos de clasificación y funciones de transferencia.

Un generador de clima estocástico es un modelo numérico que produce series de tiempo sintéticas de un vector de variables del tiempo, como precipitación, temperatura y radiación solar; con ciertas propiedades estadísticas (Semenov, Brooks et al. 1998). La principal ventaja de este acercamiento es que los generadores climáticos pueden simular varias “realizaciones” del tiempo y por tanto proveer un amplio rango de situaciones factibles. Ellos representan una herramienta computacional de bajo costo para producir escenarios de cambio climático específicos al sitio con frecuencia diaria (Semenov, Brooks et al. 1998).

Aunque el principio general detrás de los métodos de clasificación es sencillo, la implementación práctica puede ser algo complicada. Desarrollado un esquema de clasificación de circulación atmosférica para un área de interés, un grupo de observaciones meteorológicas es distribuido en clases definidas. Las circulaciones simuladas por el GCM es enlazado con una de las clases definidas, incluyendo las observaciones subyacentes de la variable local (Zorita and Von Storch 1999).

La función de transferencia es uno de los primeros métodos de reducción de escala y es considerado de importancia práctica en el contexto de cambio climático (Kim, Chang et al. 1984). La técnica está basada en relaciones lineales y no lineales entre la variable predictora y una variable local denominada como predicando. La función de transferencia de alguna forma $y=f(X)$ se usa para derivar el clima en escala local a partir de simulaciones de amplia escala, asumiendo que la relación es válida para la simulación y que se mantiene

estable aunque el clima cambie. El método no necesariamente implica usar la misma variable como predictor y predicando. Como ejemplo, Von Storch y otros (Von Storch, Zorita et al. 1993) relacionaron la lluvia de invierno en la Península Ibérica a patrones de presión al nivel del mar en el Atlántico Norte usando una técnica de correlación canónica.

En la práctica, las técnicas de reducción de escala no están restringidas al uso de una sola de ellas, sino más bien a la combinación de ellas. Aunque todas estas técnicas ayudan a incrementar el detalle de las simulaciones climáticas, dependen altamente de la calidad de la información meteorológica que entregue el modelo de mayor escala. Es por esto que primero se debe verificar si las relaciones son válidas en el pasado y para ello se pueden usar los datos del reanálisis.

El enfoque en redes neuronales artificiales (RNA) ha sido aplicado para muchas aplicaciones de EDS y existen varios ejemplos disponibles en la literatura (Wetterhall, Halldin et al. 2005, Khan, Coulibaly et al. 2006). Por ejemplo, se ha demostrado la flexibilidad de una RNA para la reducción de escala de los datos de temperatura del modelo acoplado de circulación global canadiense a la estación meteorológica de Alberni Robertson Creek en Vancouver (Gaitan Ospina 2009).

Otros trabajos (Tolika, Anagnostopoulo et al. 2008) muestran la capacidad de las RNA para reducir de escala datos de precipitación y temperatura a partir de datos de GCM. Inclusive Tomasconjuntoti y otros (Tomassetti, Verdecchia et al. 2009) mostraron en su trabajo una reducción de escala de 27km a 3km a través del uso de un modelo de RNA partiendo de datos climatológicos del modelo numérico a mesoescala MM5.

Un modelo de regionalización estadística que emplee redes neuronales artificiales (RNA) es considerado analógicamente como uno de regresión múltiple aunque, por su carácter no lineal, hace que sean más eficientes en la representación de relaciones a partir

de datos ruidosos (Hewitson and Crane 1996). Las características no lineales de los datos no requieren ser pre-establecidas ya que la RNA busca la mejor relación posible y proporciona la función más óptima entre los predictores y los predictandos (Crane and Hewitson 1998). En otras palabras, no se requiere elegir una función de transferencia inicial y el sistema se auto-entrena (Cavazos 1999).

2.2. CLIMA COLOMBIANO

2.2.1. Generalidades De La Región Andina Colombiana

La Cordillera de los Andes forma, en el territorio Nacional de Colombia, tres ramales: a) la Cordillera Oriental con una extensión de 1.500 Kilómetros de longitud, 200 Kilómetros de ancho en su parte más grande y a una altura en promedio de 1.500 y 4.500 metros, b) Cordillera Central con 850 Kilómetros de extensión, un ancho aproximado de 100 Kilómetros y alturas entre los 3.000 y 5.100 metros, y c) la Cordillera Occidental con 820 Kilómetros de extensión, 70 Kilómetros de ancho y entre 1.500 y 5.000 metros de altura. El eje principal de la cadena montañosa se extiende de sur a norte y al nordeste (este último para la Cordillera Oriental).

Esta región montañosa que divide la región del Pacífico de la Orinoquia y de la Amazonia, es importante para el estudio del clima debido a que el incremento orográfico modifica las condiciones del aire procedente de estas partes y también afecta la meteorología del Magdalena y Cauca, que se encuentran en el mismo sentido que las cadenas montañosas.

Teniendo en cuenta que el clima establece en forma primordial la expansión de la vegetación, en la zona Andina se establecen cuatro grupos de vegetación: páramo, sub andino, ecuatorial y andino. Estos grupos están distribuidos en diferentes pisos térmicos de

forma desigual, porque las condiciones del clima en las faldas de la cordillera son distintas, por lo que se obtienen media docena de perfiles de variación climática y de vegetación por cada uno de los lados de las cordilleras. En algunos casos hay diferencias en la composición de la vegetación entre los pisos térmicos. Esto es debido a la manera como cada una de las vertientes es o no afectada por los vientos, como es el caso de los Alisios, las corrientes ascendentes húmedas desde el valle del Magdalena o determinados por cambios en el Océano Pacífico. Por esto el límite superior de la flora del ecuador sube a mayor altura en el occidente de la cordillera oriental.

El porcentaje de humedad da paso a que la temperatura descienda en forma parecida al grado húmedo, manteniendo una temperatura alta que cruza las laderas y permite el crecimiento de este tipo de flora en la región. Por otro lado hay otras características en el lado de la cordillera oriental que realizan la concentración de los pisos térmicos. En las faldas de la cordillera oriental hay más humedad que en la cordillera occidental, por otro lado, es más pronunciada y sus pendientes son más inclinadas lo que afecta los cambios de la temperatura con el ascenso, alcanzando menos temperatura a medida que se desciende, en el lado del Magdalena.

Las variaciones en los cambios de temperatura sobre la región Andina están relacionadas con el cambio de la altitud. La clasificación del clima está basada en los porcentajes de la humedad y las precipitaciones ocasionadas por los ciclos regionales y las zonas que la componen.

2.2.1.1. Regiones andinas del territorio amazónico y de la Orinoquía

Están condicionados por los vientos Alisios, en la Cordillera Oriental funciona como un límite que obliga al incremento forzoso de las corrientes de aire llevadas por las corrientes de aire Alisios del este y noreste. Por este motivo se determina un sistema

lluvioso que se va desde el piedemonte de los llanos y de la amazonia hasta llegar a los 2.500 metros de altitud. Según el Instituto Geográfico Agustín Codazzi (IGAC) (1989), desde las partes bajas y altas de esta región se presentan precipitaciones entre los 3.000 y 4.000 mm por año y a la mitad de ladera se puede presentar lluvias superiores a 4.000 mm por año. Mesa, Poveda y Carvajal (Mesa, Poveda et al. 1997), resumen que en estas franjas de territorio, la máxima caída de lluvia ocurre a una altura no superior a los 1.500 m. Si la precipitación es en las noches, esta elevación es menor, si en cambio sucede en las tardes el máximo se establece por encima de los 1.000 m. Sobre los 3.000 m de altitud, las precipitaciones son inferiores a los 1.000 mm por año. La temperatura se disipa dependiendo de las diferencias de altitud.

Piedemonte andino – pacífico

En esta parte del territorio se desarrollan sistemas convectivos que poseen un ciclo de vida de 18 a 36 horas. Estos sistemas dan pie a nubosidades con lluvias de variada intensidad y por consecuencia disminuyen el incremento de la temperatura en el día por la ausencia de la radiación solar directa. De una forma parecida en la vertiente del lado oeste en la Cordillera Occidental, las corrientes de aire cargadas de humedad que se desarrollan en el Océano Pacífico y que son llevadas hacia el lado este, presentan un incremento orográfico forzado, produciendo fuertes precipitaciones en este lado de la Cordillera. La mayor parte de la lluvia no está presente en las laderas (cosa contraria que en las laderas de la Cordillera Oriental, donde las precipitaciones llegan a ser superiores a los 4.000 mm) la localización en el piedemonte del Pacífico presenta precipitaciones superiores a los 6.000 mm. El cambio de la temperatura está regulado por el incremento de la altitud, sin embargo el alto porcentaje de humedad en la región del pacífico determina que la temperatura cambie de acuerdo con el porcentaje del gradiente adiabático húmedo.

Existe una estribación húmeda con caídas de lluvias que se localizan al norte de la Cordillera Occidental y la Cordillera Central que trasciende entre la región del Caribe y la Andina. El origen de esta concentración de humedad es principalmente por el viento bajo, que viene del Caribe. Existen dos mecanismos que inician las fuertes lluvias, en las horas del día, el incremento orográfico seco sobre las estribaciones determina las áreas con precipitaciones anuales en incremento excesivo de 4.000 mm en las laderas de la Cordillera Occidental y parte de la Cordillera Central. En la noche, las zonas bajas que se encuentran cercanas a las riberas del río Nechí con el Cauca, están viéndose afectadas por la declinación del aire frío que viene de la región montañosa de Antioquia, lo cual provoca una convergencia a bajo nivel (Mesa, Poveda et al. 1997). La temperatura se regula por el cambio de altitud, no obstante tiende a presentar un incremento alto de la temperatura de día por la influencia de las corrientes de aire ascendentes que provienen del Caribe. Las épocas secas y húmedas de esta zona están establecidas por el paso de la Zona de Confluencia Intertropical (ZCIT) con un par de períodos con lluvias, uno en el segundo trimestre del año y el otro en octubre.

Altiplano frío con circulaciones locales

Esta región está compuesta por los altiplanos de Cundinamarca, Boyacá, Paletará, y Túquerres e Ipiales. Están ubicados entre 2.500 y 3.000 m sobre el nivel del mar ubicándolos con unas temperaturas entre los 9° y 12°C, con alta predominancia de heladas en las épocas secas, las cuales poseen características de cielos abiertos o con mínima nubosidad y porcentaje de humedad en el aire muy escasa favoreciendo la pérdida de radiación terrestre entre las horas de la noche y madrugada. Las épocas con condiciones óptimas para las temperaturas por debajo de 0°C, se da entre diciembre y febrero para el centro del país, y entre los meses de la mitad del año para las regiones del sur. También

están influenciados por la ZCIT, la cual determina un par de períodos de lluvia que van entre marzo y mayo, y en octubre y noviembre; previsto por Mesa, Poveda y Carvajal (1997), las regiones de los altiplanos especialmente los de Cundinamarca y Boyacá presentan alto porcentaje de nubes en el día lo cual no es pre determinante frecuente de precipitaciones en comparación con las otras áreas en la región Andina. En estas regiones caen precipitaciones alrededor de 1.000 mm por año, y es normal que se presenten procesos convectivos locales que estimulan el desarrollo de nubes de lluvia y tormentas en la tarde. Para el caso de las regiones de Túquerres y de Paletará, el terreno es un componente que supedita el viento, por eso, aunque la dirección de los vientos durante el año es de oriente, hay meses que se presentan corrientes desde el norte y a veces del occidente, este último suceso está seguido a la influencia de vientos ascendentes desde el Océano Pacífico hacia el altiplano en Nariño, entre tanto que en el altiplano de Cundinamarca y Boyacá los vientos del occidente son influenciados por el valle y la montaña. Estos vientos suben desde el Valle del Magdalena, aproximadamente luego de las 11:00 de la mañana como resultado del incremento de temperatura diferencial entre estas dos regiones. Según Eslava(Eslava 1993), estas brisas son por el resultado de la diferenciación entre la temperatura del aire sobre las laderas y el del centro del valle, son corrientes de superficie que se forman incrementando por encima de las montañas durante el día y descendiendo hacia el valle por la noche. En las áreas más extensas de los valles se presentan nubes estratos durante la mañana y nubes de convección pronunciada durante las tardes y en las horas de la noche generando tormentas eléctricas.

Valles calientes y semihúmedos

Corresponden al Magdalena y el Valle del Cauca, es normal que sucedan brisas de valle y de montaña. Por otro lado, en los valles del Magdalena y del Cauca en su base

conducen la humedad hacia las vertientes de las cordilleras. Según Mesa, Poveda & Carvajal (1997), las precipitaciones lluviosas, derivadas de las vinculaciones del aire húmedo hacia la cordillera, acontece más que todo en las laderas y no en el fondo del valle. Entre 6° y 7° de latitud norte caen en promedio 3.000 mm por año.

Rio arriba por el Magdalena medio, el porcentaje de humedad disminuye encontrándose solo algunas zonas con más porcentaje de humedad hacia las estribaciones altas de la cuenca del Saldaña. Es de tener en cuenta la región de predominancia seca del fondo del valle, encontrada entre los departamentos del Tolima, Huila y Cundinamarca el cual se debe esencialmente a la fuerza de los vientos en el flanco occidental de la Cordillera Oriental. La región de alta pluviosidad encontrada en el nororiente de Antioquia se debe en principio a la convección de la noche ocasionada por los altiplanos del oriente, de donde llegan corrientes de aire frío que aumentan la convergencia en la superficie. Mesa, Poveda & Carvajal (1997), afirman, que el Valle del Cauca por estar más cerrado y alto, es más seco proporcionando lluvias al fondo del valle en una escala promedio de 1.000 mm por año, y en las laderas se convierten en cerradas bandas con lluvias de hasta 2.000 mm por año.

Alta montaña y vertientes húmedas

Se ubican por encima de 2.500 m de altitud, la temperatura oscila entre 12°C y 17°C en la parte baja hasta ser inferior a 6°C por encima de los 3.500 m. Estas corrientes se encuentran en la Cordillera Central y en la parte superior de la Cordillera Occidental. Para la primera hay dos condiciones que acceden a la predominancia de pluviosidad considerable en altura: La aportación de humedad de los valles y los canales de aire húmedos que vienen desde la Orinoquia y el Océano Pacífico mejorado por las alturas bajas de la Cordillera Oriental y Occidental.

En la Cordillera Occidental existe una estrecha brecha de pluviosidad superior a las que existen en el valle, lo cual está reflejado con la interdicción de la nubosidad de tipo orográfico que se extienden en la vertiente pacífica de la cordillera. De igual manera que toda la región Andina, estas vertientes están bajo efecto de la ZCIT con las épocas húmedas y secas anteriormente mencionadas.

Zonas andinas con efecto de la ZCIT

En general toda la región Andina está supeditada por el efecto de la ZCIT, las regiones mencionadas son aquellas que no dan efecto a parámetros específicos de rotación sino que sus épocas de lluvia y sequías, así como las corrientes de vientos se someten únicamente al comportamiento de la ZCIT. Estas regiones hacen parte al Nudo de los Pastos, al Macizo Colombiano, hacia el norte de la Cordillera Oriental y las partes mínimas en las tres cordilleras. Estas zonas están en un rango de altitud amplio, que va desde los 500 hasta los 5.000 m. Las caídas de lluvias, al igual que en toda la región Andina, no son altas y van entre los 600 mm en los valles secos del Patía y Chicamocha, 1.200 mm en la parte del centro de Santander y 2.200 mm en la región del Macizo Colombiano, esta última es una región productora de nubes de tipo orográfico, ya que suben por esta las corrientes de aire que son llevadas por los valles del Magdalena, Cauca, Patía y la región amazónica.

2.2.2. Clasificación Climática de Köppen-Geiger

El mapa de clasificación climática de uso más frecuente es el de Wladimir Köppen, presentado en su última versión 1961 por Rudolf Geiger. Basados en los datos de la Unidad de Investigación Climática (CRU por sus siglas en inglés) de la Universidad de Anglia Oriental y el Centro Climatológico Mundial de Precipitaciones (GPCC), se realizó un nuevo mapa digital de la clasificación mundial de Köppen-Geiger, en este se detalla una

malla de datos más precisa dirigidos por GPS (Global Positional System) con el cual se registran cambios y patrones de comportamiento climático sectorial. Aunque varios autores han publicado diversas clasificaciones Köppen mejoradas o nuevas, la distribución climática originalmente desarrollada por Köppen (aquí llamada como la clasificación de Köppen-Geiger), sigue siendo la clasificación del clima de uso más frecuente.

Algunos ejemplos del uso de esta clasificación son: los estudios hidrológicos de Peel et al. (2001) que identifican y explican la variabilidad de escala continental, en escorrentía anual. Aplicaciones de modelación del clima han sido presentados, por ejemplo, por Lohmann et al. (1993) para validar el control de un modelo de circulación general. Tanto Lohmann (1993) como Kleidon (2000) aplicaron la clasificación de Köppen a modelos de simulación para ilustrar la diferencias en los resultados de la simulación, los resultados aunque no del todo satisfactorios dentro de las expectativas, marcaron un leve cambio en el comportamiento de los cultivos en cuanto a rotación de tierras y tiempos de cosecha, datos que para la actualidad son importantes debido al aceleramiento en la emisión de gases de invernadero. La clasificación climática de Köppen-Geiger actualizada pretende continuar apoyando futuros estudios similares a los expuestos anteriormente.

Dos conjuntos globales de datos de observación del clima han sido seleccionados para actualizar el mapa histórico mundial con las regiones climáticas de Köppen-Geiger. Ambos están disponibles en una resolución regular de 0,5 grados de latitud / longitud. El primer conjunto de datos es proporcionada por la Unidad de Investigación Climática (CRU) de la Universidad de Anglia Oriental (Mitchell y Jones, 2005) a través de una red de estaciones de observación climática y meteorológica con datos mensuales de nueve variables. Este conjunto de datos cubre áreas a nivel mundial excluyendo la Antártida, está públicamente disponible (www.cru.uea.ac.uk) y se conoce como CRU TS 2,1. El segundo

conjunto de datos (Beck et al., 2005) es proporcionado por el Centro Climatológico Global de Precipitaciones (GPCC), ubicado en el Servicio Meteorológico Alemán. Esta grilla fue desarrollada sobre la base de datos de precipitaciones mensuales construida por el GPCC y cubre un conjunto global de datos con exclusión de Groenlandia y la Antártida. Todas las observaciones en estas estaciones, están sujetas a múltiples controles de calidad para minimizar el riesgo de generar homogeneidades temporales en los datos reticulados debido a las diferentes densidades de estaciones. Este conjunto de datos se conoce como VASClimO.

El Mapa Mundial de la clasificación climática de Köppen-Geiger que se presenta en la figura (Ver Figura 1-3) esta actualizado con la media mensual de temperatura de CRU TS 2.1 y la información v1.1 VASClimO correspondiente a datos de precipitación para el período 1951 a 2000 de forma regular 0,5 grados de latitud / longitud de la cuadrícula.

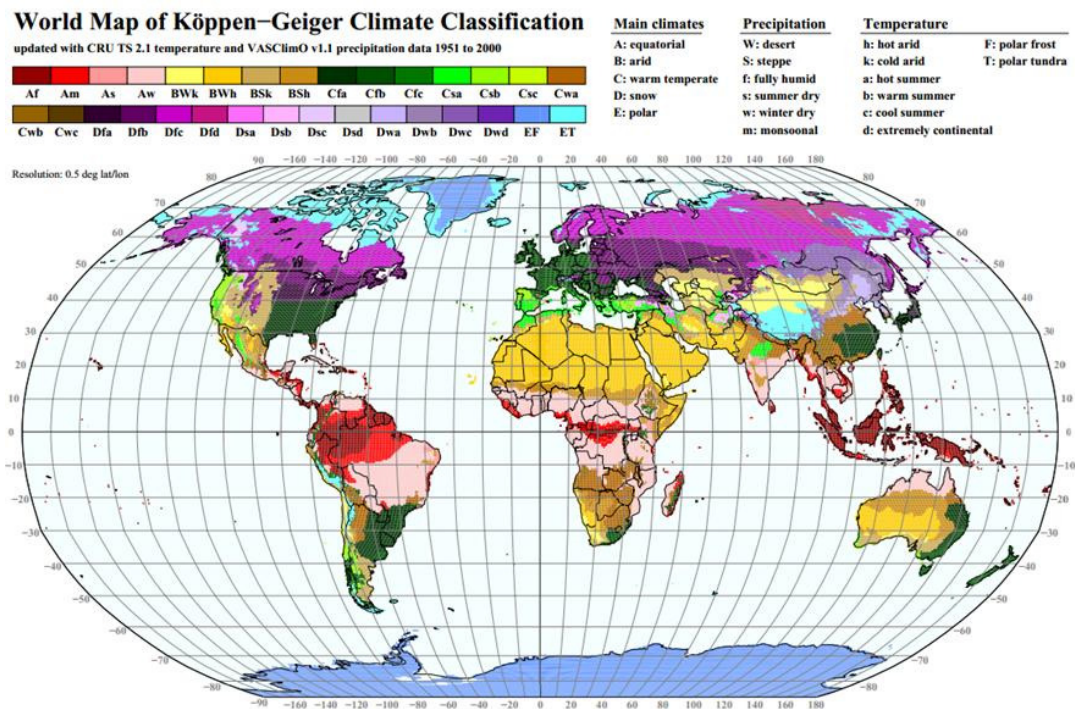


FIGURA 2-3: CLASIFICACION CLIMATICA KÖPPEN-GEIGER, MAPA MUNDIAL

Según la versión 2.006 del Mapa Mundial Climático de Köppen-Geiger (Ver figura 1-3) en Colombia se encuentran 6 zonas:

BSh: Clima principal, árido. Precipitación, estepa. Temperatura, árida caliente. Está ubicada en la alta Guajira.

Aw: Clima principal, ecuatorial. Precipitación, invierno seco. Comprende principalmente la zona norte del país y el departamento de Arauca y Casanare.

Am: Clima principal, ecuatorial. Precipitación, monzónica. Principalmente en los departamentos de Meta y Vichada y la parte norte de Antioquia.

Af: Clima principal, ecuatorial. Precipitación, completamente húmeda. Mayoritaria en el país comprendiendo la costa pacífica, el centro y sur del país exceptuando partes altas (cordilleras).

Cfc: Clima principal, temperado tibio. Precipitación, completamente húmedo. Temperatura verano frío. Ubicado en las partes altas de las cordilleras.

Csa: Clima principal, temperado tibio. Precipitación verano seco. Temperatura, verano caliente. Una pequeña zona en la cordillera central en los alrededores de Pasto.

2.2.3. Medición del Clima en Colombia

El Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia- IDEAM- obtiene, analiza, estudia, procesa y divulga la información básica sobre hidrología, hidrogeología, meteorología, geografía básica sobre aspectos biofísicos, geomorfología, suelos y cobertura vegetal para el manejo y aprovechamiento de los recursos biofísicos de la Nación y tiene a su cargo el establecimiento y funcionamiento de infraestructuras meteorológicas e hidrológicas nacionales para proveer informaciones, predicciones, avisos y servicios de asesoramiento a la comunidad. La predicción climática generada por el IDEAM se basa en el análisis de modelos procedentes de los centros

internacionales y de análisis nacionales del grupo de predicción climática. Combinando información de distintos centros meteorológicos de todo el mundo, con datos aportados por sus cerca de 3.000 estaciones dispersas por todo el país, y usando modelos matemáticos, el grupo puede acercarse con cierto grado de confianza tan solo a lo que sucederá en uno, tres y hasta seis meses. Los resultados emitidos por el IDEAM son generados con el modelo de mesoescala WRF (Weather Research and Forecasting) con datos iniciales del modelo global GFS de baja resolución, de las 00 UTC y llevados a una resolución espacial de 27, 20 y 5 kilómetros. Este modelo es la siguiente generación del modelo MM5.

3. REGRESIÓN LINEAL

3.1. DEFINICIÓN DE REGRESIÓN LINEAL

En estadística la regresión lineal o ajuste lineal es un método matemático que modela la relación entre una variable dependiente Y , las variables independientes X_i y un término aleatorio ε . En la regresión lineal, los datos se modelan utilizando funciones lineales de predicción, y los parámetros desconocidos del modelo se calculan a partir de los datos (Ver ecuación abajo). Estos modelos se denominan modelos lineales. Por lo general, la regresión lineal se refiere a un modelo en el que la media condicional de Y dado el valor de X es una función afín de X . Con menor frecuencia, la regresión lineal podría referirse a un modelo en el que la mediana, o algún otro cuantil condicional de la distribución de Y dado X se expresa como una función lineal de X . Como todas las formas de análisis de regresión, la regresión lineal se centra en la distribución de probabilidad condicional de Y dado X , en lugar de en la distribución de probabilidad conjunta de Y y X , que es el dominio del análisis multivariante.

$$Y_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon$$

Donde,

Y_i : es la variable dependiente, explicado o regresando

X_1, X_2, \dots, X_p : son las variables explicativas, independientes o regresores

$\beta_0, \beta_1, \beta_2, \dots, \beta_p$: son parámetros que miden la influencia que las variables explicativas tienen sobre el regresando donde β_0 es la intersección o término “constante”, la β_i ($i > 0$) son los parámetros respectivos a cada variable independiente a tener en cuenta en la regresión.

Este método matemático tiene muchas aplicaciones prácticas. La mayoría de las aplicaciones de regresión lineal tiene que ver con una de las siguientes dos categorías:

Si el objetivo es predicción o pronóstico, la regresión lineal se puede utilizar para ajustar un modelo predictivo a un conjunto de datos Y y X observados. Después de desarrollar dicho modelo, si un valor adicional dado de X aparece sin su respectivo valor de Y , el modelo ajustado se puede utilizar para hacer una predicción del valor de Y .

Dada una variable Y y un número de variables X_1, \dots, X_p que puede estar relacionadas con Y , el análisis de regresión lineal se puede aplicar para cuantificar la fuerza de la relación entre Y y X_j , para evaluar que X_j no tiene ninguna relación con Y y para identificar qué grupos de la X_j contienen información redundante sobre Y .

Para poder crear un modelo de regresión lineal, es necesario que se cumpla con los siguientes supuestos (Compostela 2003): la relación entre las variables es lineal, los errores en la medición de las variables explicativas son independientes entre sí, los errores tienen varianza constante (Homocedasticidad), los errores tienen una esperanza matemática igual a cero (los errores de una misma magnitud y distinto signo son equiprobables), el error total es la suma de todos los errores.

3.2. REGRESIÓN LINEAL SIMPLE

Tiene como objeto estudiar cómo los cambios en una variable, no aleatoria, afectan a una variable aleatoria, en el caso de existir una relación funcional entre ambas variables que puede ser establecida por una expresión lineal, es decir, su representación gráfica es una línea recta.

Sólo se maneja una variable independiente, por lo que sólo cuenta con dos parámetros de la forma:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Donde ε_i es el error asociado a la medición del valor X_i y siguen los supuestos de modo que $\varepsilon_i \sim N(0, \sigma^2)$ (media cero, varianza constante e igual a un σ y $\varepsilon_i \perp \varepsilon_j$ con $i \neq j$).

3.3. REGRESIÓN LINEAL MÚLTIPLE

La regresión lineal nos permite trabajar con una variable a nivel de intervalo o razón, así también se puede comprender la relación de dos o más variables y nos permitirá relacionar mediante ecuaciones, una variable en relación a otras variables llamándose Regresión múltiple. Constantemente en la práctica de la investigación estadística, se encuentran variables que de alguna manera están relacionados entre sí, por lo que es posible que una de las variables pueda relacionarse matemáticamente en función de otra u otras variables.

Maneja varias variables independientes. Cuenta con varios parámetros. Se expresan de la forma:

$$Y_i = \beta_0 + \sum \beta_i X_{ip} + \varepsilon_i$$

Donde ε_i es el error asociado a la medición i del valor X_{ip} y siguen los supuestos de modo que $\varepsilon_i \sim N(0, \sigma^2)$ (media cero, varianza constante e igual a un σ y $\varepsilon_i \perp \varepsilon_j$ con $i \neq j$).

4. REDES NEURONALES ARTIFICIALES EN LA PREDICCIÓN DE SERIES DE TIEMPO

4.1.DEFINICIÓN DE RED NEURONAL ARTIFICIAL

En la literatura no se encuentran definiciones claras o totalmente ajustadas a lo que se llama una Red Neuronal Artificial (RNA). Muchas de las definiciones explican una RNA refiriéndose a figuras en vez de darle una definición matemática explicativa (Guresen and Kayakutlu 2011). Las definiciones más ampliamente utilizadas hacen referencia a las RNA como unidades de procesamiento y algoritmos de aprendizaje (Bradley 1995, Muller, Reinhardt et al. 1995, Rojas 1996, Gurney 2003, Eberhart and Shi 2007). Cuando se habla de aprendizaje se definen como modificación de los pesos sinápticos (la fuerza de la conexión que se pueden adaptar de tal forma que la respuesta de la neurona coincida con la repuesta deseada) para capturar la información (Bradley 1995, Muller, Reinhardt et al. 1995, Rojas 1996, Gurney 2003, Eberhart and Shi 2007). Algunos adicionan a la definición, que las RNA pueden modificar su propia topología mientras aprenden (Bradley 1995, Muller, Reinhardt et al. 1995, Rojas 1996, Gurney 2003, Eberhart and Shi 2007).

Rojas (Rojas 1996) hace referencia a que aún no se tiene total entendimiento del mecanismo computacional de una neurona biológica y por tanto es preferible referirnos a Elementos de Procesamiento más que a neurona artificial. Las definiciones de RNA hacen énfasis en los Elementos de Procesamiento (EP) y el algoritmo de aprendizaje

Una buena definición es la propuesta por Bradley (Bradley 1995) que describe una RNA como una combinación masiva de unidades de procesamiento simples que pueden adquirir conocimiento del ambiente a través de un proceso de aprendizaje y almacenar el conocimiento en sus conexiones. Pero tal vez la definición más elaborada es la propuesta

por Guresen y Kayakutlu (Guresen and Kayakutlu 2011) quienes descomponen su enunciado en seis definiciones: arquitectura, elementos de entrada, elementos de salida, interconexiones, algoritmos de aprendizaje y adaptabilidad.

En términos prácticos se puede describir una Red Neuronal Artificial como un sistema de interconexiones, altamente distribuidas, de elementos de procesamiento de adaptación no lineal (EPs). Las RNA son un sistema de interconexión de EP en red que colaboran para producir un estímulo de salida (ver Figura 4-1).

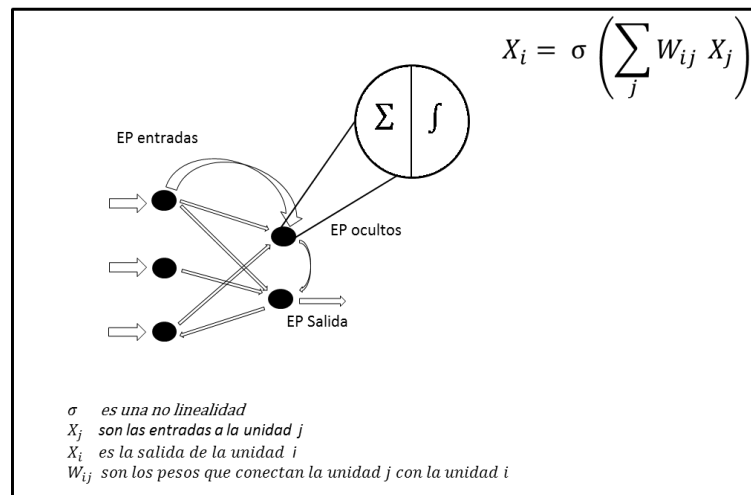


FIGURA 4-1: BLOQUES CONSTITUTIVOS DE LAS RNA

Estos sistemas de computación distribuida tienen ventajas como: fiabilidad, tolerancia a fallos, alto rendimiento (división de tareas de computación) y computación cooperativa. Pero generan problemas de localidad de la información y la subjetividad en la selección de la topología de interconexión (Khashei and Bijari 2012).

La adaptabilidad es la habilidad de cambiar los parámetros del sistema de acuerdo a algunas reglas (normalmente la minimización de una función de error). La adaptabilidad

permite al sistema buscar su desempeño óptimo, pero tienen problemas para responder de una manera repetible a cantidades absolutas.

La no linealidad de las RNA es una ventaja enorme en el campo de control dinámico de variables sin restricciones y produce sistemas de computación más potentes (en comparación con el procesamiento lineal), como la separación característica. Sin embargo, se complica el análisis teórico enormemente.

Estas características de procesamiento distribuido, adaptación y no linealidad, son el sello de los sistemas biológicos de procesamiento de información. Las RNA, por lo tanto, trabajan con los mismos principios básicos que los cerebros biológicos, pero probablemente la analogía debería detenerse aquí.

A diferencia de otros métodos de procesamiento de la información más analíticos, la computación neuronal explora la información contenida en los datos de entrada, sin necesidad de suposiciones adicionales como los métodos estadísticos que se basan en supuestos acerca de conjuntos de datos de entrada (es decir, probabilidades a priori, funciones de densidad de probabilidad, etc.). En el caso de la Inteligencia Artificial, se decodifica a priori el conocimiento humano con reglas simples SI ENTONCES (IF THEN), desarrollando una inferencia (búsqueda) en estas reglas para llegar a una conclusión. En cambio con las redes neuronales se descubren relaciones entre los datos de entrada, a través de la presentación iterativa de la información, y las características intrínsecas de la topología neuronal. Existen dos fases primordiales en la operación de redes neuronales. La fase de aprendizaje donde la información es presentada repetitivamente a la red mientras sus pesos se ajustan para obtener la respuesta deseada; y la fase de recuperación donde la red entrenada, con los pesos congelados, es aplicada a información nunca antes presentada. La fase de aprendizaje demanda tiempo debido a su naturaleza iterativa de búsqueda del

mejor desempeño, pero una vez se ha entrenado la red, la fase de recuperación puede ser muy rápida puesto que el procesamiento puede ser distribuido.

En general, las redes neuronales ofrecen soluciones viables cuando existen grandes volúmenes de información. Cuando es difícil (o imposible) formular un problema analíticamente y están las condiciones para obtener información experimental, entonces una solución de red neuronal es normalmente apropiada.

Las principales aplicaciones de redes neuronales son las siguientes:

Clasificación de patrones: La necesidad de clasificación de la información en clases es un problema común en el procesamiento de información. Se puede encontrar en control de calidad, predicción financiera, investigación de laboratorio, marketing dirigido, predicción de bancarrota, reconocimiento óptico de caracteres, etc. Las RNA de tipo feedforward, normalmente llamadas perceptrones multicapa (MLPs por sus siglas en inglés) han sido aplicadas en estas áreas pues son excelentes mapeadores funcionales (Lippmann 1989).

Memoria Asociativa: Los principios de la memoria humana parecen ser de este tipo. En la memoria asociativa las entradas son agrupadas por características comunes, o hechos relacionados. Las redes que implementan la memoria asociativa pertenecen generalmente al tipo de taxonomía recurrente, como la red de Hopfield o la memoria asociativa bidireccional (Anderson and Rosenfeld 1987, Kohonen 1988, Freeman and Skapura 1991).

Extractores de características: Un aspecto importante del tratamiento de la información es simplemente usar la información pertinente y desechar el resto. Esto normalmente se logra en una etapa de pre-procesamiento, donde la RNA se puede utilizar como analizador de componentes principales, cuantificadores vectoriales, o red de

agrupación. Los extractores de características se basan en la idea de competencia y normalmente tienen una topología simple de una capa (Kohonen 1988).

Redes dinámicas: un número importante de aplicaciones ingenieriles requieren el procesamiento de información que varía en el tiempo como por ejemplo el reconocimiento de voz, el control adaptativo, predicción de series de tiempo, pronósticos financieros y la modelación dinámica no-lineal. Para hacer frente a las señales que varían en el tiempo, las topologías de red neuronal se refuerzan con un mecanismo de memoria de corto plazo. Esta área aún se encuentra en fase de investigación pero parece ser prometedora para las redes neuronales (Vries and Principe 1992, Haykin 1999).

4.2. TAXONOMÍA

Como se mencionó anteriormente, una red neuronal no es más que una interconexión de elementos de procesamiento. La forma de estas interconexiones provee una de las variables claves para la división de las redes neuronales en familias. El caso más general es la red neuronal completamente interconectada. Por definición cualquier EP puede alimentar o recibir activaciones de cualquier otro EP incluyéndose a sí mismo. Por tanto, cuando los pesos de las conexiones son presentados en forma de matriz (pesos de la matriz), ésta estará completamente poblada. A manera de ejemplo, una red neuronal completamente interconectada de 6 EP se puede ver en la figura abajo (ver Figura 4-2)

Esta red es llamada Red Recurrente (recurrent network). En estas redes recurrentes algunas de las conexiones pueden estar ausentes, pero hay conexiones de retroalimentación. Una entrada presentada a una red recurrente en el tiempo t , afectará la salida de la red para futuros tiempos mayores a t . Por tanto, las redes recurrentes necesitan ser operadas a través del tiempo (Cao, Ewing et al. 2012).

Si la matriz de interconexiones está restringida a activaciones de alimentación directa (no hay retroalimentación ni auto-conexiones) la red neuronal se define como de alimentación directa (feedforward). Por tanto una red neuronal de este tipo es aquella donde sus EP no forman ciclos directos. Esto las diferencia de las redes recurrentes.

La red neuronal feedforward fue el primer tipo de red y posiblemente el más simple ideada. En esta red, la información se mueve en una sola dirección, hacia adelante, desde los nodos de entrada, a través de los nodos ocultos (si lo hay) y los nodos de salida.

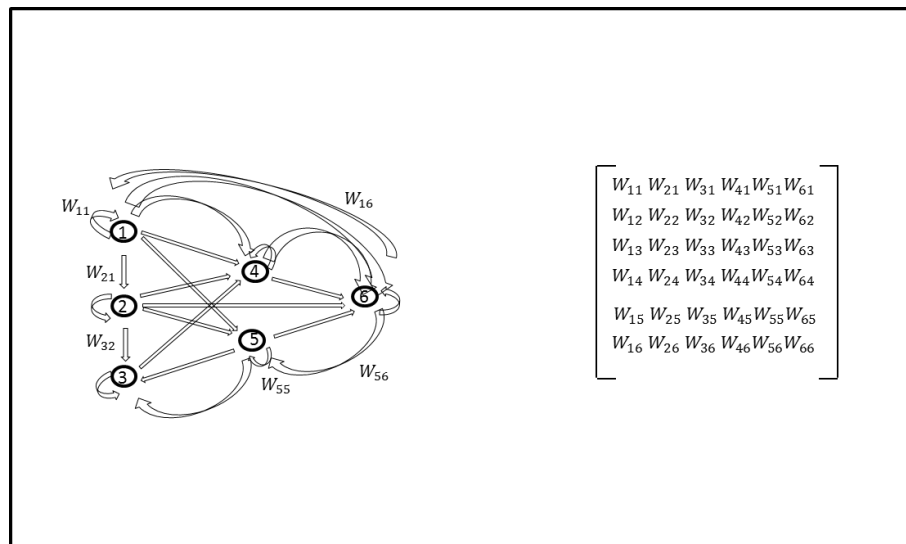


FIGURA 4-2: RNA TOTALMENTE CONECTADA Y SU MATRIZ DE PESOS

4.3. TOPOLOGÍA DE LAS REDES NEURONALES

De acuerdo al patrón de conexiones es factible clasificar las redes neuronales artificiales. Las redes necesitan varios EPs para patrones de clasificación complejos pero estos EPs significan muchos pesos que requerirán cantidades importantes de datos para generalizar el problema. La forma que se ha encontrado para superar este tema es a través de técnicas de reducción de datos o topologías especiales (Análisis de Componentes Principales o Mapas de Auto-Organización). Esto reduce el número de entradas para la red

que es normalmente la capa más grande. Otro método es usar topologías alternativas que requieren menos pesos por EP (Ej.: Funciones de Base Radial) o usar diseños modulares de MLP que reducen el número de pesos por EP porque su topología no está completamente conectada. Estos métodos alternativos dan origen a la clasificación topológica aquí propuesta:

4.3.1. Perceptrón

El modelo biológico más simple de un perceptrón es una neurona y vice versa. Es decir, el modelo matemático más simple de una neurona es un perceptrón, La neurona es una célula especializada y caracterizada por poseer una cantidad indefinida de canales de entrada llamados dendritas y un canal de salida llamado axón. Las dendritas operan como sensores que recogen información de la región donde se hallan y la derivan hacia el cuerpo de la neurona que reacciona mediante una sinapsis que envía una respuesta hacia el cerebro, esto en el caso de los seres vivos.

El perceptrón fue probablemente el primer neurocomputador exitoso (Rosenblatt 1957). Rosenblatt construyó el MARK I para la clasificación de imágenes. Esta topología no es más que una red neuronal feed-forward sin capas ocultas. Sus capacidades de procesamiento de información son limitadas. La regla de aprendizaje del perceptrón es sencilla: se presenta un patrón, si la salida es la deseada, no hace nada. Si la respuesta es incorrecta, se cambian los pesos de las unidades de procesamiento activas hacia la respuesta deseada. Se repite el proceso hasta que todas las unidades tengan salidas aceptables.

Existen dos tipos de aprendizaje, el primero utiliza una tasa de aprendizaje mientras que el segundo no la utiliza. Esta tasa de aprendizaje amortigua el cambio de los valores de los pesos.

El primer tipo de aprendizaje utiliza la siguiente regla de actualización de los pesos:

$$w(j)' = w(j) + \alpha(\delta - y)x(j)$$

El segundo tipo de aprendizaje, sin tasa de aprendizaje, utiliza como regla de actualización de los pesos la siguiente:

$$w(j)' = w(j) + (\delta)x(j)$$

Donde,

x_j denota el elemento en la posición j en el vector de la entrada

w_j el elemento en la posición j en el vector de peso

y denota la salida de la neurona

δ denota la salida esperada

Por lo cual, el aprendizaje es modelado como la actualización del vector de peso después de cada iteración, lo cual sólo tendrá lugar si la salida y difiere de la salida deseada δ .

Novikoff (1962) probó que el algoritmo de aprendizaje converge después de un número finito de iteraciones si los datos son separables linealmente y el número de errores es limitado. Sin embargo si los datos no son separables linealmente, la línea de algoritmo anterior no se garantiza que converja.

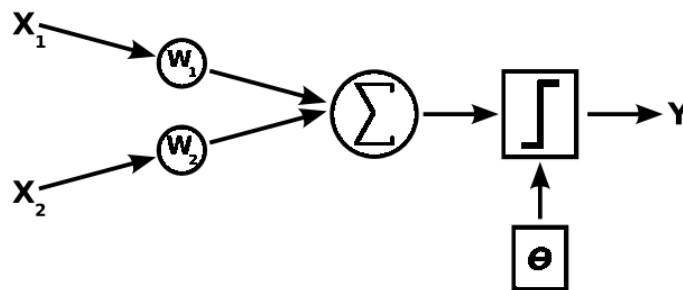


FIGURA 4-3: PERCEPTRON

4.3.2. Perceptrón Multicapa (MLP)

El perceptrón multicapa (MLP) es una de las topologías de redes más ampliamente usadas. El artículo de Lippman (Lippmann 1987) es tal vez una de las mejores referencias para captar las capacidades computacionales de los MLPs. En general, para clasificación de patrones estáticos, los MLP con dos capas ocultas son muy adecuados. En otras palabras, las funciones discriminantes pueden tomar cualquier forma, según el requerimiento de los clústeres de datos de entrada. Es más, cuando los pesos están apropiadamente normalizados y las clases de salida son normalizadas a 0/1, los MLP logran el mismo desempeño del llamado receptor máximo a posteriori, que es un óptimo desde un punto de vista de clasificación (Makhoul 1991). En términos de capacidades de mapeo, los MLP se cree son capaces de aproximar funciones arbitrarias. Esto ha sido de importancia en el estudio dinámico no lineal (Lapedes and Farber 1987) y otros problemas de mapeo.

Los MLP son entrenados comúnmente con el algoritmo de retroalimentación (Backpropagation algorithm) (Rumelhart, Hinton et al. 1986). De hecho el interés en las redes neuronales fue impulsado por la aparición de este algoritmo. El algoritmo de aprendizaje de mínimos cuadrados (Least mean squares, LMS) propuesto por Widrow no puede extrapolarse a los EPs de la capa oculta, ya que no se conocen los valores deseados de la señal. En el caso del algoritmo de retroalimentación, el error se propaga a través de la red y permite la adaptación de los EPs ocultos.

Dos de las características más importantes de los MLP son: sus elementos de procesamiento no lineales tienen una linealidad que debe ser suavizada (la función logística y la tangente hiperbólica son los más usados); y su interconectividad masiva (ej.: cualquier elemento de una capa dada, puede alimentar la siguiente capa). Entre sus limitaciones se

puede decir que el perceptrón multicapa no extrapola bien, es decir, si la red se entrena mal o de manera insuficiente, las salidas pueden ser imprecisas.

El perceptrón multicapa se entrena con la corrección del error de aprendizaje, lo cual significa que la respuesta deseada para el sistema debe ser conocida.

El aprendizaje por corrección de error funciona de la siguiente manera: de la respuesta del sistema en el elemento de procesamiento i en la iteración n , $y_i(n)$, y la respuesta deseada $d_i(n)$ para un patrón de entrada dado, un error instantáneo e_i se define por:

$$e_i(n) = d_i(n) - y_i(n)$$

Usando la teoría del aprendizaje de gradiente descendente, cada peso en la red puede ser adaptado corrigiendo el valor presente con un término que es proporcional a la entrada actual y al error en el peso. Ej.:

$$w_{ij}(n + 1) = w_{ij}(n) + \eta \delta_i(n) x_j(n)$$

El error local $\delta_i(n)$ se puede computar directamente del $e_i(n)$ en el EP de salida o puede ser calculado como una suma ponderada de errores en los EPs internos. La constante η se llama “tamaño de paso”. Este procedimiento es el que se conoce como Algoritmo de retroalimentación.

La retroalimentación computa la sensibilidad de costo funcional con respecto a cada peso en la red y actualiza cada peso proporcionalmente a su sensibilidad. Lo interesante de este procedimiento es que puede ser implementado con información local y requiere pocas multiplicaciones por peso, lo cual es muy eficiente. Como este es un procedimiento de gradiente descendente, solo utiliza información local por tanto puede caer en mínimos

locales. Es más, el procedimiento es inherentemente ruidoso ya que se está usando un estimado pobre del gradiente causando una lenta convergencia.

4.3.3. Redes de Función de Base Radial (RBF):

Las redes RBF son comúnmente utilizadas para propósitos de modelación de funciones no lineales e inciertas. Utilizar las RBF con propósitos de modelación puede ser entendido como un problema de aproximación en un espacio altamente dimensional (Zemouri, Gouriveau et al. 2010).

Una red RBF consiste de una capa de entrada, una oculta y una de salida con una función de activación de las unidades ocultas llamada “Función de Base Radial”. Las salidas de sus dos capas forman una combinación lineal que producen una respuesta significativa no cero a la entrada solo cuando cae dentro de una pequeña región localizada del espacio de entrada (Li and Shi 2010). Uno de los tipos de funciones radiales más comunes es la llamada distribución de campana Gaussiana. Cada función radial tiene dos parámetros: un centro y una amplitud. La amplitud de la función radial determina la dispersión de la función y que tan rápido decrece la activación del nodo oculto con la distancia del centro a la entrada (Bishop 1995). Las neuronas de la capa de salida son una combinación lineal ponderada de la RBF en la capa oculta. Una red RBF puede ser modelada por las siguientes ecuaciones:

$$y_j(x) = \sum_{i=1}^n w_{ij} \psi_i(x) + b_j$$

Donde, $y_j(x)$ es la salida en el j -ésimo nodo en la capa de salida, n es el número de nodos ocultos, w_{ji} es el peso del i -ésimo nodo oculto al j -ésimo nodo de salida., $\psi_i(x)$ es la función de activación básica de la capa oculta y b_j es el parámetro bias del j -ésimo nodo.

Muchos de los tipos comunes de RBF son funciones lineales, radial cubica de Duchon, cuadrática radial más funciones de activación cubicas o gaussianas. La última función tiene la forma:

$$\psi_i(x) = \exp\left(\frac{-\|X - u_i\|^2}{2\sigma_i^2}\right)$$

Donde X es el vector de entrada, u_i es el vector centro del i -ésimo nodo oculto y σ_i es la amplitud de la función básica. Existen dos tipos distintivos de arquitecturas RBF Gaussianas. El primero utiliza la función de activación exponencial, por lo que la activación de la unidad es una perturbación gaussiana como función de las entradas. El segundo tipo utiliza la función de activación softmax por lo que las activaciones de todas las unidades de entrada son normalizadas para sumar uno. Este tipo de redes comúnmente se llaman “RBF normalizadas” o red NRBF (Aminian 2010).

Las redes RBF tienen un fundamento matemático fuerte soportado en la teoría de regularización para resolver problemas mal condicionados (problemas altamente sensibles a cambios en los datos finales). Una red de este tipo se compone por una serie de entradas p que se conectan en un nodo oculto. Cada nodo en la capa oculta es una función p gaussiana multivariable de la forma:

$$G(x, x_i) = \exp\left[\frac{-1}{2\sigma_i^2} \sum_{k=1}^p (x_k - x_{ik})^2\right]$$

De media x_i (cada dato) y varianza s_i . Estas funciones se llaman funciones base radial. Finalmente se linealiza la salida ponderada de los nodos ocultos para obtener

$$F(x) = \sum_{j=1}^N w_j [G(x, x_j)]$$

El problema con esa función es que puede llevar a obtener una capa oculta muy grande (el número de muestras del conjunto de entrenamiento)

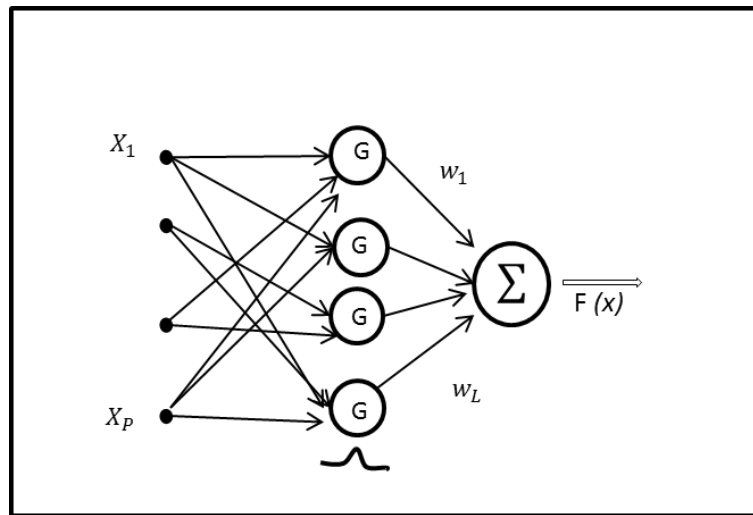


FIGURA 4-4: RED DE FUNCIÓN DE BASE RADIAL (RBF)

4.3.4. Redes Análisis de Componentes Principales (ACP)

El problema principal en reconocimiento de patrones es definir características importantes de los datos para la clasificación (extracción de características). Es deseable transformar las muestras de entrada en un nuevo espacio (espacio de características) donde la información acerca de las muestras aun permanezca pero se reduzca la dimensionalidad. El Análisis de Componentes Principales (ACP) también llamado Transformada de Descomposición de Valor Singular de Karhunen-Loeve, es una técnica de este tipo. El ACP encuentra un grupo de direcciones ortogonales en el espacio de entrada y proporciona una forma de encontrar esas proyecciones en estas direcciones en forma ordenada. El primer componente principal es el que tiene la proyección más larga (pensando en la proyección como una sombra del clúster de datos en cada dirección). Las direcciones

ortogonales se llaman vectores propios de la matriz de correlación del vector de entrada, y las proyecciones los correspondientes valores propios.

Ya que el ACP ordena las proyecciones, se puede reducir la dimensionalidad truncando las proyecciones a un determinado orden. El error de reconstrucción es igual a la suma de proyecciones que se dejan fuera. Las características en el espacio de proyección se convierten en valores propios. Nótese que este espacio de proyección es lineal.

El ACP se hace normalmente analíticamente resolviendo un problema de valores propios de la función de correlación de entrada. De todas formas, Sanger (Sanger 1989) y Oja (Oja 1982) demostraron que el ACP puede ser alcanzado por una red neuronal lineal de una capa entrenada con la regla de aprendizaje de Hebbian modificada.

Nótese en la Figura 4-5 que la red tiene p entradas (se asume que la muestra tiene p componentes) y $m < p$ EPs lineales de salida. La salida está dada por:

$$y_j(n) = \sum_{i=0}^{p-1} w_{ij}(n)x_i(n) \quad j = 0, 1, \dots, m - 1$$

Para entrenar los pesos, se usa la siguiente regla de Hebbian modificada:

$$\Delta w_{ji} = \eta \left[y_i(n)x_i(n) - y_i(n) \sum_{k=0}^j w_{ki}(n)y_k(n) \right] \quad i = 0, 1, \dots, p - 1; \quad j = 0, 1, \dots, m - 1$$

Donde η es el tamaño de paso.

Lo interesante de esta red es que se calculan los vectores propios de la función de correlación sin nunca calcular la función de correlación. Sanger mostró que este procedimiento de aprendizaje converge a la solución correcta (ej.: los pesos de la red ACP se acercan a los primeros m componentes principales de la matriz de datos de entrada. Las

salidas están por lo tanto relacionadas a los valores propios y pueden ser usados como entradas en otra red neuronal para clasificación.

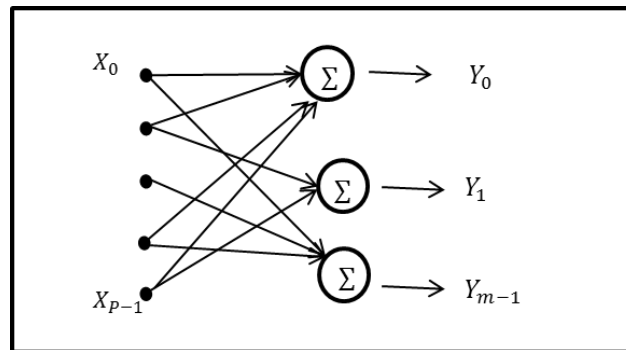


FIGURA 4-5: RED ANALISIS DE COMPONENTE PRINCIPALES (ACP)

Las redes ACP pueden ser utilizadas para compresión de información entregando las mejores m características lineales. También pueden ser usadas para reducción de información en conjunto con clasificadores MLP. En ese caso, la separabilidad de las clases no está siempre garantizada pues si los clústeres de datos están suficientemente separados es factible pero si las clases están una encima de otra, el ACP hará las proyecciones pero, en algunos casos, estará traslapada con otras proyecciones. Otro problema con las redes ACP es el punto de datos extremos. Estos datos extremos distorsionaran la estimación de los vectores propios y crearan proyecciones de datos sesgadas. Las redes no lineales son mejores para manejar estos casos.

La importancia del análisis de componentes principales es que el número de entradas para un clasificador MLP puede ser reducido en gran medida, lo cual impacta positivamente en el número de patrones de entrenamiento y los tiempos de entrenamiento.

4.3.5. Redes Dinámicas

Las redes dinámicas son una clase importante de topología neuronal, capaces de procesar señales que varían en el tiempo. Pueden ser vistas como una extensión no lineal

de filtros lineales adaptativos, o una extensión de las redes neuronales estáticas con entradas de tiempo variables. Como tal ellas llenan un nicho importante en las aplicaciones de las redes neuronales.

Una red neuronal dinámica es una red neuronal estática con una extensión de un mecanismo de memoria, que es capaz de almacenar valores pasados de la señal de entrada. En muchas aplicaciones (identificación de sistemas, clasificación de patrones temporales, predicción no lineal) la memoria es importante para permitir decisiones basadas en el comportamiento de la entrada en un periodo de tiempo. Un clasificador estático toma decisiones basadas solo en la información presente, por lo tanto puede que no desarrolle funciones que involucren conocimiento acerca de la historia de las señales de entrada.

4.3.5.1. Red Neuronal con Retraso en el Tiempo (TDNN)

En redes neuronales, la estructura de memoria más común son los filtros lineales. En la red neuronal de retraso en el tiempo (TDNN por su nombre en inglés) la memoria es una línea de retraso superpuesta ej.: un conjunto de posiciones de memoria que almacenan el pasado de la entrada (Waibel, Hanazawa et al. 1989). Las conexiones auto-recurrentes (alimentación de la salida de un EP con la entrada) también han sido utilizadas como memoria y esas unidades son llamadas “unidades contexto” (Elman 1990, Jordan 1997).

La memoria gamma (véase Figura 4-6, $g_i(t)$ son entradas para los siguientes EP) es una estructura que cae en cascada de conexiones auto-recurrentes (de Vries and Principe 1990). Por lo tanto es una estructura con retroalimentación local, que extiende la unidad de contexto con almacenamiento más versátil y acepta la línea de retraso como un caso especial ($\mu=1$).

μ es un parámetro adaptativo que controla la profundidad de la memoria. Esta estructura tiene una profundidad de memoria K/μ , donde K es el número de retrasos en la cascada. Su resolución es μ (de Vries and Principe 1990). Ya que la topología es recurrente, una forma de aprendizaje temporal debe ser usada para adaptar el parámetro gamma μ (ej.: aprendizaje recurrente en tiempo real o retroalimentación a través del tiempo). La ventaja de esta estructura en las redes dinámicas es que ellas pueden, con un número definido de retrasos, proveer una memoria controlable. Y debido a que la red adapta el parámetro gamma para minimizar el error cuadrático medio, se alcanza la mejor profundidad/resolución buscada.

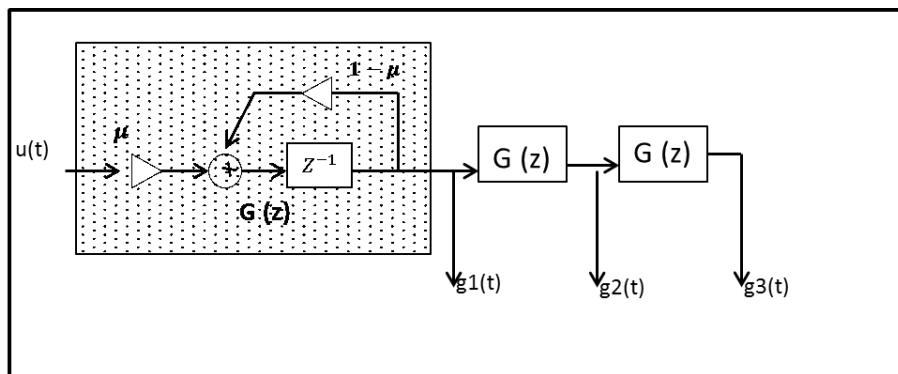


FIGURA 4-6: MEMORIA GAMMA

La memoria gamma puede ser implementada a las entradas (memoria gamma focalizada), a los EPs ocultos o a las EPs de salida. En cada caso esta almacenará las activaciones de los respectivos EPs y usará sus valores pasados para calcular la salida neta. Una red neuronal dinámica con memoria gamma se llama Modelo Neuronal Gamma. Este modelo se puede aplicar a clasificación de patrones que varían con el tiempo, detectores de señal, predicción de series de tiempo caóticas e identificación de sistemas no lineales (Waibel, Hanazawa et al. 1989).

2.4.5.2. Redes Recurrentes con Retardo Temporal (TLRN)

La TLRNs que confían una estructura de memoria en su capa de entrada pueden ser entendida como preprocesadores de las entradas. Pero ahora el problema es la representación de la información en vez de la información entre los patrones de entrada, como sucede en las redes ACP. Cuando se tiene una señal en el tiempo (tal como una serie de tiempo financiera o una señal de monitoreo de una proceso industrial) no se sabe a priori, en el tiempo, cual es la información relevante. El procesamiento de la señal puede ser utilizado en estos casos en sentido general y puede ser sustituido para predicción, identificación de la dinámica o clasificación.

La red TLRN más estudiada es el modelo gamma. El modelo gamma se caracteriza por la estructura de memoria que es una cascada de integradores con fuga, ej.: una extensión de las unidades de contexto de las redes de Jordan y Elman (Ver Figura 4-7)

Las señales en los retrasos de la memoria gama pueden ser representados por:

$$x_0(n) = u(n)$$

$$x_k(n) = (1 - \mu)x_k(n - 1) + \mu x_{k-1}(n - 1) \quad k = 1, \dots, K$$

Pensando en las entradas como corrientes y los pesos como potenciómetros que el usuario puede controlar, los EPs se pueden entender como simples transistores y la salida como voltaje. Nótese que la señal en el retraso k es una versión suavizada de la entrada, que retiene el voltaje de un evento pasado, creando una memoria. Adicional el punto en el tiempo donde la respuesta tiene un pico es aproximadamente k/m donde m es un parámetro de retroalimentación. Esto significa que la red neuronal puede controlar la profundidad de la memoria cambiando el valor del parámetro de retroalimentación en vez de cambiar el número de entradas. El parámetro m puede ser adaptado usando un procedimiento de gradiente descendiente como los otros parámetros de la red, pero debido a que este es un

parámetro recursivo, una regla de aprendizaje más poderosa requiere ser aplicada (ej.: retropropagación a través del tiempo (BPTT)).

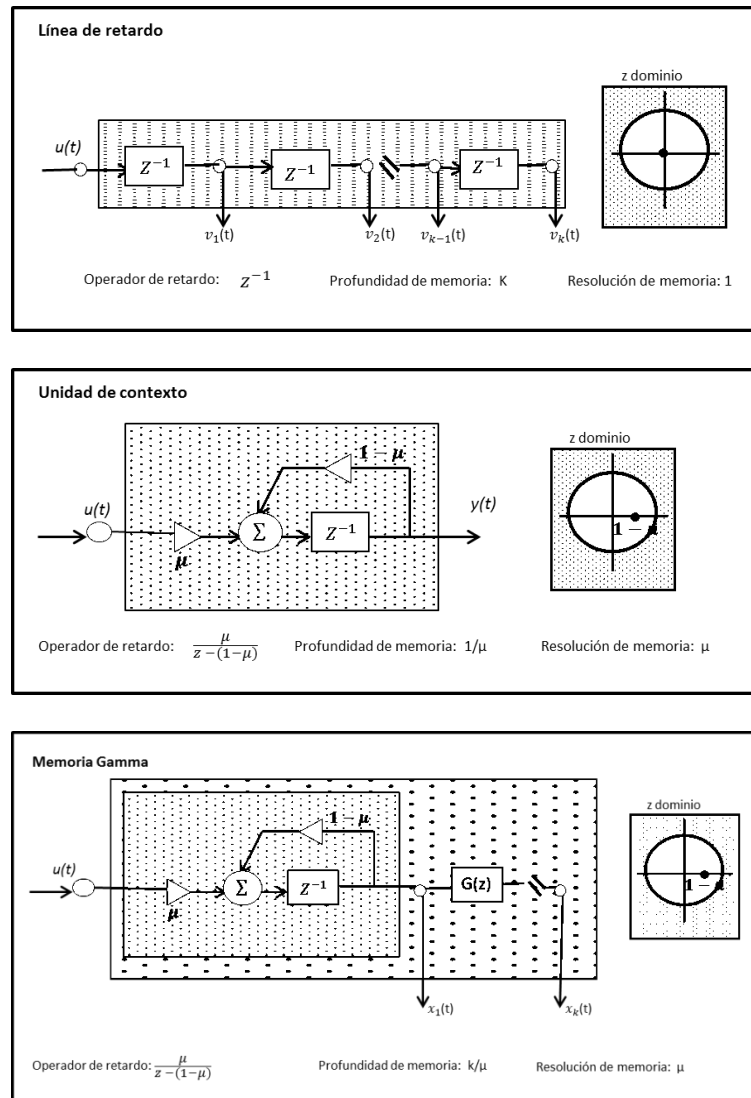


FIGURA 4-7: ESTRUCTURAS DE MEMORIA CONEXIONISTA Y LA LOCALIZACIÓN DEL DOMINIO DE FRECUENCIA DEL POLO

Las memorias se pueden añadir a cualquier capa de la red, produciendo topologías neuronales muy sofisticadas que son muy útiles para la predicción de series de tiempo y sistemas de identificación y reconocimiento de patrones temporales (ver Figura 4-8).

En vez de usar la memoria gamma, existen otras estructuras de memoria recientemente aplicadas con algunas ventajas. Una de estas es la memoria Laguarre, basada en la función de Laguarre. Las funciones de Laguarre son una serie de funciones ortogonales que se construyen a partir de filtros lowpass seguidas por una cascada de funciones de allpass (Oliveira e Silva 1995). Esta familia de funciones constituye un tramo ortogonal del espacio gama, por lo que tienen las mismas propiedades que las memorias gama, pero pueden mostrar una convergencia más rápida para algunos problemas. La ecuación de las funciones de Laguarre es:

$$L_i(z, \mu) = \sqrt{1 - (1 - \mu)^2} \frac{z^{-1} - (1 - \mu)^{j-1}}{(1 - (1 - \mu)z^{-1})^i} \quad i = 1, 2, \dots$$

Nótese que esto da una ecuación recursión de la forma

$$x_0(n) = (1 - \mu)x_0(n - 1) + \sqrt{1 - (1 - \mu)^2}u(n)$$

$$x_k(n) = (1 - \mu)x_k(n - 1) + x_{k-1}(n - 1) - (1 - \mu)x_{k-1}(n)$$

Donde $u(n)$ es la señal de entrada.

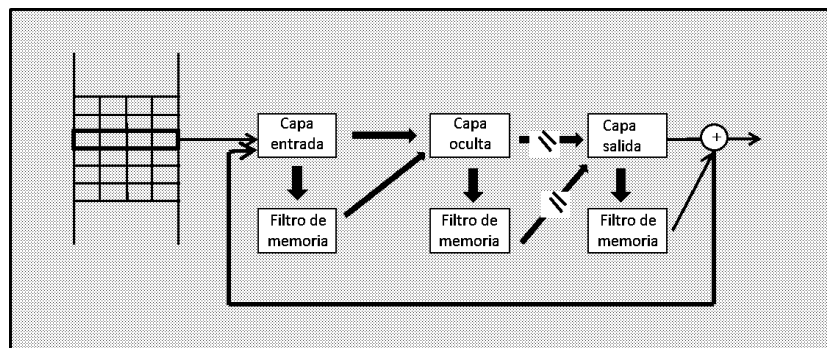


FIGURA 4-8: USO DE GAMMA KERNELS EN UNA ARQUITECTURA MLP

Recientemente autores como Taghi han encontrado mejores desempeños de las redes TLRN con estructuras de memoria gamma para el pronóstico de datos temporales al ser comparados con otras estructuras (Taghi Sattari, Yurekli et al. 2012).

4.4. PARADIGMAS DE APRENDIZAJE

El proceso de modificación de los parámetros de la red para mejorar el desempeño comúnmente se llama aprendizaje. Tomando como ejemplo la figura del Perceptrón Multicapa con 6 EP (ver Figura 4-3), la meta es obtener el valor de 1 voltio en la salida, cuando diferentes corrientes son presentadas a la red. Lo que un usuario haría es conectar una entrada y verificar el valor a la salida. Si el valor no es 1 voltio, entonces algunos de los potenciómetros requerirán ser cambiados hasta que la meta sea lograda. Luego la segunda corriente es presentada y se repite el proceso hasta obtener la respuesta deseada para todas las entradas. Cuando la red neuronal se ha entrenado, este proceso de cambiar los pesos es automático.

Aprender requiere diferentes consideraciones. Primero, al cambiar los parámetros de la red, el desempeño debe mejorar. Por tanto, se requiere una medida de desempeño. Segundo, las reglas de cambio de parámetros deben ser especificadas. Tercero, el procedimiento de entrenamiento debe hacerse con información conocida.

La aplicación de procedimientos de medición del desempeño produce otra división taxonómica importante. Cuando la función de desempeño está basada en la definición de una medida de error, se dice que es un aprendizaje supervisado. Normalmente el error está definido como la diferencia de la salida de la RNA y una señal externa deseada predefinida. En aplicaciones de ingeniería donde el desempeño deseado es conocido, los paradigmas del aprendizaje supervisado se vuelven importantes.

Otros métodos de aprendizaje modifican los pesos de la red de acuerdo a alguna regla interna preestablecida de iteración (aprendizaje no supervisado). Por tanto no hay un entrenador. Esta es la razón por la cual el aprendizaje no supervisado se llama auto-organización (self-organization). Este tipo de aprendizaje puede ser apropiado para el

descubrimiento de características (extracción de características) en señales complejas con redundancia. Un tercer caso intermedio de aprendizaje es el llamado Aprendizaje por Refuerzo (reinforcement learning). Es este caso un entrenador externo indica la calidad (buena o mala) de la respuesta. Este tipo de aprendizaje aún se encuentra en la fase de investigación, pero puede tener la respuesta al aprendizaje on-line.

Para el aprendizaje supervisado se requieren hacer tres tipos básicos de decisiones: selección del criterio de error, cómo se propagará el error a través de la red y que restricciones (estáticas o a través del tiempo) se le imponen a la salida de la red. El primer tema está relacionado con la fórmula (función de costo) que calculará el error. El segundo aspecto está asociado con el mecanismo que modificará los parámetros de la red en una forma automatizada. El tercero hace referencia a como se restringe la salida de la red frente a la señal deseada. Se puede especificar sólo el comportamiento en el momento final (aprendizaje de punto fijo), es decir, no restringir los valores que toma la salida para llegar a la conducta deseada. O bien, se pueden restringir los valores intermedios y lo que se llama aprendizaje de trayectoria. Nótese que una red feedforward, debido a que es un asignador instantáneo (la respuesta se obtiene en un solo paso) solo puede ser entrenada por un aprendizaje de punto fijo. En cambio, las redes recurrentes pueden ser entrenados especificando el comportamiento en el tiempo final (punto fijo de aprendizaje) o el comportamiento a lo largo de una trayectoria (la trayectoria de aprendizaje).

El aprendizaje requiere de la especificación de un grupo de datos de entrenamiento. El desempeño del entrenamiento debe ser chequeado contra un grupo de datos llamado “Grupo de Prueba”. Es de fundamental importancia la selección del tamaño del conjunto de datos de entrenamiento y que este provea un cubrimiento representativo de todas las posibles condiciones. Durante el aprendizaje, la red va a descubrir la mejor asignación

entre los datos de entrada y los valores deseados. Si los datos usados en el entrenamiento no son representativos de la clase de datos de entrada, se puede esperar un desempeño bajo en el Grupo de Prueba, aunque el desempeño de la red haya sido excelente para el grupo de entrenamiento.

4.5. PARÁMETROS DEL APRENDIZAJE

El control de los parámetros del aprendizaje es un problema en investigación de Redes Neuronales. El punto es que se desea entrenar una red de la manera más rápida posible obteniendo el mejor desempeño. Aumentar el parámetro de tasa de aprendizaje, disminuirá el tiempo de entrenamiento pero también incrementará la posibilidad de divergencia y de oscilación alrededor del valor óptimo. Debido a que la corrección del error depende del desempeño de las características de superficie y de la tasa de aprendizaje, para obtener un aprendizaje constante, un parámetro de aprendizaje adaptativo es necesario. La modificación de las tasas de aprendizaje son posibles bajo ciertas condiciones, pero un número adicional de parámetros experimentales también deberían ser incluidos. Por tanto el procedimiento tiende a ser frágil y las ganancias son dependientes de los problemas (Fahlman 1988, Silva and Almeida 1990, Lecun, Simard et al. 1992).

El procedimiento convencional es simplemente seleccionar una tasa de aprendizaje y un término de momentum. El término de momentum imparte un “factor de memoria” en la adaptación y ha mostrado que acelera la adaptación mientras evita el quedar atrapado en mínimos locales hasta cierto punto. Entonces, la ecuación de aprendizaje se convierte en:

$$w_{ij}(k) = \gamma \Delta w_{ij}(k-1) - \mu \nabla_{w_{ij}} E(k)$$

Donde γ es una constante (normalmente ajustada entre 0.5 y 0.9) y μ es la misma tasa de aprendizaje.

También el Delta barra Delta de Jacob es un procedimiento versátil, pero requiere mayor cuidado en la especificación de los parámetros de aprendizaje. La idea es que cuando hay iteraciones consecutivas que producen el mismo signo de actualización del peso, la tasa de aprendizaje es muy pequeña. Por otra parte, si las iteraciones consecutivas producen actualizaciones de los pesos que tienen signos opuestos, la tasa de aprendizaje es muy rápida. Jacob propuso las siguientes fórmulas para la actualización de la tasa:

$$k \quad \text{si } S_{ij}(n-1)D_{ij}(n) > 0$$

$$\Delta n_{ij}(n+1) = -\beta \eta_{ij}(n+1) \quad \text{si } S_{ij}(n-1)D_{ij}(n) < 0$$

$$0 \quad \text{en otro caso}$$

Donde, n es la tasa de aprendizaje para cada peso, $D_{ij}(n)$ es el gradiente y

$$S_{ij}(n) = (1 - \xi)D_{ij}(n-1) + \xi S_{ij}(n-1)$$

Donde ξ es una pequeña constante.

La actualización se puede hacer al finalizar la presentación de todos los elementos del Grupo de Entrenamiento (Aprendizaje batch) o en cada iteración (tiempo real). La primera modalidad “suaviza” el gradiente y puede conllevar a aprendizajes más rápidos para datos ruidosos, sin embargo, también puede promediar el gradiente a cero y el aprendizaje se estanca. En la mayor parte de los casos es más preferida la opción de modificación de los pesos en cada iteración con pequeñas tasas de aprendizaje.

El aprendizaje llamado Momentum es una mejora al gradiente descendente en el sentido en que el término de memoria (el incremento pasado en el peso) se utiliza para acelerar y estabilizar la convergencia. En el aprendizaje con momentum la ecuación de actualización del peso se convierte en:

$$w_{ij}(n+1) = w_{ij}(n) + \eta \delta_i(n)x_j(n) + \alpha(w_{ij}(n) - w_{ij}(n-1))$$

Donde α es el momentum.

El entrenamiento se puede implementar en dos formas: se presenta un patrón y se adaptan los pesos (On-line training), o se presentan todos los patrones del archivo de entrada (una época), se acumula la actualización de los pesos y luego se actualizan todos los pesos con el promedio de la actualización. Esto se llama aprendizaje batch. Son teóricamente equivalentes pero el primero algunas veces tiene ventajas en problemas difíciles (Ej.: muchas parejas similares entrada – salida).

Para iniciar la retroalimentación se requiere un valor inicial para cada peso (normalmente un valor aleatorio pequeño), continuando el procedimiento hasta que se alcance algún tipo de “criterio de parada”. Los tres criterios más comunes son: a) llegar a un determinado número de iteraciones, b) sobrepasar el error cuadrático medio de la salida o c) usar una validación cruzada. La validación cruzada es la más poderosa de los tres ya que termina el entrenamiento en el punto donde se logra la mejor generalización (ej.: el desempeño en el grupo de prueba). Para implementar la validación cruzada se debe separar una pequeña porción de los datos de entrenamiento y usarlos para ver cómo se comporta la red entrenada (ej.: cada 100 épocas de entrenamiento, probar la red con los datos de validación). Cuando el desempeño comienza a degradarse en el grupo de validación, el entrenamiento debe suspenderse.

Revisar el progreso del entrenamiento en cualquier proceso iterativo de entrenamiento es fundamental. La curva de aprendizaje (cómo evoluciona el error cuadrático medio con la iteración de entrenamiento) es una forma. Se puede determinar la dificultad de esta tarea y cómo controlar los parámetros de aprendizaje a partir de esta curva. Cuando se obtiene una curva plana, el tamaño del paso debe incrementarse para

acelerar el aprendizaje. Por otra parte, cuando la curva oscila arriba y abajo el paso debe decrecerse. En el extremo, el error puede tender al incremento constante, mostrando que el aprendizaje es inestable. En ese punto la red debe ser reiniciada. Cuando la curva de aprendizaje se estabiliza tras varias iteraciones en un nivel de error no aceptable, se sugiere pensar en otro tipo de topología (más capas ocultas y/o una topología diferente) u otro proceso de entrenamiento (alguna técnica de gradiente más sofisticada).

La existencia de mínimos locales en la función de error dificulta considerablemente el entrenamiento, pues una vez alcanzado un mínimo el entrenamiento se detiene aunque no se haya alcanzado la tasa de convergencia fijada.

Cuando caemos en un mínimo local sin satisfacer el porcentaje de error permitido se puede considerar: cambiar la topología de la red (número de capas y número de neuronas), comenzar el entrenamiento con unos pesos iniciales diferentes, modificar los parámetros de aprendizaje, modificar el conjunto de entrenamiento o presentar los patrones en otro orden.

Otro problema es cuando parar el entrenamiento. Los criterios de parada están basados en el monitoreo del error cuadrático medio (ECM). La curva del ECM en función del tiempo se llama “Curva de Aprendizaje”. El criterio más utilizado es probablemente seleccionar un número de iteraciones, pero también se puede predefinir un error final. Estos dos métodos tienen inconvenientes. Un acuerdo que se hace en la práctica es poner un valor umbral de incremento mínimo en el aprendizaje. Cuando, entre dos iteraciones consecutivas, el error no cae por lo menos en una determinada cantidad, el entrenamiento para. Este criterio puede ser usado para comparar las diferentes topologías.

Pero aún existe otra posibilidad y es monitorear el ECM en el Grupo de Prueba, como en la validación cruzada. En este caso se debe parar el aprendizaje cuando el error en

el Grupo de Prueba comienza a incrementar (ver Figura 4-9). En ese punto se da la generalización máxima.

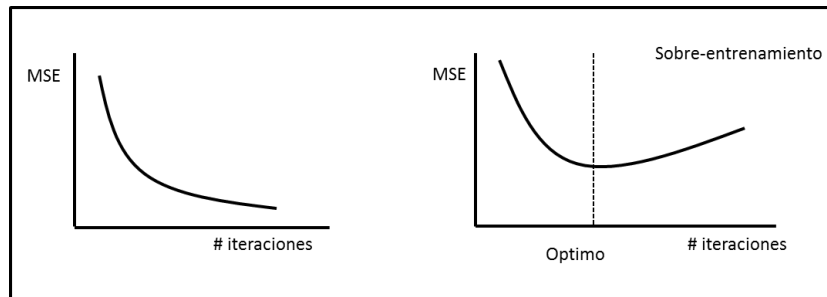


FIGURA 4-9: COMPORTAMIENTO DEL MSE PARA LOS CONJUNTOS DE ENTRENAMIENTO Y PRUEBA

Para poder implementar este procedimiento se debe entrenar la red hasta un cierto número de iteraciones, congelar los pesos y evaluar el desempeño con el Grupo de Prueba. Después volver al Grupo de Entrenamiento y continuar el aprendizaje. Es un poco más incómodo implementar este criterio ya que, para un segmento de iteraciones de entrenamiento, es requerido hacer un cálculo de desempeño adicional sobre el Grupo de Prueba.

5. APLICACIÓN DE LAS RNA AL PRONÓSTICO DE TEMPERATURAS

5.1. SELECCIÓN ESTACIONES METEOROLÓGICAS IDEAM

Actualmente el IDEAM cuenta con una red de cerca de 3.000 estaciones dispersas por todo el país divididas, en dos grandes tipos: principales y ordinarias. Las estaciones principales son aquellas en la cual se hacen observaciones de lluvia, temperatura del aire, temperaturas máxima y mínima a dos metros, humedad, viento, radiación, brillo solar, evaporación temperaturas extremas del tanque de evaporación, cantidad de nubes y fenómenos especiales (ej.: reventón cálido, corrientes en chorro o máximos de viento). En

las ordinarias se hacen observaciones de temperatura del aire y precipitación, primordialmente. Poseen muy poco instrumental registrador. Algunas llevan instrumentos adicionales tales como tanque de evaporación, heliógrafo y anemómetro (Sabogal, N., 1985).

La temperatura del aire en superficie medida por el IDEAM es la temperatura del aire libre a una altura comprendida entre 1,25 y 2,00 metros sobre el nivel del suelo. En sus estaciones meteorológicas suelen instalarse termómetros de máxima y mínima y el sicrómetro, formado por los termómetros seco y húmedo (López, L. 1968). La temperatura del aire es la que indica el termómetro seco del sicrómetro con graduaciones cada 0.2°C. Las lecturas se efectúan a la 07:00, 13:00 y 19:00 HLC (López, L. 1988) y se registran en el “Diario de Observaciones” que incluye el trazo continuo del termógrafo de donde se obtienen la temperatura mínima (T_{\min}) y temperatura máxima (T_{\max}).

La variable de estudio e interés para el desarrollo del modelo de reducción de escala estadístico basado en redes neuronales artificiales es la temperatura media diaria (T), obtenida como la media entre la T_{\max} y T_{\min} diaria. Esta variable se encuentra disponible tanto en estaciones principales como ordinarias por lo que no se hizo necesaria hacer una identificación de ellas.

Desafortunadamente para el periodo comprendido entre el 1 de enero de 1961 y el 31 de diciembre de 1977 no hay información disponible de temperatura en el IDEAM. Adicionalmente en el periodo del 1 de enero de 1978 hasta el 31 de diciembre de 2003 se encontró gran discontinuidad de datos de temperatura media (T) en las diferentes estaciones. Los datos perdidos en ocasiones alcanzaron hasta un 52.48%, es decir, más de la mitad de la información de temperatura media diaria en el periodo comprendido entre el 01/01/1978 y 31/12/2003 no estuvo disponible. Por tanto la continuidad de la información

(datos de T diaria sin periodos de omisión) y el número de datos faltantes se convirtieron en los primeros dos criterios de selección de las estaciones meteorológicas usadas en el proyecto.

El tercer criterio para la selección de las estaciones meteorológicas de estudio fue su ubicación según la regiones climatológicas de Köppen-Geiger (Kottek, Grieser et al. 2006). El propósito de este criterio es evidenciar si debido a la complejidad orográfica de Colombia la resolución actual del Reanálisis es suficiente para la descripción y pronóstico climático del país.

Basados en los tres criterios anteriormente mencionados se seleccionaron las estaciones: Sta Leticia (E1), Isla del Santuario (E2), Margarita Las Hda (E3), Manaure (E4), Apto Rafael Barvo (E5). Los información de datos faltantes, ubicación y zona climática además de otros se encuentran descritos en el Anexo 1.

La estación meteorológica Santa Leticia se encuentra ubicada en la zona sur del país, en el municipio de Puracé (Lat. 02°14'N Long. 76°09'W) sobre la cordillera central a 2.085 m.s.n.m. en un piso térmico medio correspondiente a un clima frío y a clima de páramo. La temperatura media de la zona es de 16°C, con precipitaciones anuales entre 1.600 y 2.500 mm.; regularmente las épocas de verano se presentan en los meses de junio, julio, agosto y parte de diciembre y enero; el resto del año es invierno intermitente.

La mayor parte del territorio es montañoso y su relieve corresponde a la cordillera central de los Andes, en los que cuenta entre otros, con los accidentes orográficos de la Sierra Nevada de los Coconucos (4.235 m.), ubicada en los límites con el departamento del Huila, los volcanes de Pan de Azúcar (4.800 m.), Puracé (4.445 m) y Sotará, el pico de Paletará, el Páramo Blanco y los cerros Agua Blanca, Chagartón (4.000 m.), Canelo, Cargachiquillo, Cuaré, Los Coconucos, Peñas Blancas y Puzná.

Según la clasificación climática de Köppen-Geiger se encuentra ubicada en la zona Af caracterizada por lluvias repartidas a lo largo del año por lo que no hay una estación seca. Este clima se identifica por ser un clima de alta vegetación verde, además de ser uno de los más húmedos, cuenta con distintas plantas y especies vegetales. En este clima existen lluvias abundantes por la evaporación del agua contenida en las plantas. La flora se compone de distintos niveles de altura, es decir, existen árboles que superan los 30m; árboles de altura media y niveles bajos de arbustos y pastos.

La estación meteorológica Isla del Santuario se encuentra ubicada en la zona centro del país (Lat. 05°28'N Long. 73°44'W) sobre la cordillera Oriental a 2.580 m.s.n.m. ubicada en el municipio de Fuquene al Norte de Cundinamarca. El municipio de Fúquene tiene dos grandes áreas, una de ellas es la planicie fluvial y lacustre y la otra son las formas montañosas del entorno, existen laderas dependientes, suaves, aptas para el desarrollo agropecuario; por otro lado existen sectores escarpados donde se han evidenciado derrumbes y caída de rocas. La vegetación natural del municipio de Fúquene pertenece a dos ecosistemas: andino y páramo. El páramo se localiza a más de 3.300 m de altitud y cubre aproximadamente el 20% del territorio del municipio este ecosistema presenta un tipo de vegetación abierta dominada principalmente por pastos, frailejones y pequeños arbustos de chite. El piso andino se localiza hacia la parte plana y parte de la zona montañosa del municipio y ocupa el 80% del territorio está comprendido en la franja superior a los 2.600 m y corresponde a las zonas del bosque húmedo montañoso. Existen amplias zonas reforestadas con árboles exóticos como acacias, eucaliptos y pinos. Temperatura media 13°C.

La zona climática está considerada como Cfc según la clasificación de Köppen-Geiger donde los meses con temperatura media superior a 10 °C son menos de cuatro al

año. Es un clima frío, con una oscilación anual muy escasa, temperaturas siempre por encima de $-3\text{ }^{\circ}\text{C}$ y lluvias abundantes.

La estación meteorológica Hacienda Las Margaritas se encuentra ubicada en la zona oriental del país (Lat. $04^{\circ}20'\text{N}$ Long. $72^{\circ}09'\text{W}$) a 150 m.s.n.m. en el municipio de Puerto Gaitán que topográficamente comprende una extensa zona de estructura plana conocida como las Sabanas de la Altillanura, cuyo eje articulador principal es el Río Meta. Su clima es cálido, la temperatura promedio es de $27\text{ }^{\circ}\text{C}$ y puede alcanzar hasta $40\text{ }^{\circ}\text{C}$ en el periodo de diciembre a marzo, que provoca que la zona se vaya secando y aparecen bosques de rebalse que en invierno se encuentran inundados hasta 6 m o más. Pisos Térmicos: Clima cálido (9 meses de invierno y 3 meses de verano).

Por su ubicación en la zona tropical, este municipio está bajo la influencia del desplazamiento de la Zona de Confluencia Intertropical. Entre diciembre y marzo se presenta el período más seco, debido a que los vientos alisios del noreste son los dominantes en esta época del año y desplazan hacia el sur la ZCIT. El período de lluvias se extiende de marzo a noviembre, debido a que en esta época los vientos alisios del sureste empiezan a ser los dominantes, desplazando la ZCIT hacia el norte. El prolongado período de lluvias se debe al doble paso de la ZCIT por la alternancia de los vientos alisios dominantes. La vegetación de la llanura está conformada por pastos y pajonales con abundantes arbustos y árboles de baja altura. En las riberas de los ríos se encuentran los bosques de galería de gran variedad florística

La zona climática está considerada como Am, monzónico, según la clasificación de Köppen-Geiger es decir, con algún mes por debajo de 60 mm y si la precipitación del mes más seco es superior a la fórmula $[100 - (\text{Precipitación anual}/25)]$. Cálido todo el año, con una estación seca corta seguida por una húmeda con fuertes lluvias.

La estación meteorológica Manaure se encuentra ubicada en el extremo norte de Colombia (Lat. 11°46'N Long. 72°28'W) a 1 m.s.n.m. en el municipio de su mismo nombre donde el clima es caliente, refrescado por las brisas marinas del Noreste, permanece todo el año con temperaturas que oscilan entre los 28 grados centígrados y los 38 grados centígrados. Debido a las altas temperaturas y a la intensidad de los vientos, la evaporación es muy alta, los vientos del noreste ocasionan la época de mayor sequía, las lluvias son muy escasas y solo se presentan en los meses de Mayo, Junio, Octubre y Noviembre. Las precipitaciones son influidas por la dirección y velocidad de los vientos y son cortas y fuertes, llegando a caer en algunos lugares hasta 150 milímetros en un solo aguacero. En general es una llanura plana pedregosa y suavemente ondulada, en la que predominan las dunas del litoral, las lagunas marinas, playones arcillosos y salinas entre el Pájaro y Mayapo, sometidas a inundaciones prolongadas; entre Manaure y Ahuyama la costa se recorta en acantilados bajos las capas arcillosas que conforman la planicie de la Media y Alta Guajira.

La zona climática está considerada como BSh, Estepario cálido, según la clasificación de Köppen-Geiger, donde la vegetación es escasa debido a las condiciones meteorológicas. Este clima es conocido en algunas regiones como mediterráneo seco pues es, en muchas ocasiones, un clima de transición entre el Csa (mediterráneo) y el BW (desértico). La temperatura media anual está por encima de los 18°C.

La estación meteorológica Apto Rafael Barvo se encuentra ubicada en la zona norte de Colombia (Lat. 9°20'N Long. 75°16'W) en la franja de la Costa Atlántica a 166 m.s.n.m. en el municipio de Corozal, departamento de Sucre, Sub-Región Sabanas. El clima del Departamento de Sucre es cálido, con temperaturas medias entre 27°C y 30°C, suavizado por los vientos alisios del noreste y las brisas marítimas; la humedad relativa es del orden

del 85% y las lluvias están distribuidas durante el año en dos períodos, alternados con épocas secas; las precipitaciones aumentan de este a oeste; así, en la faja costera pueden ser inferiores a 1.000 mm y en la región del bajo San Jorge y bajo Cauca, sobrepasar los 3.000 mm. Sus tierras están comprendidas en el piso térmico cálido.

La zona climática está considerada como Aw, Sabana (invierno seco), según la clasificación de Köppen-Geiger, Cálido todo el año, con estación seca. Es el clima propio de la sabana. Caliente subhúmedo con lluvias en verano. Se da entre la zona ecuatorial y en los desiertos cálidos entre los 10 y 25 ° de latitud norte y sur. Tiene una estación invernal seca que aumenta a medida que se aleja del ecuador. La precipitación mínima es superior a 100 mm.

5.2. PRE-PROCESAMIENTO DE DATOS

5.2.1. Reconstrucción datos

Evidentemente la falta de datos de temperatura media diaria y principalmente en periodos continuos prolongados, es una fuerte limitación a la construcción del modelo propuesto. Para eliminar las discontinuidades se propuso una metodología basada en la estacionalidad anual de los datos y el desplazamiento de la tendencia de la serie de tiempo obtenida.

Básicamente se construyó la serie de tiempo anual de la temperatura media en cada fecha (día) en el periodo de tiempo de estudio (ver Tabla 15). Es decir, por ejemplo se toman todos los 1 de enero del año 1978 al 2003 y se obtiene una serie de tiempo. Si esta serie de tiempo contiene datos faltantes, se construye una serie de tiempo basada en la media móvil con intervalo 3. La serie de datos reconstruida es aquella en la que se combinan los datos de la serie original y los datos de la serie tendencial. Solo los datos faltantes en la serie original fueron remplazados con los datos de la serie tendencial.

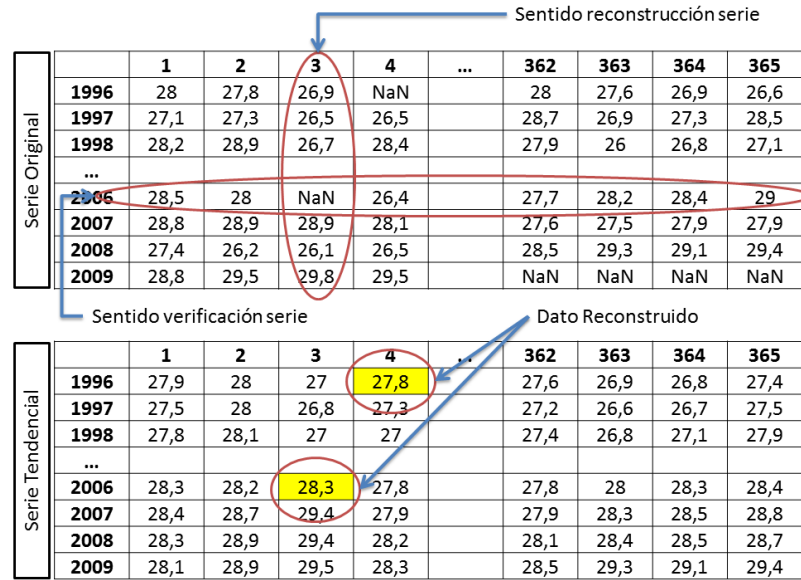


FIGURA 5-1: RECONSTRUCCIÓN SERIES DE TIEMPO

Los datos obtenidos en la serie tendencial y que entraron a ser parte de la serie reconstruida fueron comparados con la tendencia de los datos de ese año (ver Figura 5-1). Es decir, el dato reconstruido se verifica frente a la tendencia de temperatura media diaria del año correspondiente (temperaturas del periodo 1 de enero a 31 de diciembre de ese año). Se encontró que los cambios de temperatura diaria se encuentran dentro de la variación normal de la serie.

Una vez verificados los datos que entraron a la serie reconstruida se validó mediante la prueba de Karikal-Wallis si estadísticamente se puede afirmar que la serie de tiempo original y la serie de tiempo reconstruida vienen de las misma población o tienen propiedades idénticas y se aplicó la prueba de Kolmogorov-Smirnov para determinar si la reconstrucción de la serie altera de manera significativa la distribución original. En ninguno de los casos (estaciones) hubo alteración significativa de la distribución original y

estadísticamente mantuvieron propiedades idénticas. Por lo tanto se aceptan las series reconstruidas como representativas para los modelos. (Ver Anexo 2)

5.2.2. Separación anomalías

Los modelos desarrollados en este proyecto se refieren únicamente a las anomalías de temperatura, no la temperatura absoluta. El término anomalía de temperatura se refiere a la desviación atípica de un valor de referencia o promedio a largo plazo. Una anomalía positiva indica que la temperatura observada estaba más caliente que el valor de referencia, mientras que una anomalía negativa indica que la temperatura observada fue más fría que el valor de referencia. Las anomalías de temperatura se calcularon en relación con el período base de estudio para cada estación (ver Tabla 15). La razón para trabajar con las anomalías, en lugar de la temperatura absoluta, es que la temperatura absoluta varía considerablemente en distancias cortas (debido a la orografía), mientras que las anomalías térmicas mensuales o anuales son representativas de una región mucho más grande. De hecho, se ha demostrado (Hansen and Lebedeff 1987) que las anomalías de temperatura están estrechamente correlacionados a distancias del orden de 1000 km.

Un modelo clásico para una serie de tiempo, supone que una serie $x(1), \dots, x(n)$ puede ser expresada como suma o producto de tres componentes: tendencia, estacionalidad y un término de error aleatorio. Existen tres modelos de series de tiempo, que generalmente se aceptan como buenas aproximaciones a las verdaderas relaciones, entre los componentes de los datos observados. Estos son:

$$\text{Aditivo } X(t) = \tau(t) + E(t) + A(t)$$

$$\text{Multiplicativo: } X(t) = \tau(t) \cdot E(t) \cdot A(t)$$

$$\text{Mixto: } X(t) = \tau(t) \cdot E(t) + A(t)$$

Donde,

$X(t)$ serie observada en instante t

$\tau(t)$ componente de tendencia

$E(t)$ componente estacional

$A(t)$ componente aleatoria (accidental)

Una suposición usual es que $A(t)$ sea una componente aleatoria o ruido blanco con media cero y varianza constante.

Para la selección del modelo se sabe que el aditivo, es adecuado, por ejemplo, cuando $E(t)$ no depende de otras componentes, como $\tau(t)$, sí por el contrario la estacionalidad varía con la tendencia, el modelo más adecuado es un modelo multiplicativo. En el caso de la temperatura media diaria en región ecuatorial, donde no hay estaciones, se supone una estacionalidad anual debido al ciclo de rotación del planeta. Y basados en las observaciones de incremento de la temperatura global (Doherty, Bojinski et al. 2009, Hartell and Pearman 2010) y cambios extremos en la temperatura (Friedel 2012) se acepta para el proyecto, como más adecuado, un modelo mixto de la forma:

$$T(t) = \tau(t)E(t) + A(t)$$

Donde,

$T(t)$ es la temperatura media en el instante (día) t

$\tau(t)$ componente de tendencia

$E(t)$ componente estacional

$A(t)$ anomalía

Basados en este modelo, se descompusieron las series de tiempo de cada estación para el periodo de estudio en sus tres componentes.

Para la componente tendencial $\tau(t)$ se realizó una regresión lineal de las temperaturas medias observadas a través del tiempo y se derivó una ecuación lineal de la forma:

$$\tau(t) = \beta t_i + C$$

Donde,

$\tau(t)$: componente tendencial en el periodo t

t: periodo t, día de observación.

β, C : coeficiente de correlación lineal.

El índice estacional, que es el factor en el cual se desplaza la componente tendencial de la serie, se derivó del promedio de desplazamientos diarios año a año, como el cociente de la observación y el promedio móvil calculado para esa fecha:

$$E(t) = \frac{\sum_{i=1}^n \frac{T_i(t)}{MAV_i(t)}}{n}$$

Donde,

$E(t)$: Índice estacional en el periodo (t)

$T_i(t)$: Temperatura observada en el periodo (t) en el año i

$MAV_i(t)$: Promedio móvil en el periodo (t) en el año i

n: número de años del periodo de estudio.

Las anomalías se calculan como la diferencia de la temperatura media observada en el periodo t y su respectiva componente tendencial desplazada por $E(t)$ en el mismo período:

$$A(t) = T(t) - \tau(t)E(t)$$

Las anomalías $A(t)$ resultantes (Ver Anexo 3) son las que de aquí en adelante se utilizarán para la construcción de los modelos.

5.2.3. Análisis exploratorio de información

Las metodologías clásicas de inferencia estadística dependen fuertemente de la suposición que los datos están libres de valores atípicos (outliers), que están aproximadamente normalmente distribuidos y que están claramente no-correlacionados si son obtenidos a intervalos de serie regulares. Si la información no es cercana a una distribución normal o si no está libre de valores atípicos, los resultados de los métodos clásicos de inferencia estadística pueden ser engañosos (Khan, Coulibaly et al. 2006).

Por tanto, para este proyecto se realizó un análisis gráfico exploratorio de los datos de anomalías de la temperatura media diaria en cada una de las 5 estaciones trabajadas, para determinar tres puntos:

1. Si la distribución es cercana a una normal.
2. Si la información contiene outliers
3. Como la información se recolecta en el tiempo, si hay evidencia de correlaciones seriales (correlaciones entre datos sucesivos de la información)

Una buena idea de la forma de la distribución generada por los datos y la presencia de valores atípicos se puede deducir de la observación de cuatro gráficas: gráfica de densidad, histograma, gráficas de barras (boxplot) y gráficas cuantil-cuantil (qq-plot) normal. Las gráficas de densidad son, esencialmente, gráficas suavizadas de un histograma. Las boxplot se usan para localización y variación de información como identificación del medio (50%) de la información, la mediana y los puntos extremos. Una gráfica qq-plot normal consiste de una gráfica de los datos ordenados versus los cuantiles correspondientes de una distribución normal (distribución normal con media cero y

varianza uno). Si la gráfica qq-plot es aproximadamente lineal, los datos son razonablemente Gaussianos (Khan, Coulibaly et al. 2006). De entre estas cuatro gráficas, el histograma y la gráfica de densidad dan la información sobre la forma de la distribución mientras que las boxplot y qq-plot informan sobre los valores atípicos. El chequeo de la correlación de datos en la serie de tiempo se hace a través de gráficas de la función de autocorrelación (ACF). Así mismo de aquí se pueden deducir características de tendencia o estacionalidad (ciclos).

Los datos de anomalías fueron estandarizados (varianza 1) y normalizados (media 0) al 31 de diciembre de 2003. Los datos del reanálisis NCEP/NCAR como los del CGCM 3.1 fueron entregados normalizados al 31 de diciembre de 1990. De las gráficas de análisis exploratorio (Ver Anexo 4) se observa que las estaciones E1 y E2 son cercanas a una distribución normal mientras que las estaciones E3, E4 y E5 presentan colas. En las 5 estaciones meteorológicas hay presencia de outliers. En cuanto autocorrelación se observó que la correlación de datos desaparece rápidamente y se mantiene en el desplazamiento analizado de 360 días.

5.2.4. Organización de la información

El modelo de reducción de escala se construyó basado en dos grupos de datos: variables de entrada o predictores y variables de salida o predicandos. Los predictores corresponden a las entradas del modelo de reducción de escala que son a su vez las salidas comunes al modelo global canadiense CGCM3.1 y al reanálisis NCEP/NCAR. La lista de variables de los escenarios A1B y A2, disponible para el CGCM 3.1 y el reanálisis NCEP/NCAR se encuentra detallada en la Tabla 1. Estas 25 variables meteorológicas corresponden a: velocidad del viento, vorticidad, dirección del viento, divergencia,

componentes U y V, geopotencial, humedad específica a diferentes alturas y temperatura a 2m.

TABLA 1: PREDICTORES

No.	Predictores	No.	Predictores
1	Presión en el nivel medio del mar (Pa)	14	500hPa Divergencia
2	1000hPa Velocidad del viento	15	850hPa Velocidad del viento
3	1000hPa componente U	16	850hPa componente U
4	1000hPa componente V	17	850hPa componente V
5	1000hPa Vorticidad	18	850hPa Vorticidad
6	1000hPa Dirección del viento	19	850hPa Geopotencial
7	1000hPa Divergencia	20	850hPa Velocidad del viento
8	500hPa Velocidad del viento	21	850hPa Divergencia
9	500hPa componente U	22	500hPa Humedad específica
10	500hPa componente V	23	850hPa Humedad específica
11	500hPa Vorticidad	24	1000hPa Humedad específica
12	500hPa Geopotencial	25	Temperatura a 2m
13	500hPa Dirección del viento		

5.2.4.1. Variables de entrada o Predictores

Las variables de entrada utilizadas en la construcción de los diferentes modelos evaluados corresponden a las 25 variables del Reanálisis de la celda en la cual se encuentra ubicada la estación meteorológica del IDEAM. Pero adicionalmente se usaron como variables de entrada para la construcción de más modelos, las 25 variables de cada una de las 8 celdas circundantes a la celda de ubicación de la estación (Ver Figura 5-2).

Debido a la dependencia en el parámetro Coriolis para el cálculos de corrientes de viento, las celdas donde $Y=24$ (cerca al Ecuador donde las fuerzas Coriolis tienden a cero) contiene solo un grupo limitado de predictores. Cuando el aire se acerca al Ecuador, los vientos Alisios del noreste y sureste convergen sobre una estrecha zona a lo largo de él, denominada Zona de Confluencia Intertropical; en ella la inversión se debilita y el aire se

eleva, el desarrollo vertical de las nubes aumenta y la inestabilidad se extiende a mayores altitudes. Las precipitaciones se hacen más fuertes y más frecuentes (Hidrográficas 2012).

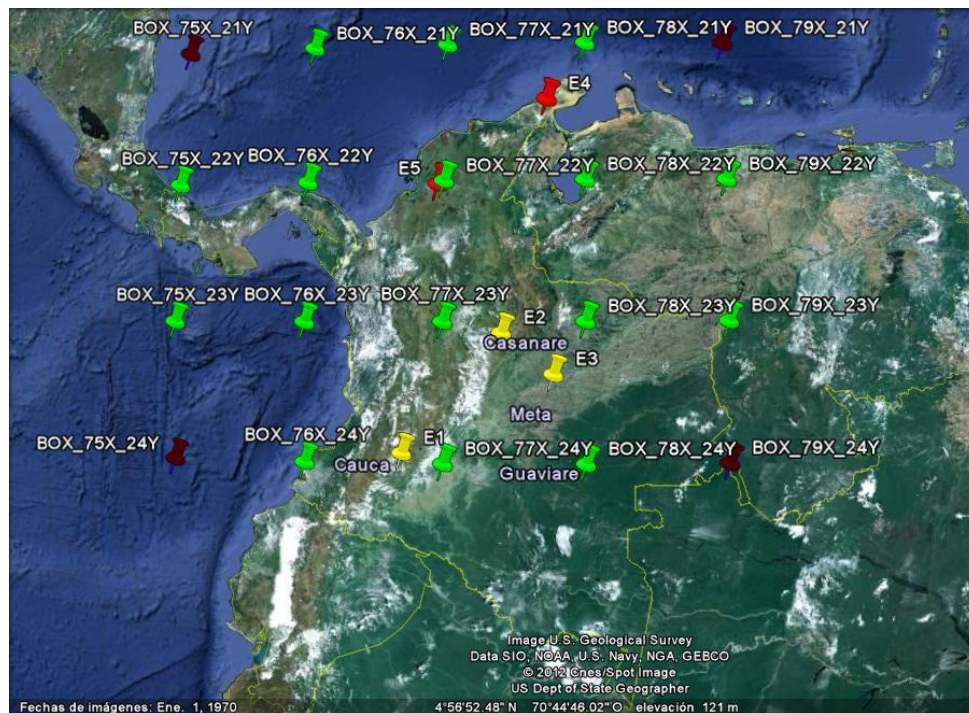


FIGURA 5-2: UBICACIÓN CELDAS Y ESTACIONES DE ESTUDIO

Posteriormente se utilizaron dos metodologías para la condensación de los predictores: Un Análisis de Componente Principal (ACP) y un método alternativo llamado regresión Stepwise.

El objetivo del Análisis de Componentes Principales fue identificar, a partir de un conjunto de p variables, otro conjunto de k ($k < p$) variables no directamente observables, denominadas factores, tal que: k sea un número pequeño, se pierda la menor cantidad posible de información y la solución obtenida sea interpretable. Se aceptó para todos los modelos el número de componentes principales que redujera el número de variables explicando más del 85% de la varianza de los datos. Las gráficas del ACP se presentan en el Anexo 5.

La técnica estadística Stepwise para el análisis de regresión relaciona un predicando como una combinación lineal de predictores del modelo $f_j(x)$ ($j = 1, \dots, p$) en cada una de las observaciones $(x_1, y_1), \dots, (x_n, y_n)$. En este trabajo se utilizó un método sistemático para agregar y quitar términos a los modelos multilineales en función de su significancia estadística del 95% en la regresión, probando los modelos con y sin un término potencial (Ver Tabla 17).

5.2.4.2. Variables de salida o Predicandos

El siguiente grupo de datos, predicandos, es la variable de salida del modelo de reducción de escala que se denominó anomalías $A(t)$, derivadas de los datos reportados por el IDEAM para la temperatura media diaria en cada una de las 5 estaciones meteorológicas. Estas anomalías se encuentran graficadas en el Anexo 3.

5.2.4.3. Entrenamiento y Validación del Modelo

Usando la información histórica del reanálisis disponible para el periodo 1961-2000, tres tipos diferentes de datos fueron seleccionados: datos de entrenamiento, datos de validación y datos de prueba. Los datos de prueba (DP) corresponden a un 10% del total de datos del periodo de evaluación. Los datos en entrenamiento (DE) corresponden al 67,5% de los datos y los datos de validación (DV) al 22,5% que queda disponible. El modelo se validó en el DV usando una validación cruzada.

5.3. SELECCIÓN DE ARQUITECTURAS

Según la revisión bibliográfica que se hizo para este trabajo (en ScienceDirect, ISI Web of Science y Scopus), son escasos los trabajos específicos en reducción de escala de variables climáticas en Colombia, donde su orografía y ubicación global es particular y compleja para el pronóstico de variables climáticas (Ej: efecto coriolis). En este sentido se propuso para el trabajo utilizar las arquitecturas estudiadas por otros autores, para

homologarlas bajo las condiciones colombianas. Según este criterio se presentan a continuación las arquitecturas que se probaron.

5.3.1. Taxonomías y topologías evaluadas

Las arquitecturas disponibles en NeuroSolutios 6 (Principe, Euliano et al. 2000) son:

Por taxonomía: Perceptrones multicapa y Redes Recurrentes. Para esta versión no están disponibles las redes dinámicas TLRN y TDNN aunque se corrieron en una versión de evaluación para ver su comportamiento.

Por topologías: Perceptron Multicapa (MLP), Red Recurrente (RN), Generalized FeedForwrd (GFF), Función de Base Radial (RBF) y Red Neuronal Probabilística (PNN).

La aplicación de redes neuronales para la reducción de escala en Perceptrones Multicapa es reportada por (Miksovsky and Raidl 2005, Coulibaly and Evora 2007, Krasnopolsky 2007, Chattopadhyay, Jhajharia et al. 2011, De, De et al. 2011, Do Hoai, Udo et al. 2011, Chronopoulos, Kamoutsis et al. 2012). En Redes Recurrentes en predicción de nevadas (Sauter, Weitzenkamp et al. 2010), niveles de agua (Coulibaly 2010) y niveles de ozono (Solaiman, Coulibaly et al. 2008). En redes Generalized FeedForward para predicción de lluvias (Hung, Babel et al. 2009) y modelación de humedad relativa (Białobrzewski 2008). Se encuentran también trabajos comparativos para la predicción de temperatura donde las Redes RBF tienen buen desempeño (Miksovsky and Raidl 2005, Ustaoglu, Cigizoglu et al. 2008, Aksornsingchai and Srinilta 2011). No se reporta el uso frecuente o que demuestre una mejora al usar una red neuronal probabilística para la reducción de escala de variables meteorológicas. Tal vez el trabajo más relacionado es el modelo QRNN desarrollado por Cannon (Cannon 2011). Una de las mayores limitaciones de este tipo de arquitectura es la inviabilidad de generar reportes de escenarios probabilísticos.

Basado en estos dos criterios se decidió utilizar las topologías MLP, RBF, RN y GFF.

Los tipos de aprendizaje evaluados en cuanto a algoritmo fueron LevenbergMarquardt (L) y Momentum (M) con las ventajas y características de cada uno descritas en el capítulo 2. En cuanto a la presentación de datos para el aprendizaje se utilizó presentación Batch (B) y en tiempo Real u On-line (O).

La identificación de cada uno de los modelos (nombre del modelo) está compuesta por 4 partes: AAA-B-C-D, donde AAA se refiere a la sigla de la topología, B al número de capas ocultas, C al tipo de presentación de datos y D al algoritmo de aprendizaje. En el Anexo 7 aparece la lista de los 100 modelos evaluados para cada red, resultantes de la combinación del número de variables de entrada (9 celdas, 1 celda, PCA o Stepwise), topología (MLP, RBF, RN, GFF, TDNN, TLRN), presentación de datos de aprendizaje (B ó O) y algoritmo de aprendizaje (L ó M).

5.3.2. Desarrollo del experimento

La metodología seguida para la construcción del modelo de red neuronal artificial constó de dos pasos: selección mejor arquitectura y optimización de la arquitectura. Al final se guardó la mejor arquitectura en combinación con los EP de la capa oculta optimizada y los mejores pesos resultantes del entrenamiento y prueba.

La construcción de los modelos de redes neuronales para la selección de la arquitectura fue desarrollado usando el entorno de NeuroSolutions(Principe, Euliano et al. 2000). El conjunto de variables predictoras se describe en la Tabla 1. Las variables de entrada para los diferentes modelos son las 25 variables derivadas del Reanálisis NCEP/NCAR correspondiente a la celda de ubicación de la estación meteorológica, 225 variables correspondientes a las 25 variables de cada una de las 9 celdas circundantes y

propia de la estación meteorológica, las variables derivadas de la regresión Stepwise y los componentes principales derivados de la reducción de un 50% de las variables de entrada a través de la aplicación del análisis de componentes principales aplicado a los anteriores conjuntos de 25 y 225 variables. Debido a que las celdas de la zona 24Y del reanálisis no cuentan con algunas de las 25 variables, por el efecto coriolis según se explicó anteriormente, el número exacto de variables por conjunto de datos del modelo se describirá en el Capítulo 5.3.3. Las variables de salida son las anomalías $Y_{N(t)}$, normalizadas a cero y estandarizadas a uno, derivadas del análisis de la serie temporal reconstruida.

El principal propósito del entrenamiento de las redes es lograr generalizar las características de la serie de tiempo. Si una red neuronal aprendió correctamente las características de los datos, entonces se entiende que el modelo alcanzó una buena generalización. Pero, dependiendo de la complejidad de la red, un modelo puede caer, bien sea en un sobre-estimación o sub-estimación. Una red muy compleja puede no detectar todas las características del conjunto de datos, ocasionando una sub-estimación o puede ajustarse hasta captar el ruido en la información, no solo las características, llevando a un sobre-estimación. Para evitar esto se utilizó la validación cruzada, dividiendo los datos así: 10% prueba, 22,5% datos para validación y 67,5% datos para el entrenamiento de la red. La validación cruzada se corrió sobre el mismo grupo de datos tanto en las regresiones lineales como en los modelos de redes neuronales para garantizar la comparación entre los dos métodos de ajuste. Este grupo de validación se utilizó para ir computando el error al mismo tiempo que se entrena la red previniendo problemas de “memorización” de patrones en el grupo de entrenamiento y adicionalmente funciona como criterio de parada del entrenamiento cuando en el grupo de validación inicia a crecer el error.

Los Elementos de Procesamiento en capa de entrada están constituidos por el número de variables de entrada del conjunto correspondiente. Se utilizó una capa o dos capas (para el caso de MLP-2-O-M, MLP-2-B-L y MLP-2-B-M) ocultas. Tanto en la capa oculta como en la capa de salida se utilizaron funciones de activación Tangente Hiperbólica. La capa de salida tiene un solo nodo que representa la Anomalía $Y(t)$ y para el aprendizaje se usaron dos algoritmos de retroalimentación del error (EBP, Error Backpropagation) modificados, LevenbergMarquardt (L) y Momentum (M). La función de costo utilizada fue el Error Cuadrático (L2Criterion). Una descripción más detallada de las arquitecturas construidas y evaluadas aparece en el Anexo 7. En la Tabla 2 se muestran las características de las arquitecturas.

Cada una de las 5 estaciones meteorológicas fue modelada por separado. Primero se entrenaron cada una de las topologías, con los 5 conjuntos de variables de entrada descritos. Se usaron semillas aleatorias para la iniciación de los pesos.

La selección de la arquitectura se hace basada en tres parámetros: el coeficiente de correlación (r), el error medio absoluto (MAE) y el error cuadrático medio (MSE). Estos tres parámetros fueron evaluados en el conjunto de entrenamiento, validación cruzada y prueba, siendo el de validación el de mayor interés para seleccionar el modelo.

Para correr los modelos se entregaron al software Neuro Solutions 6 for Excel (Express Builder) archivos construidos con los conjunto de datos de entrenamiento, validación y prueba. Se corrieron entrenamientos en computador Intel Core i5 a 2.30GHz, uno a la vez, en cada una de las arquitecturas descritas en el capítulo anterior, hasta alcanzar los criterios de parada descritos en el Capítulo 4.5. En ese punto se corrió una prueba para ver el comportamiento de la arquitectura entrenada en datos nunca antes

presentados a la red neuronal. Esta información fue colectada como idea del comportamiento del modelo para la predicción de datos.

TABLA 2: CARACTERISTICAS DE LAS ARQUITECTURAS

Topología	EP	Función Activación
MLP-1-B-L	CONJUNTO - 10 - 1	Tanh - Tanh - L2Criterion
MLP-1-B-M	CONJUNTO - 10 - 1	Tanh - Tanh - L2Criterion
MLP-1-O-M	CONJUNTO - 10 - 1	Tanh - Tanh - L2Criterion
MLP-2-B-L	CONJUNTO - 12 - 6 - 1	Tanh - Tanh - Tanh - L2Criterion
MLP-2-B-M	CONJUNTO - 12 - 6 - 1	Tanh - Tanh - Tanh - L2Criterion
MLP-2-O-M	CONJUNTO - 12 - 6 - 1	Tanh - Tanh - Tanh - L2Criterion
GFF-1-B-L	CONJUNTO - 10 - 1	Tanh - Tanh - L2Criterion
GFF-1-B-M	CONJUNTO - 4 - 1	Tanh - Tanh - L2Criterion
GFF-1-O-M	CONJUNTO - 10 - 1	Tanh - Tanh - L2Criterion
RBF-1-B-L	CONJUNTO - 10 - 1	Tanh - Tanh - L2Criterion
RBF-1-B-M	CONJUNTO - 20 - 1	Tanh - Tanh - L2Criterion
RBF-1-O-M	CONJUNTO - 20 - 1	Tanh - Tanh - L2Criterion
RN-1-B-L	CONJUNTO - 4 - 1	Tanh - Tanh - L2Criterion
RN-1-B-M	CONJUNTO - 10 - 1	Tanh - Tanh - L2Criterion
RN-1-O-M	CONJUNTO - 10 - 1	Tanh - Tanh - L2Criterion
TDNN-1-B-L	(CONJUNTO x 5) - 4 - 1	Tanh - Tanh - L2Criterion
TDNN-1-B-M	(CONJUNTO x 5) - 10 - 1	Tanh - Tanh - L2Criterion
TDNN-1-O-M	(CONJUNTO x 5) - 10 - 1	Tanh - Tanh - L2Criterion
TLRN-1-B-L	(CONJUNTO x 5) - 4 - 1	Tanh - Tanh - L2Criterion
TLRN-1-B-M	(CONJUNTO x 5) - 10 - 1	Tanh - Tanh - L2Criterion
TLRN-1-O-M	(CONJUNTO x 5) - 10 - 1	Tanh - Tanh - L2Criterion

5.3.3. Resultados mejores arquitecturas

5.3.3.1. Estación E1: Santa Leticia (Código IDEAM 21055030)

La estación meteorológica Santa Leticia se encuentra ubicada dentro de la celda 77X24Y del reanálisis NCEP/NCAR, de resolución 3,71° Lat. x 3,75 Long. Esta celda y todas las circundantes de la franja 24Y cuentan con información solo de 13 variables de entrada: presión en el nivel medio del mar (Pa), 1000hPa componente U, 1000hPa componente V, 500hPa componente U, 500hPa componente V, 500hPa geopotencial, 850 hPa componente U, 850 componente V, 850 geopotencial, 500hPa humedad específica, 850

humedad específica, 1000hPa humedad específica y temperatura a 2m. Las variables de entrada correspondientes a velocidad del viento, dirección del viento, vorticidad y divergencia a 1000, 500 y 850hPa no se reconocen en el CGCM y por ende reanálisis.

Por tanto los 5 conjuntos de variables de entrada evaluados para la estación E1 (Santa Leticia) corresponden a:

Conjunto 1: 1 variable de entrada Temperatura a 2m celda 77X24Y

Conjunto 2: 13 variables de la celda 77X24Y (Ver Anexo 6)

Conjunto 3: 9 variables entrada resultantes del análisis Stepwise efectuado sobre las 25 variables de la celda 77X24Y

Conjunto 4: 114 variables resultantes de 75 variables circundantes de la franja 23Y (76X23Y, 77X23Y, 78X23Y) y 39 de la franja 24Y (76X24Y, 77X24Y, 78X24Y).

Conjunto 5: 21 variables de entrada resultantes del análisis Stepwise efectuado sobre las 114 variables de entrada de las celdas circundantes. (Ver Anexo 6)

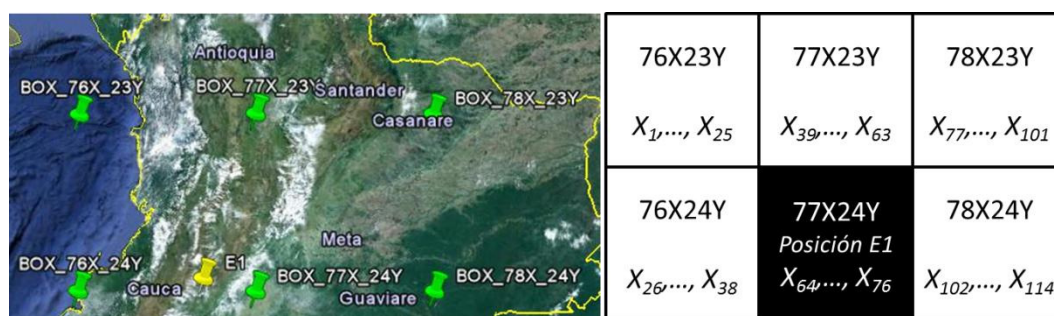


FIGURA 5-3: POSICIÓN ESTACIÓN E1 Y NUMERACIÓN DE VARIABLES

Adicionalmente, a través del Análisis de Componentes Principales (ACP) se obtuvieron dos conjuntos de entrada con 6 y 57 variables derivados de los conjunto 2 y 4 respectivamente para reducir en un 50% las variables de entrada. Se encontró que con 6 y

57 componentes principales se explica respectivamente el 86,46% y el 97,35% de la variabilidad de los datos (Anexo 5).

Los resultados de coeficiente de correlación (r), MAE y MSE en los tres grupos de datos: entrenamiento, validación y prueba para los 10 mejores modelos se encuentran en el Anexo 8. Los 100 modelos probados incluyen 5 modelos de regresión lineal que sirven para mostrar si hay mejora en el uso de redes neuronales frente a las regresiones lineales: regresión lineal de una variable (RL-T), regresión lineal de las 13 variables del conjunto 2 (RL25), regresión lineal de las 9 variables del conjunto 3 (SW25), regresión lineal de las 114 variables del conjunto 4 (RL225) y regresión lineal de las 21 variables del conjunto 5 (SW225). Los datos en la tabla 4 se presentan en orden ascendente desde el mejor coeficiente de correlación en los datos de validación al menor.

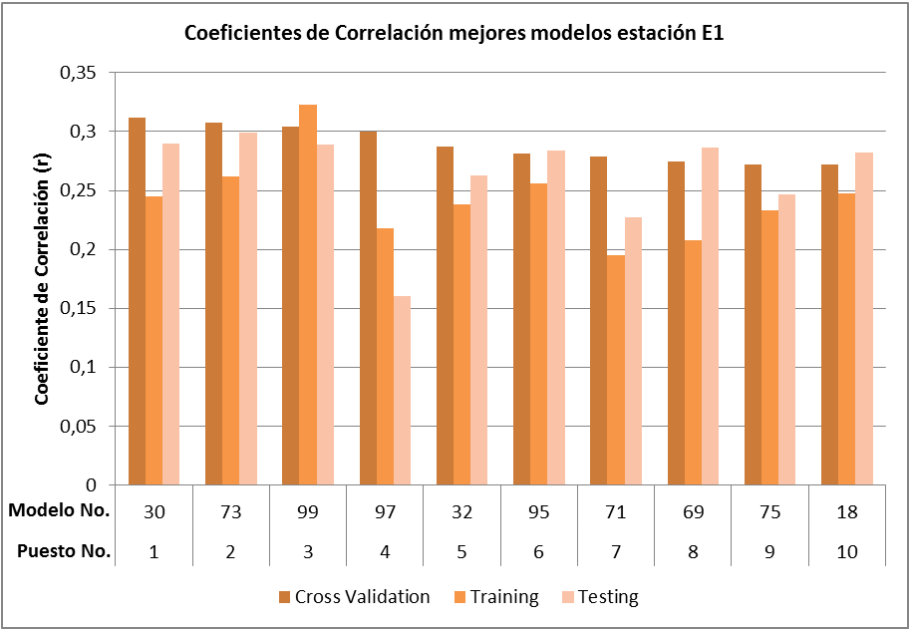


FIGURA 5-4: MEJORES RESULTADOS MODELOS ESTACIÓN E1

Ninguno de los modelos de regresión lineal se encuentra entre los diez primeros mejores modelos. Las 3 primeras arquitecturas corresponden a redes del tipo Time

Lagged Recurrent Networks (TLRN) con coeficientes de correlación en la validación de entre 0,304 a 0,312. En ninguna de las 10 mejores arquitecturas se presentan el método de aprendizaje de LevenbergMarquardt (L) o la presentación de datos en batch (B) como mejor opción.

Aparecen arquitecturas de perceptrones multicapa (MLP) con una (1) y dos (2) capas ocultas, como buenos modelos que aunque presentan coeficientes de correlación más bajos respecto a otras arquitecturas, también tienen una mejor precisión (menor error MAE ó MSE). Las redes recurrentes (RN) aunque presentan similares coeficientes de correlación respecto a las redes TLRN en el grupo de validación, tienen caídas importantes de este coeficiente en el grupo de prueba. Así mismo las arquitecturas Time delay Neural Network (TDNN) presentaron coeficientes de correlación muy cercanos a los de TLRN en el grupo de datos de prueba pero con una precisión inferior (MAE y MSE) comparado a los otros modelos.

Las arquitecturas de menor desempeño fueron las Generalized Feed Forward (GFF) y Radial Basis Function (RBF). También se observó que la reducción de variables en un 50% por el método de Análisis de Componentes Principales disminuye la correlación de los datos y reduce la precisión del modelo, inclusive el caso de ACP para el conjunto 4 donde se logró con 57 componentes principales explicar el 97,35% de la variabilidad de los datos. Con esta metodología el modelo cayó del puesto 5 ($r = 0,287$ MAE = 0,835) al puesto 17 ($r = 0,261$ MAE = 0,843). [Prueba de Mann-Whitney: p-valor = 0,026 $\alpha = 0,05$; La diferencia de posición entre las muestras es diferente de 0]

Al comparar la presentación de los datos en el proceso de aprendizaje de la red de forma On-line (O) y batch (B) para el mejor modelo se encontró que la arquitectura cayó

del puesto 1 al puesto 77 ó en el caso de la arquitectura similar pero con solo 9 variables cayó del puesto 2 al 72. Mostrando un efecto importante de la presentación de los datos durante el aprendizaje. Las mejores arquitecturas que trabajaron con presentación de datos batch fueron los perceptrones multicapa con pocas variables (menor a 57 variables) que corresponden a los puestos 14 a 16. En este caso el algoritmo de aprendizaje Momentum (M) versus LevenbergMarquardt (L) mostró una mejora del desempeño. Esto se replica en las redes recurrentes (RN) pero no en las arquitecturas RBF, GFF y TDNN donde el algoritmo LevenbergMarquardt (L) presentó mejor desempeño.

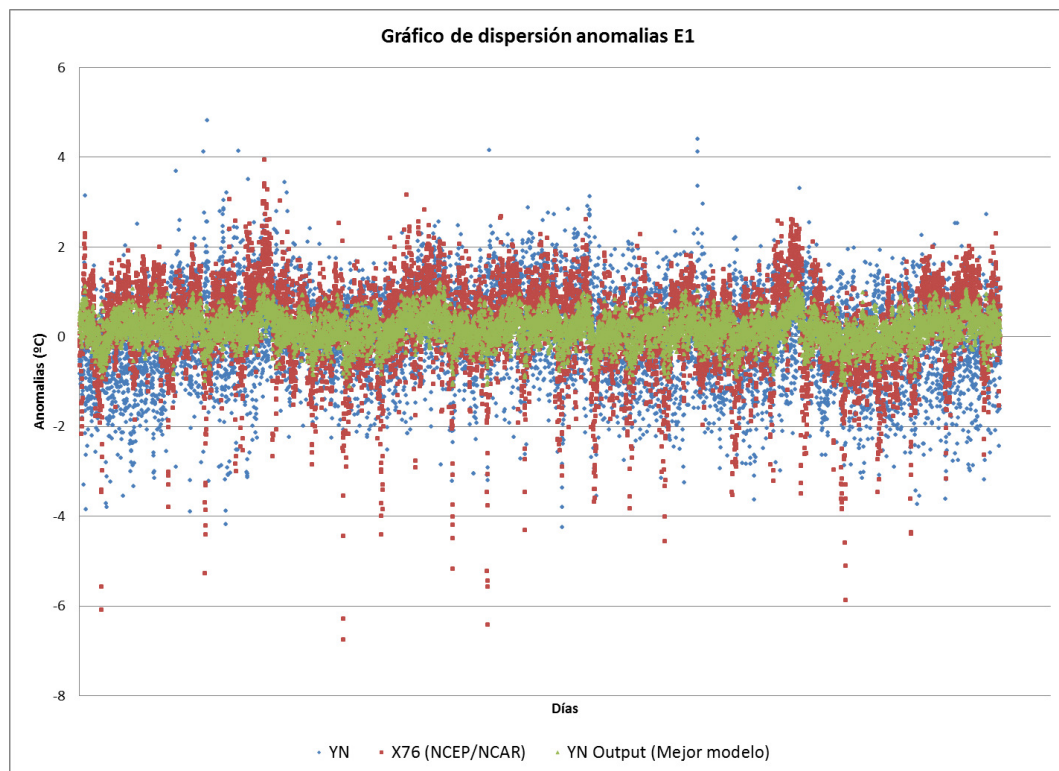


FIGURA 5-5: COMPARATIVO SERIES DE TIEMPO E1

El modelo del puesto No. 2 TLRN-1-O-M que a diferencia del modelo del puesto No. 1 trabajó con 9 variables en vez de 13 además de presentar un coeficiente de correlación similar en la validación (ver Figura 5-4), mostró un mejor desempeño en el

grupo de prueba (correlación y precisión). Se considera este modelo de mayor eficiencia por su precisión con menor número de variables de entrada.

Se seleccionó el modelo TLRN-1-O-M con 9 variables, resultantes de la regresión Stepwise: 1000hPa componente V (X66), 500hPa componente U (X67), 500 hPa geopotencial (X69), 850hPa componente U (X70), 850 componente V (X71), 850 geopotencial (X72), 500hPa humedad específica (X73), 850 humedad específica (X74), 1000hPa humedad específica (X75).

A partir de la arquitectura determinada como mejor modelo se hizo una optimización cambiando la función de transferencia de las capas ocultas y de salida por la función Sigmoidal. El efecto de esta variación sobre el coeficiente de correlación se encuentra en la **¡Error! No se encuentra el origen de la referencia..**

TABLA 3: EFECTO DE LA FUNCION DE TRANSFERENCIA - E1

Función de transferencia		Coef. Corr. (r)
Capa Oculta1	Capa Salida	E1
Tanh	Tanh	0,294
Sigmoidal	Tanh	0,337
Tanh	Sigmoidal	0,005
Sigmoidal	Sigmoidal	-0,08

SE OBTUVO EL MAYOR COEFICIENTE DE CORRELACIÓN UTILIZANDO LA FUNCIÓN DE TRANSFERENCIA SIGMOIDAL EN LA CAPA OCULTA Y UNA FUNCIÓN DE TRANSFERENCIA TANGENTE HIPERBÓLICA (TANH) EN LA CAPA DE SALIDA. ESTA ARQUITECTURA SE ENTRENÓ Y PROBÓ 20 VECES CONSECUTIVAS, INICIANDO CADA VEZ CON PESOS DIFERENTES (VER FIGURA 5-6) PARA ENCONTRAR EL MEJOR AJUSTE DEL MODELO. LOS RESULTADOS DE LA MEJORA EN EL COEFICIENTE DE CORRELACIÓN FRENTE AL REANÁLISIS SE PRESENTAN EN LA

La arquitectura de mejor desempeño corresponde a TLRN-1-O-M con 32 EP en la capa oculta la cual mejoró en un 71,33% la correlación que ofrece actualmente el reanálisis NCEP/NCAR.

Tabla 4.

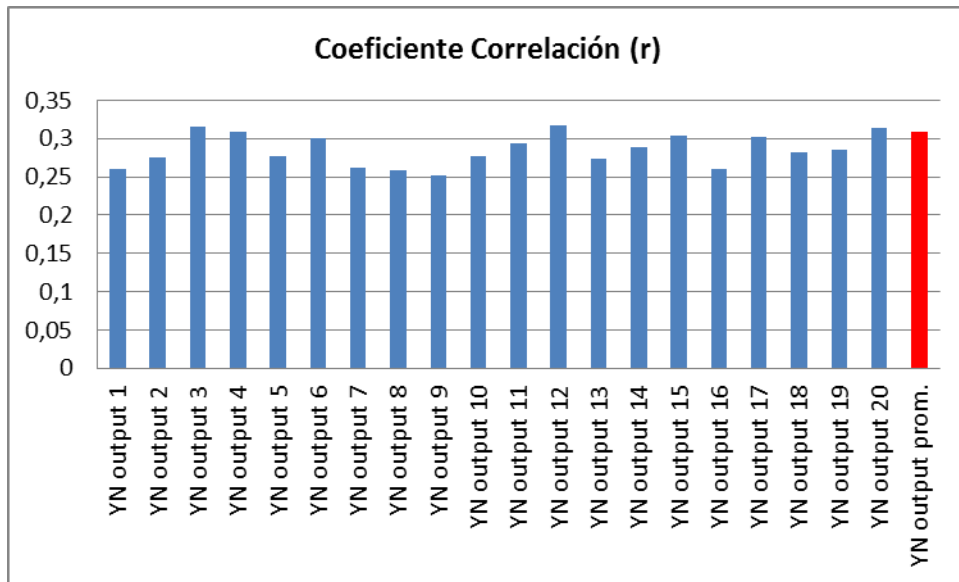


FIGURA 5-6: ENTRENAMIENTOS MEJOR MODELOS PARA SELECCIÓN DE PESOS - E1

La arquitectura de mejor desempeño corresponde a TLRN-1-O-M con 32 EP en la capa oculta la cual mejoró en un 71,33% la correlación que ofrece actualmente el reanálisis NCEP/NCAR.

TABLA 4: MEJORA MODELOS RNA FRENTE A REANÁLISIS NCEP/NCAR

	NCEP/NCAR		Mejor Modelo		Mejora
	Coef. Corr.	Coef. Det.	Coef. Corr.	Coef. Det.	
Pearson	0,164	0,027	0,275	0,076	67,41%
Spearman	0,153	0,023	0,262	0,069	71,33%

5.3.3.2. Estación E2: Isla del Santuario (Código IDEAM 24015120)

La estación meteorológica de Isla del Santuario se encuentra sobre la celda 77X23Y del NCEP/NCAR con una resolución de 3,71° Lat. x 3,75 Long. Esta celda cuenta con las veinticinco variables de entradas completas pero sus celdas circundantes, 76X24Y, 77X24Y y 78X24Y tan solo con 13 cada una, como se describió en la estación E1.

De aquí se desprenden los siguientes 5 conjuntos de variables de entrada:

Conjunto 1: 1 variable correspondiente a la Temperatura a 2m de la celda 77X23Y

Conjunto 2: 25 variables correspondientes a la celda 77X23Y

Conjunto 3: 11 variables resultantes de la regresión Stepwise sobre las 25 variables de la celda 77X23Y (Anexo 6)

Conjunto 4: 189 variables resultantes de las 25 variables de cada una de las celdas circundantes de la franja 22Y y 23Y y la propia celda, más 13 variables de cada una de las celdas circundantes de la franja 24Y.

Conjunto 5: 30 variables resultantes de la regresión Stepwise aplicada a las 189 variables circundantes. (Anexo 6)

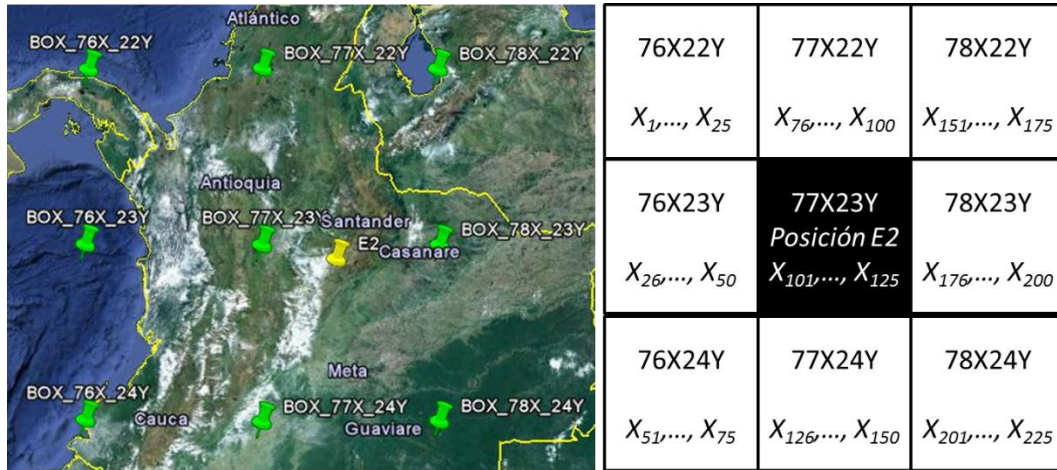


FIGURA 5-7: POSICIÓN ESTACIÓN E2 Y NUMERACIÓN DE VARIABLES

A través del Análisis de Componentes Principales (ACP) se obtuvieron dos conjuntos de entrada con 12 y 94 variables derivados de los conjuntos 2 y 4 respectivamente para reducir en un 50% las variables de entrada. Se encontró que con 12 y 94 componentes principales se explica respectivamente el 87,10% y el 98,22% de la variabilidad de los datos (Anexo 5).

Los resultados de coeficiente de correlación (r), MAE y MSE en los tres grupos de datos: entrenamiento, validación y prueba para los 10 mejores modelos se encuentran en el Anexo 8. Los 100 modelos evaluados incluyen, igualmente a como se hizo con la estación meteorológica E1, 5 modelos de regresión lineal que sirven para mostrar si hay mejora en el uso de redes neuronales frente a las regresiones lineales. Los datos en la tabla se presentan en orden ascendente desde el mejor coeficiente de correlación en los datos de validación al menor.

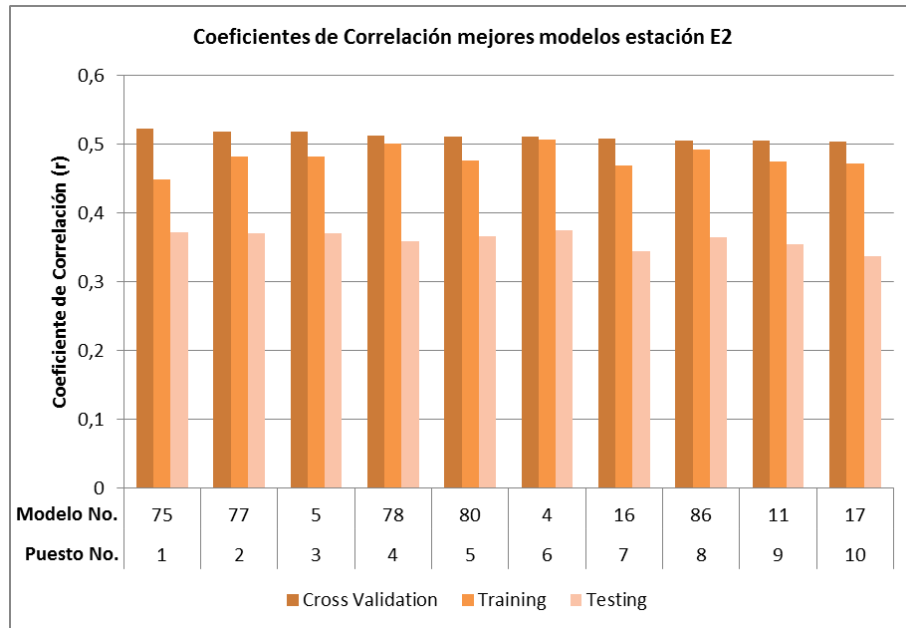


FIGURA 5-8: MEJORES RESULTADOS MODELOS ESTACIÓN E2

Para la estación E2 se observó que las regresiones lineales tuvieron un desempeño bueno frente a los modelos de redes neuronales y se posicionaron en los puestos 3 y 6 dentro de los 10 primeros, con coeficientes de correlación y precisión cercana al mejor modelo que resultó ser un perceptrón multicapa (MLP) con una capa oculta (1), presentación de datos On-line (O) y algoritmo de aprendizaje Momentum (M). Se alcanzó una mejor correlación respecto a lo obtenido en los modelos de la estación E1 con valor de $r = 0,523$ y $MAE = 0,667$.

Los 5 primeros modelos se dan en las arquitecturas o regresiones construidas con las 30 variables de entrada del conjunto 5, que resultó de la aplicación de la metodología de Stepwise a las 189 variables circundantes de la celda correspondiente a la estación meteorológica de Isla del Santuario. Siete de las ocho arquitecturas encontradas en los 10 primeros lugares se basaron en presentación de datos en batch (B) mientras que seis de estas mismas ocho utilizaron el algoritmo de aprendizaje de LevenbergMarquardt (L).

La caída de desempeño por cambio del algoritmo de aprendizaje entre el modelo puesto No. 4 MLP-1-B-L y el puesto No. 8 MLP-1-B-M fue leve ($r = -0,006$ y $MAE = +0,003$), mientras que el cambio de presentación de datos de Batch a On-line tuvo un impacto mayor al comparar el modelo en el puesto No. 1 MLP-1-O-M contra el modelo en el puesto No. 8 MLP-1-B-M ($r = -0,017$ $MAE = 0,013$) [Prueba de Mann-Whitney: p-valor $< 0,0001$ $\alpha = 0,05$; La diferencia de posición entre las muestras es diferente de 0].

Se encontró para este caso que la arquitectura de red neuronal Generalized Feedforward de la forma GFF-1-B-L tiene un buen desempeño cuando se entrenó con pocas variables (menos de 94 variables de entrada).

Las arquitecturas de menor desempeño fueron las RBF, redes dinámicas (TDNN y TLRN), redes recurrentes (RN) y modelos con 189 variables de entrada.

Se observa en los datos que aunque arquitecturas de Análisis de Componentes Principales no obtuvieron los mejores coeficientes de correlación en el grupo de validación, si ajustaron muy bien el grupo de prueba. Caso puntual MLPPCA-1-O-M puesto No. 12 (94 variables de entrada) que obtuvo un r en el grupo de prueba de 0,370 ($r = 0,372$ para la arquitectura de mejor desempeño) y puesto No. 28 (12 variables de entrada) que obtuvo un r en el grupo de prueba de 0,362. Igualmente estos dos modelos obtuvieron MAE menores, en el grupo de prueba, que las 10 arquitecturas de mejor desempeño. Caso similar ocurrió con el modelo en el puesto No. 59 TLRM-1-O-M que tuvo en el grupo de prueba un $r = 0,368$ y $MAE = 0,634$, es decir, tuvo un mejor desempeño con datos nunca antes presentados a la red, comparado a la mejor arquitectura en esta misma condición.

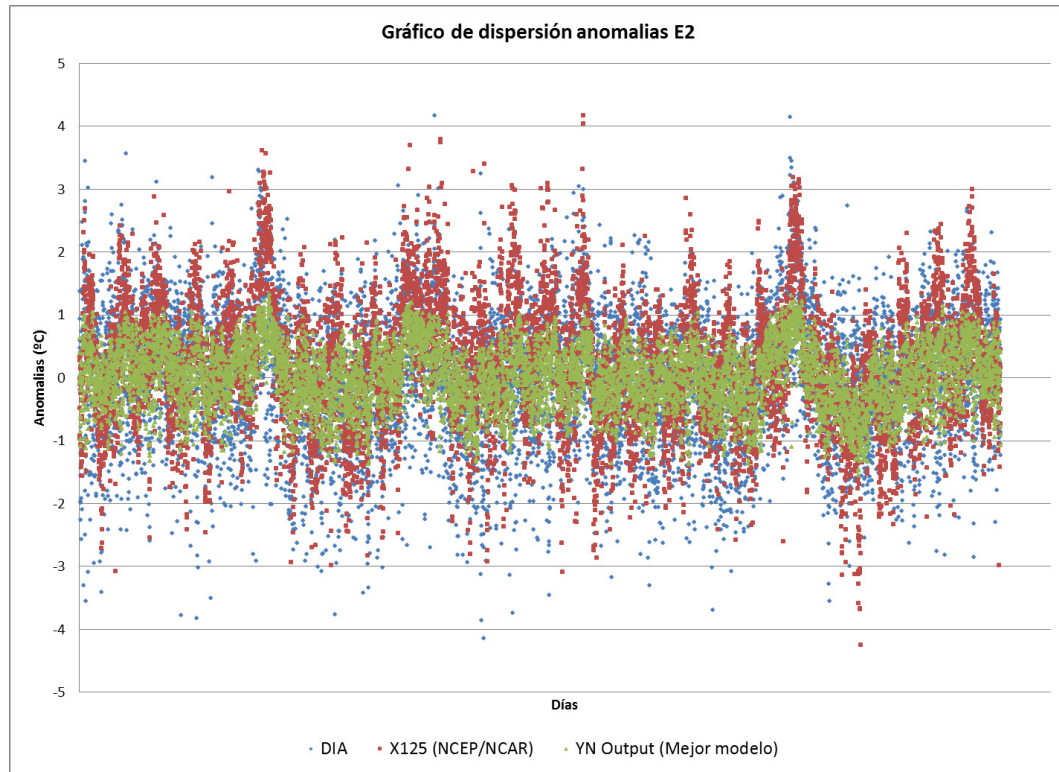


FIGURA 5-9: COMPARATIVO SERIES DE TIEMPO ESTACIÓN E2

En general se observa que los MLP tuvieron mejor desempeño frente a otras arquitecturas y pocos de estos modelos aparecen en la parte inferior de la tabla de posicionamiento (ultimo MLP aparece en el puesto No. 70).

Basado en lo anterior se seleccionó el modelo puesto No. 1 MLP-1-O-M con 30 variables de entrada para la optimización.

A partir de la arquitectura determinada como mejor modelo se hizo una optimización cambiando la función de transferencia de las capas ocultas y de salida por la función Sigmoidal. El efecto de esta variación sobre el coeficiente de correlación se encuentra en la Tabla 5.

TABLA 5: EFECTO DE LA FUNCIÓN DE TRANSFERENCIA - E2

Función de transferencia		Coef. Corr. (r)
Capa Oculta1	Capa Salida	E2
Tanh	Tanh	0,361
Sigmoidal	Tanh	0,351
Tanh	Sigmoidal	0,348
Sigmoidal	Sigmoidal	0,251

SE OBTUVO EL MAYOR COEFICIENTE DE CORRELACIÓN UTILIZANDO LA FUNCIÓN DE TANGENTE HIPERBÓLICA EN LA CAPA OCULTA Y UNA FUNCIÓN DE TRANSFERENCIA EN LA CAPA DE SALIDA. ESTA ARQUITECTURA SE ENTRENÓ Y PROBÓ 20 VECES CADA VEZ CON PESOS DIFERENTES (VER FIGURA 5-10) PARA ENCONTRAR EL MEJOR AJUSTE DEL MODELO. LOS RESULTADOS DE LA MEJORA EN EL COEFICIENTE DE CORRELACIÓN FRENTE AL REANÁLISIS SE PRESENTAN EN LA

Tabla 6.

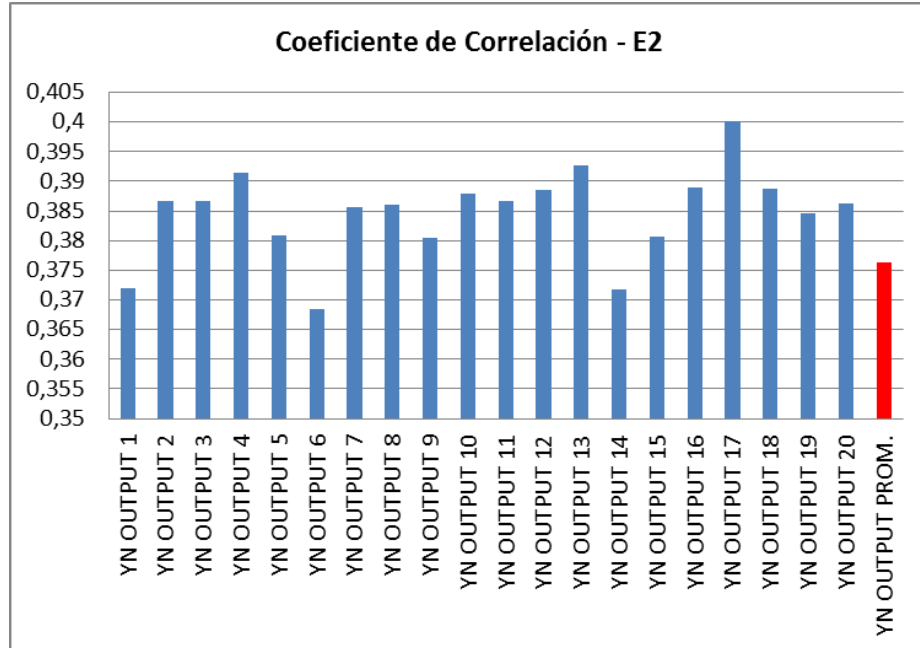


FIGURA 5-10: ENTRENAMIENTOS MEJOR MODELO PARA SELECCION DE PESOS - E2

La arquitectura de mejor desempeño corresponde a MLP-1-O-M con 13 EP en la capa oculta la cual mejoró en un 51,52% la correlación que ofrece actualmente el reanálisis NCEP/NCAR.

TABLA 6: MEJORA MODELOS RNA FRENTE A REANÁLISIS NCEP/NCAR

	NCEP/NCAR		Mejor Modelo		Mejora
	Coef. Corr.	Coef. Det.	Coef. Corr.	Coef. Det.	
Pearson	0,300	0,090	0,463	0,214	54,06%
Spearman	0,297	0,088	0,450	0,202	51,52%

5.3.3.3. Estación E3: Hacienda Las Margaritas (Código IDEAM 35125010)

La estación meteorológica de Hacienda Las Margaritas se encuentra sobre la celda 78X23Y del NCEP/NCAR con una resolución de 3,71° Lat. x 3,75 Long. Esta celda cuenta con las veinticinco variables de entradas completas pero sus celdas circundantes, 77X24Y, 78X24Y y 79X24Y tan solo con 13 cada una, como sucedió en la estación E2.

De aquí se desprenden los siguientes 5 conjunto de variables de entrada:

Conjunto 1: 1 variable correspondiente a la Temperatura a 2m de la celda 78X23Y

Conjunto 2: 25 variables correspondientes a la celda 78X23Y

Conjunto 3: 12 variables resultantes de la regresión Stepwise sobre las 25 variables de la celda 77X23Y (Anexo 6)

Conjunto 4: 189 variables resultantes de las 25 variables de cada una de las celdas circundantes de la franja 22Y y 23Y y la propia celda, más 13 variables de cada una de las celdas circundantes de la franja 24Y.

Conjunto 5: 38 variables resultantes de la regresión Stepwise aplicada a las 189 variables circundantes. (Anexo 6)

A través del Análisis de Componentes Principales (ACP) se obtuvieron dos conjuntos de entrada con 12 y 94 variables derivados de los conjunto 2 y 4 respectivamente para reducir en un 50% las variables de entrada. Se encontró que con 12 y 94 componentes principales se explica respectivamente el 87,27% y el 98,32% de la variabilidad de los datos (Anexo 5).

Se muestran los resultados de coeficiente de correlación (r), MAE y MSE en los tres grupos de datos: entrenamiento, validación y prueba para los 10 mejores modelos se encuentran en el Anexo 8. Dentro de los 100 modelos evaluados se incluyen 5 modelos comparativos de regresión lineal para la respectiva comparación de mejora aportada por las redes neuronales propuestas. Como en las anteriores estaciones, los datos en la tabla se presentan en orden ascendente desde el mejor coeficiente de correlación en los datos de validación al menor.

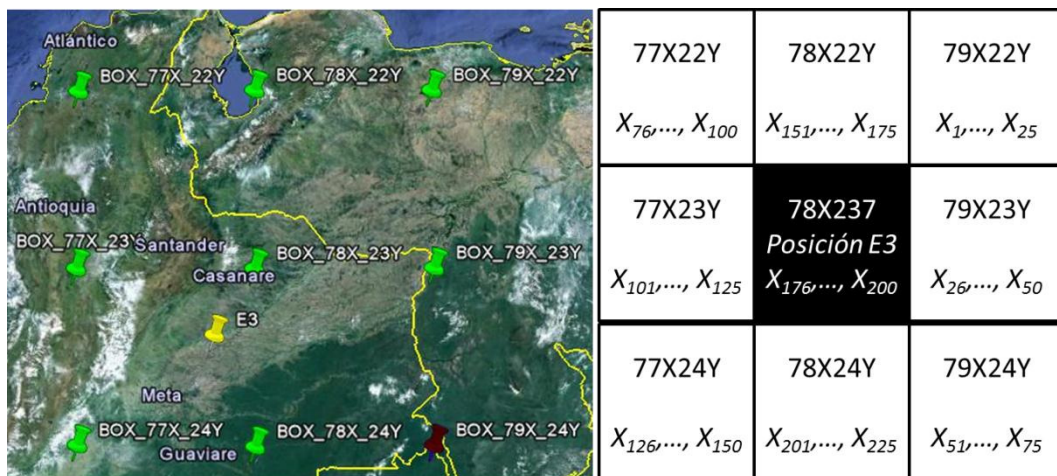


FIGURA 5-11: POSICIÓN ESTACIÓN E3 Y NUMERACIÓN DE VARIABLES

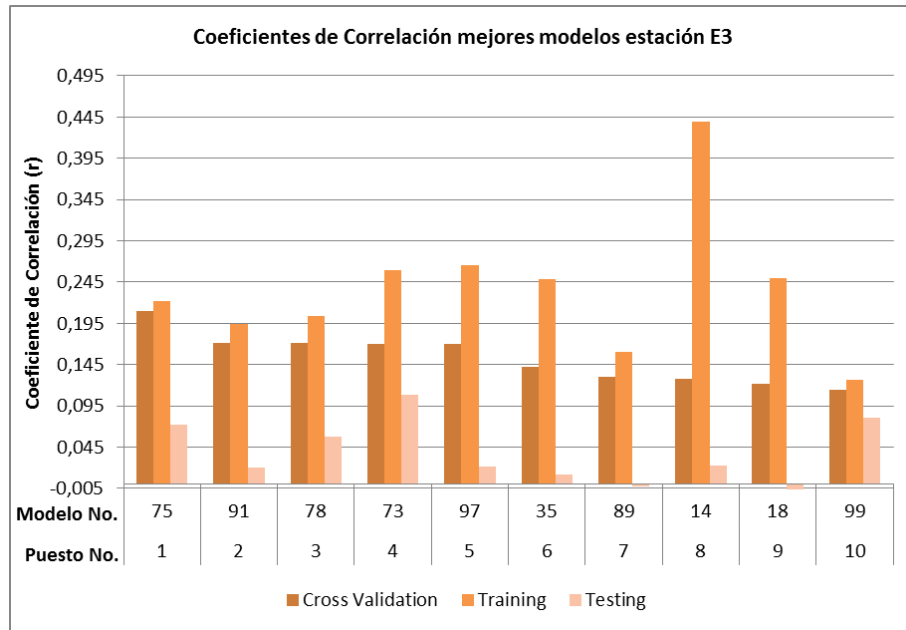


FIGURA 5-12: MEJORES RESULTADOS MODELOS ESTACIÓN E3

Se encontró que aunque la correlación tanto en el grupo de entrenamiento como en el de validación en cada una de las arquitecturas propuestas está, en general, dentro del mismo orden de valores; los coeficientes de correlación (r) en el grupo de prueba caen drásticamente (ver Figura 5-12). Por ejemplo en la arquitectura puesto No. 1, el coeficiente de correlación en el grupo de entrenamiento fue de 0,221, en el de validación 0,209 mientras que en el grupo de prueba cae a 0,072. Comparativamente a los desempeños obtenidos para las estaciones E1 y E2, el ajuste de las arquitecturas para explicar el comportamiento de los datos fue de menor calidad.

Se encontraron entre las 10 mejores arquitecturas los perceptrones multicapa (MLP), las Time Lagged Recurrent Network (TLRN), una Generalizad Feed Forward (GFF) y una Recurrent Network (RN). Ocho de las 10 primeras arquitecturas prefirieron la presentación de datos en forma On-Line (O) frente al Batch (B) y solo los puestos No. 3 y 8

entre los 10 primeros están contruidos con el algoritmo de aprendizaje de LevenbergMarquardt (L).

En términos de variables de entrada la red más eficiente fue la arquitectura correspondiente al puesto No. 4, TLRN-1-O-M, que utiliza 12 variables correspondientes al conjunto 3 de datos. Esta arquitectura a su vez es la que mejor correlaciona datos de prueba ($r = 0,109$ en grupo de prueba) e inclusive tiene una mejor correlación ($r = 0,259$) que la arquitectura del puesto No. 1 en el grupo de entrenamiento ($r = 0,222$). También se observó que la arquitectura del puesto No. 8, TLRN-1-B-L, con 25 variables de entrada (conjunto 1), tuvo el mejor ajuste y precisión en el grupo de entrenamiento ($r = 0,439$ MAE = 0,738). Se presentan las gráficas de series de tiempo en los grupos de entrenamiento, validación y prueba de estos 3 casos en el anexo 9. Se observa que las TLRN representan mejor las características de las anomalías en el tiempo y son más robustas para pronosticar datos extremos.

En términos generales, las arquitecturas con desempeños más bajos corresponden a las Radial Basis Function (RBF), Time Delay Neural Network (TDNN), Perceptrones Multicapa (MLP) con más de una capa oculta y arquitecturas con mayor número de variables de entrada (conjunto 4 y 5).

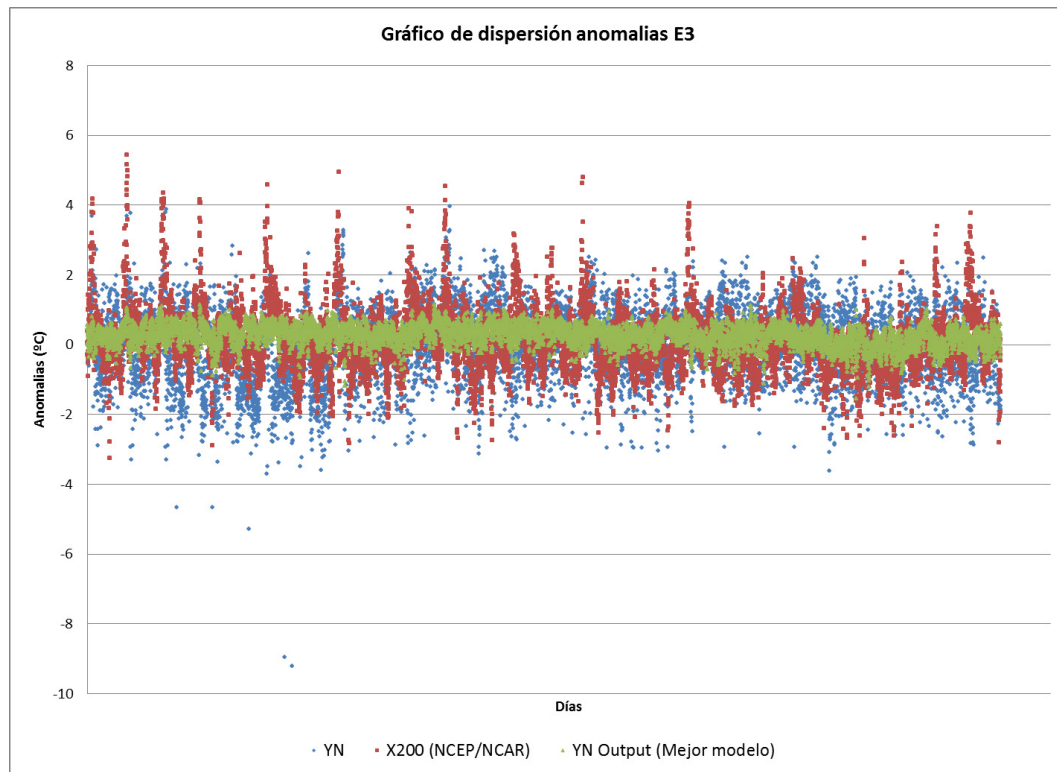


FIGURA 5-13: COMPARATIVO SERIES DE TIEMPO ESTACIÓN E3

Se consideró como mejor modelo para esta estación E3 (Hacienda Las Margaritas) la arquitectura puesto No. 4, TLRN-1-O-M con 12 variables de entrada: Presión en el nivel medio del mar (X176), 1000hPa componente U (X178), 1000hPa vorticidad (X180), 1000hPa divergencia (X182), 500hPa velocidad del viento (183), 500hPa componente V (X185), 500 hPa vorticidad (X186), 850hPa velocidad del viento (X190), 850 vorticidad (X193), 850 geopotencial (X194), 8520hPa velocidad del viento (X195), 500hPa humedad específica (X197) que aunque no presentó el mejor coeficiente de correlación en el grupo de validación, si lo hizo en el grupo de prueba demostrando mejor habilidad en la predicción con datos no presentados a la red.

A partir de la arquitectura determinada como mejor modelo se hizo una optimización cambiando la función de transferencia de las capas ocultas y de salida por la

función Sigmoial. El efecto de esta variación sobre el coeficiente de correlación se encuentra en la Tabla 7.

TABLA 7: EFECTO DE LA FUNCIÓN DE TRANSFERENCIA - E3

Función de transferencia		Coef. Corr. (r)
Capa Oculta1	Capa Salida	E3
Tanh	Tanh	0,172
Sigmoial	Tanh	0,044
Tanh	Sigmoial	0,017
Sigmoial	Sigmoial	0,021

Se obtuvo el mayor coeficiente de correlación utilizando la función de transferencia Tangente hiperbólica en la capa oculta y una función de transferencia Tangente hiperbólica en la capa de salida. Esta arquitectura se entrenó y probó 20 veces consecutivas, iniciando cada vez con pesos diferentes (ver Figura 5-14) para encontrar el mejor ajuste del modelo. Los resultados de la mejora en el coeficiente de correlación frente al reanálisis se presentan

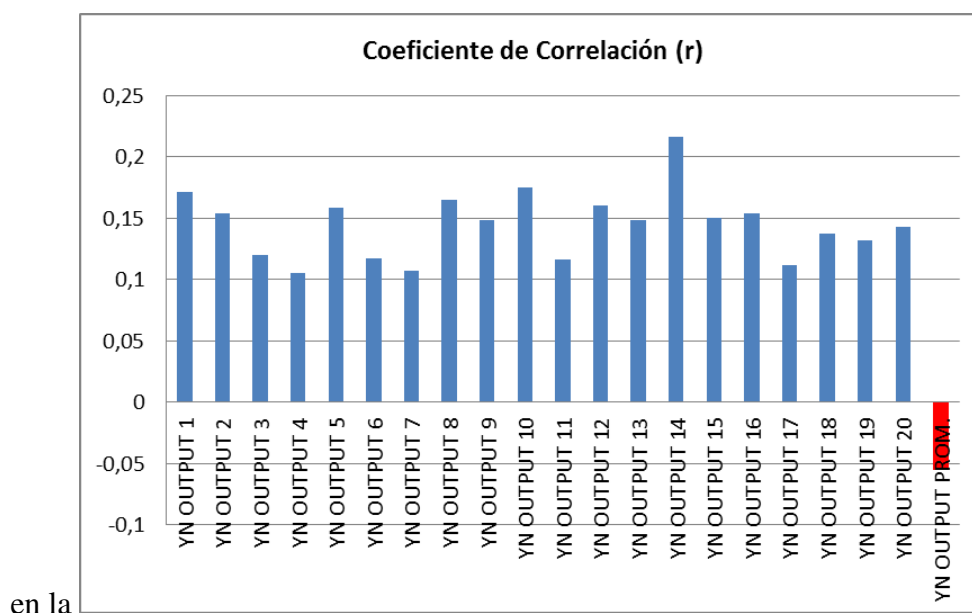


FIGURA 5-14: ENTRENAMIENTOS MEJOR MODELO PARA SELECCIÓN DE PESOS - E3

Tabla 8.

La arquitectura de mejor desempeño corresponde a TLRN-1-O-M con 5 EP en la capa oculta la cual mejoró en un 78,37% la correlación que ofrece actualmente el reanálisis NCEP/NCAR.

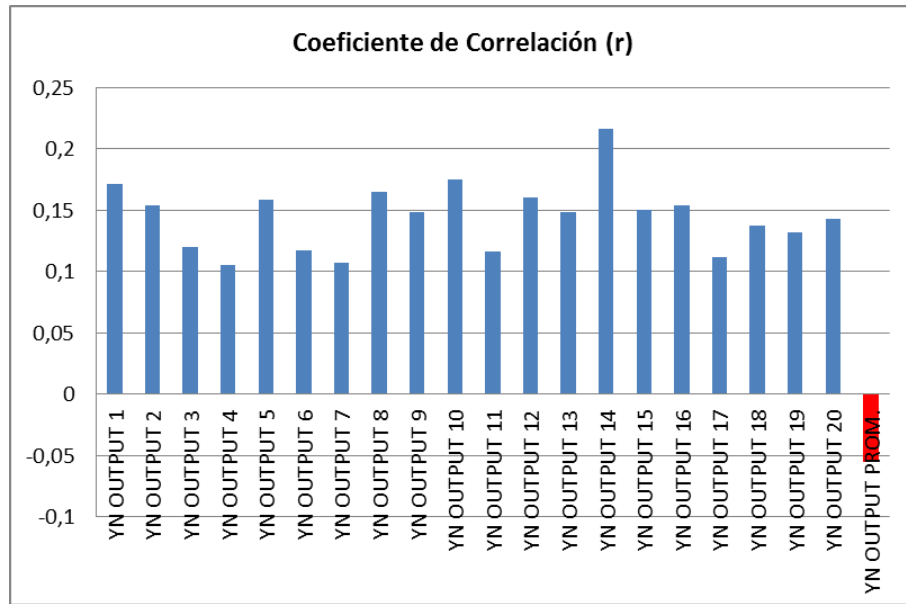


FIGURA 5-14: ENTRENAMIENTOS MEJOR MODELO PARA SELECCIÓN DE PESOS - E3

TABLA 8: MEJORA MODELOS RNA FRENTE A REANÁLISIS NCEP/NCAR

	NCEP/NCAR		Mejor Modelo		Mejora
	Coef. Corr.	Coef. Det.	Coef. Corr.	Coef. Det.	
Pearson	0,122	0,015	0,173	0,030	42,48%
Spearman	0,090	0,008	0,161	0,026	78,37%

5.3.3.4. Estación E4: Manaure (Código IDEAM 15075030)

La estación meteorológica E4 (Manaure) se encuentra ubicada sobre la celda 78X21Y del NCEP/NCAR con una resolución de 3,71° Lat. x 3,75 Long. Esta celda cuenta con las veinticinco variables de entradas completas al igual que sus celdas circundantes

77X21Y, 79X21Y y 77X22Y, 78X22Y 79X22Y, pero no se contó con datos de la franja superior 20Y.

De aquí se desprenden los siguientes 5 conjunto de variables de entrada:

Conjunto 1: 1 variable correspondiente a la Temperatura a 2m de la celda 78X21Y

Conjunto 2: 25 variables correspondientes a la celda 78X21Y

Conjunto 3: 11 variables resultantes de la regresión Stepwise sobre las 25 variables de la celda 78X21Y (Anexo 6)

Conjunto 4: 150 variables resultantes de las 25 variables de cada una de las celdas circundantes de la franja 21Y y 22Y.

Conjunto 5: 18 variables resultantes de la regresión Stepwise aplicada a los conjunto 4 de datos. (Anexo 6)

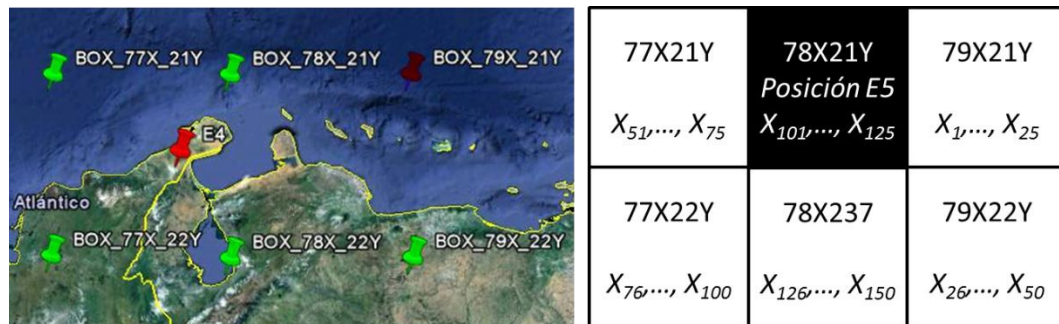


FIGURA 5-15: POSICIÓN ESTACIÓN E4 Y NUMERACIÓN DE VARIABLES

A través del Análisis de Componentes Principales (ACP) se obtuvieron dos conjuntos de entrada con 12 y 75 variables derivados de los conjunto 2 y 4 respectivamente para reducir en un 50% las variables de entrada. Se encontró que con 12 y 75 componentes principales se explica respectivamente el 91,96% y el 98,56% de la variabilidad de los datos (Anexo 5).

Se muestran los resultados de coeficiente de correlación (r), MAE y MSE en los tres grupos de datos: entrenamiento, validación y prueba para los 10 mejores modelos en la Anexo 8. Dentro de los 100 modelos se incluyen 5 modelos comparativos de regresión lineal para la respectiva comparación de mejora aportada por las redes neuronales propuestas. Los datos en la tabla se presentan en orden ascendente desde el mejor coeficiente de correlación en los datos de validación al menor.

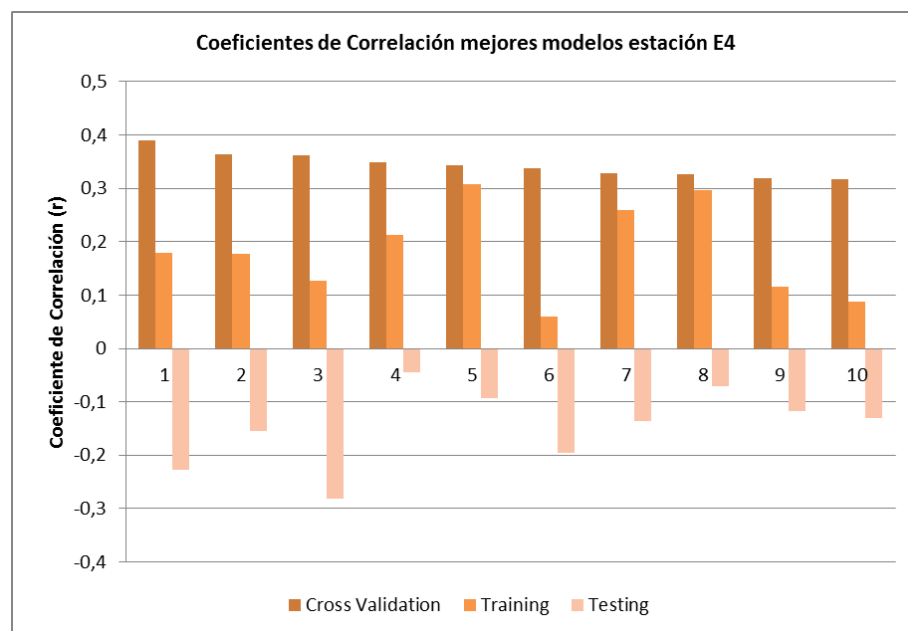


FIGURA 5-16: MEJORES RESULTADOS MODELOS ESTACIÓN E4

Se encontró que ninguno de los 10 modelos con mayor coeficiente de correlación (r) en el grupo de validación tuvo un coeficiente de correlación positivo para el grupo de prueba. Las arquitecturas con mayor coeficiente de correlación corresponden a Time Delay Neural Network (TDNN), Perceptrones Multicapa (MLP) con una y dos capas ocultas, Recurrent Network (RN), Time Lagged Recurrent Network (TLRN) y Generalized Feed Forward. Las arquitecturas TLRN tuvieron coeficientes de correlación más altos en el

grupo de entrenamiento frente a las demás arquitecturas ($r = 0,260 - 0,309$). Y ocuparon los puestos No. 5, 7 y 8 entre los diez primeros; pero presentaron coeficientes de correlación, en el grupo de prueba, bajos y negativos.

En orden de magnitud, se encontró que la arquitectura del Perceptrón Multicapa construida con 75 Componentes Principales tiene la mejor correlación (aunque negativa) en los datos de prueba frente a los 10 primeros puestos.

Nueve de los diez primeros modelos trabajaron con presentación de datos On-Line (O) y algoritmo de aprendizaje Momentum (M). Solo la arquitectura puesto No. 2, MLP-2-B-L, con 18 variables del entrada del conjunto 5 y 2 capas ocultas utilizó presentación de datos Batch (B) y el algoritmo de aprendizaje de LevenbergMarquardt (L). Esta red tuvo desempeño similar al puesto No. 1, TDNN-1-O-M con 11 variables de entrada del conjunto 3, pero con una mejor precisión en el grupo de prueba ($MAE = 0,737$).

En términos generales se observa que las arquitecturas del tipo RBF son las de menor desempeño aunque se notó que fueron las únicas capaces de interpretar con coeficientes de correlación (r) positivos al grupo de prueba. También se encontró en el puesto No. 70 una arquitectura TLRN-1-B-L con 25 variables de entrada del conjunto 2 que fue capaz de tener un coeficiente de correlación moderado ($r = 0,369$) en el grupo de entrenamiento y un coeficiente de correlación positivo en el grupo de prueba ($r = 0,117$). Cuando se cambió el tipo de algoritmo de aprendizaje de LevenbergMarquardt (L) Momentum, la correlación mejoró en el grupo de validación pero cayó en el grupo de entrenamiento y se tornó negativa en el grupo de prueba. Las semillas aleatorias de arranque son diferentes.

Basado en los resultados expuestos anteriormente, se prefiere para el pronóstico de anomalías de la temperatura en la estación meteorológica E4, Manaure, la arquitectura de tipo TLRN-1-B-L con 25 variables de entrada correspondientes al conjunto 2, por ajustar mejor los datos de entrenamiento y prueba.

Debido a los resultados sobre esta estación, la baja correlación presentada por las arquitecturas, no se hizo optimización ni pruebas con funciones de transferencia.

5.3.3.5. Estación E5: Apto Rafael Barvo (Código IDEAM 25025080)

La estación meteorológica E5 (Apto Rafael Barvo) se encuentra ubicada sobre la celda 77X22Y del NCEP/NCAR con una resolución de 3,71° Lat. x 3,75 Long. Esta celda cuenta con las veinticinco variables de entradas completas al igual que sus celdas circundantes de las franjas 21Y, 22Y y 23Y.

De aquí se desprenden los siguientes 5 conjuntos de variables de entrada:

Conjunto 1: 1 variable correspondiente a la Temperatura a 2m de la celda 77X22Y

Conjunto 2: 25 variables correspondientes a la celda 77X22Y

Conjunto 3: 7 variables resultantes de la regresión Stepwise sobre las 25 variables de la celda 77X22Y (Anexo 6)

Conjunto 4: 225 variables resultantes de las 25 variables de cada una de las celdas circundantes de la franja 21Y, 22Y y 23Y.

Conjunto 5: 31 variables resultantes de la regresión Stepwise aplicada a las conjunto 4 de datos. (Anexo 6)

A través del Análisis de Componentes Principales (ACP) se obtuvieron dos conjuntos de entrada con 12 y 112 variables derivados de los conjunto 2 y 4 respectivamente para reducir en un 50% las variables de entrada. Se encontró que con 12 y

112 componentes principales se explica respectivamente el 90,11% y el 98,66% de la variabilidad de los datos (Anexo 5).

Se muestran los resultados de coeficiente de correlación (r), MAE y MSE en los tres grupos de datos: entrenamiento, validación y prueba para los 10 mejores modelos en el Anexo 8. Dentro de los 100 modelos se incluyen 5 modelos comparativos de regresión lineal para la respectiva comparación de mejora aportada por las redes neuronales propuestas. Los datos en la tabla se presentan en orden ascendente desde el mejor coeficiente de correlación en los datos de validación al menor.

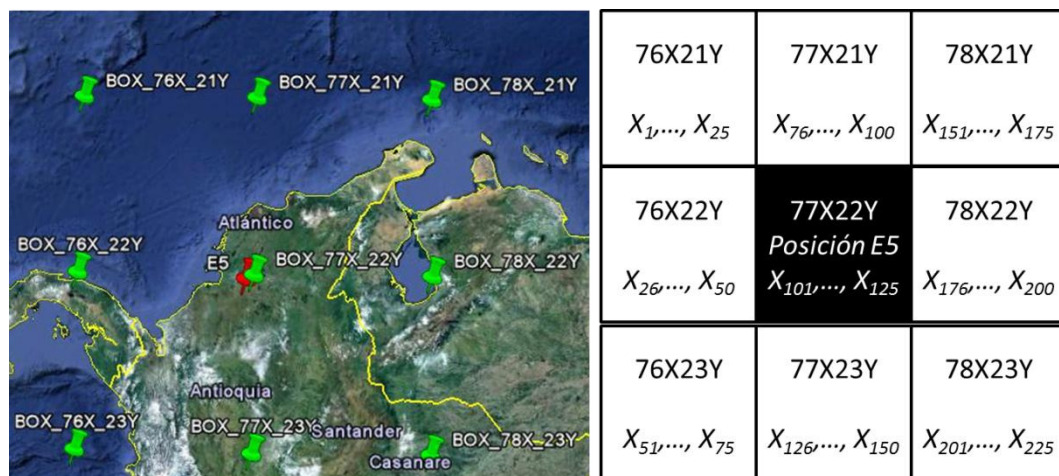


FIGURA 5-17: POSICIÓN ESTACIÓN E5 Y NUMERACIÓN DE VARIABLES

Se observó un claro dominio de los Perceptrones Multicapa (MLP) para aprender mejor las características de la serie de tiempo de anomalías de la temperatura en la estación E5 (Apto Rafael Barvo). Los 3 mejores perceptrones multicapa fueron los de dos capas ocultas con algoritmo de aprendizaje Momentum. La presentación de los datos de retroalimentación en Batch (B) mostró una leve mejora de la precisión (MAE) frente al método On-line pero perdiendo levemente la correlación.

Otras arquitecturas de buen desempeño fueron los perceptrones multicapa de una capa oculta, las Generalized Feedforward y en el puesto No. 6 apareció una Red Recurrente (RN). Aparte de los MLP las demás arquitecturas se ajustaron mejor con presentación de datos en forma On-line (O) y algoritmo de aprendizaje Momentum (M).

Los 10 modelos con mejor desempeño utilizaron como variables de entrada bien sea el conjunto 3 o el conjunto 5 de datos que corresponden a la determinación de variables derivadas por la regresión Stepwise. En general son arquitecturas muy eficientes en términos de aprendizaje de las características de la serie de tiempo con pocas variables de entrada.

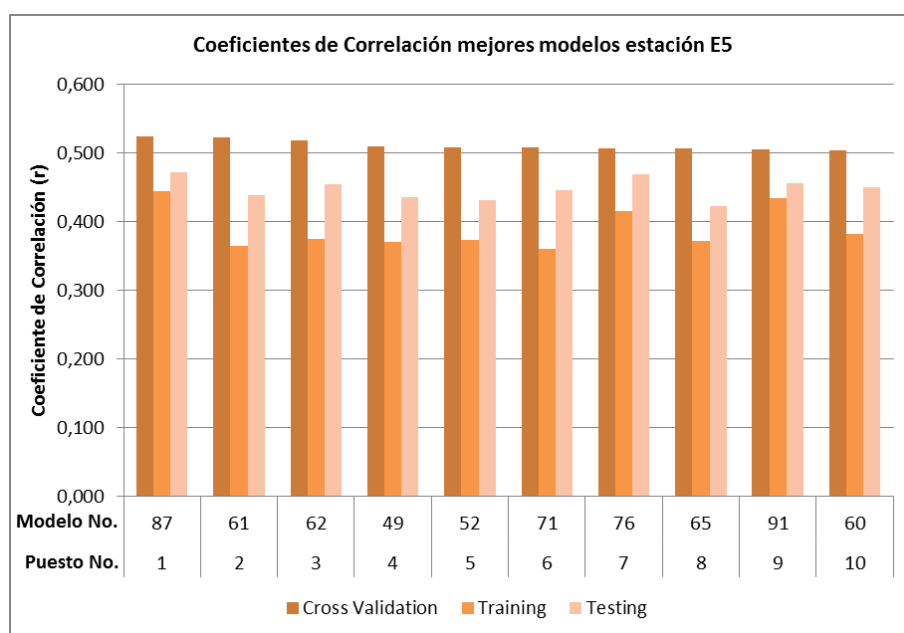


FIGURA 5-18: MEJORES RESULTADOS MODELOS ESTACIÓN E5

Se encontró que una regresión lineal con 31 variables de entrada, correspondientes al conjunto 5 de datos, ocupó el puesto No. 7 de desempeño con correlación y precisión ($r = 0,470$ MAE = 0,717) similar al modelo puesto No. 1 en el grupo de prueba ($r = 0,472$ MAE = 0,745).

En general las arquitecturas Radial Basis Function (RBF), Time Delayed Neural Network (TDNN), Time lagged Recurrent Network (TLRN) y sobre todo con alto número de variables de entrada, presentaron los menores desempeños. Excepto la Red Recurrente (RN) puesto No. 6, las RN tuvieron también desempeños inferiores a otras arquitecturas similares en algoritmo, presentación de datos y variables de entrada.

Solo una red, MLP-1-B-L con 7 variables de entrada puesto No. 5, de las 10 primeras, utilizó el algoritmo de aprendizaje de LevenbergMarquardt (L).

Se seleccionó como mejor arquitectura el puesto No. 1, MLP-2-O-M con 31 variables de entrada, por su correlación y precisión superior a las demás arquitecturas en el grupo de validación, y sus resultados buenos en los grupos de entrenamiento y prueba.

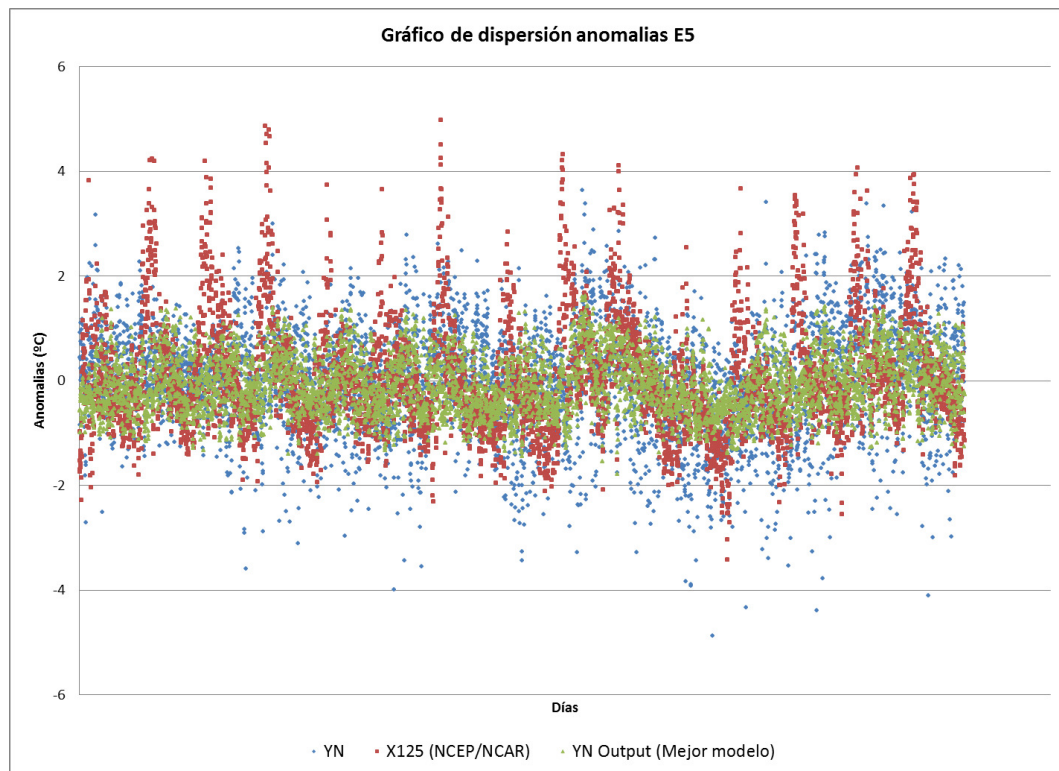


FIGURA 5-19: COMPARATIVO SERIES DE TIEMPO E5

A partir de la arquitectura determinada como mejor modelo se hizo una optimización cambiando la función de transferencia de las capas ocultas y de salida por la función Sigmoidal. El efecto de esta variación sobre el coeficiente de correlación se encuentra en la Tabla 9.

Se obtuvo el mayor coeficiente de correlación utilizando la función de transferencia Tangente hiperbólica en la capa oculta 1 y la capa de salida y una función de transferencia Sigmoidal en la capa oculta 2. Esta arquitectura se entrenó y probó 20 veces consecutivas, iniciando cada vez con pesos diferentes (ver Figura 5-20) para encontrar el mejor ajuste del modelo. Los resultados de la mejora en el coeficiente de correlación frente al reanálisis se presentan en la Tabla 10.

TABLA 9: EFECTO DE LA FUNCIÓN DE TRANSFERENCIA - E5

Función de transferencia			Coef. Corr. (r)
Capa Oculta 1	Capa Oculta 2	Capa Salida	E3
Tanh	Tanh	Tanh	0,471
Tanh	Sigmoidal	Tanh	0,48
Sigmoidal	Tanh	Sigmoidal	0,464
Sigmoidal	Tanh	Tanh	0,474
Tanh	Sigmoidal	Sigmoidal	0,472
Sigmoidal	Sigmoidal	Sigmoidal	0,469

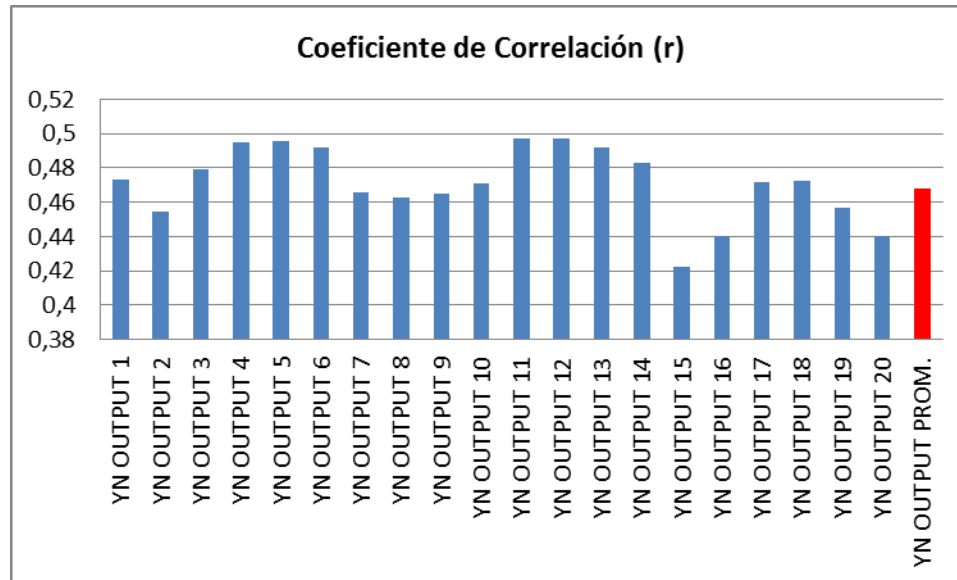


FIGURA 5-20: ENTRENAMIENTOS MEJOR MODELO PARA SELECCIÓN DE PESOS - E5

La arquitectura de mejor desempeño corresponde a MLP-2-O-M con 17 y 6 EP en las capas ocultas la cual mejoró en un 89,96% la correlación que ofrece actualmente el reanálisis NCEP/NCAR.

TABLA 10: MEJORA MODELOS RNA FRENTE A REANÁLISIS NCEP/NCAR

	NCEP/NCAR		Mejor Modelo		Mejora
	Coef. Corr.	Coef. Det.	Coef. Corr.	Coef. Det.	
Pearson	0,231	0,053	0,479	0,229	107,54%
Spearman	0,242	0,059	0,460	0,212	89,96%

5.4. DESARROLLO DEL MODELO COMPUESTO

Modelo compuesto se llamó al modelo que mediante una sola arquitectura tuvo como variables de salida, simultáneamente las anomalías de cada una de las cinco estaciones evaluadas. Para ellos se construyeron redes neuronales con cinco salidas, una para cada estación, y se corrieron entrenamientos, validaciones cruzadas y pruebas para determinar su desempeño. Los datos diarios de anomalías de la temperatura fueron

seleccionados en el período comprendido entre el 01/01/1987 y el 31/12/2003 por en este periodo en el que se encuentra información para todas las estaciones meteorológicas de estudio.

5.4.1. Arquitecturas evaluadas

Del Capítulo 5.3.3. por estación, se observó que las arquitecturas de mejor desempeño para el aprendizaje de las características de los datos de anomalías de temperatura fueron los Perceptrones Multicapa (MLP) y las redes Time Lagged Recurrent Network, siendo estas últimas las más eficientes en términos de número de variables de entrada.

Haciendo observación sobre las estaciones E1, E2, E3 y E5 se determinó que tiene mejor desempeño tanto computacional como de correlación y precisión el método de presentación de datos On-Line (O). Adicional el algoritmo de aprendizaje Momentum (M) mostró mejor convergencia al mínimo y por tanto se seleccionaron estos dos parámetros para la construcción del modelo compuesto. En cuanto al conjunto de variables de entrada las estaciones E1, E2, E3 y E5 mejoraron su desempeño al trabajar con el grupo de datos conjunto 3 ó 5 que corresponden a los conjuntos derivados de la aplicación de la regresión Stepwise sobre los conjunto 2 y 4 respectivamente, para la reducción de variables de entrada.

La estación E4 se describe en este trabajo como un caso especial donde las arquitecturas evaluadas (100 modelos) no mostraron facilidad para el aprendizaje de los patrones de comportamiento, dando alta dispersión en MAE entre corridas con semillas de pesos diferentes, bajos coeficientes de correlación y alto costo computacional para la convergencia al mínimo. Por tanto, aunque se determinó como el mejor modelo la arquitectura TLRN-1-B-L con 25 variables de entrada, no se considera que tenga un buen

desempeño comparado a las otras estaciones y si puede estar generando ruido en las observaciones sobre mejores desempeños encontradas en las demás estaciones. Motivo por el cual se excluyó de este análisis.

Las variables de entrada que se utilizaron para este modelo compuesto corresponden a las 352 variables derivadas de las 25 variables de cada una de las 12 celdas de la franja 21Y, 22Y y 23Y y 13 variables disponibles por cada celda 76X24Y, 77X24Y, 78X24Y y 79X24Y que se emplearon en este estudio, correspondientes a las variables del salida del reanálisis NCEP/NCAR.

Para la reducción del número de variables de entrada se pudo haber considerado bien sea intentar hacer regresiones Stepwise entre las 352 variables de entrada y las 5 variables de salida o utilizar las variables de los Conjunto 3 ó 5 encontradas en los modelos por estación. En cualquiera de los dos casos se considera que se puede perder información, por ejemplo de interacción entre celda continuas, por lo que se prefirió en este modelo compuesto aplicar un análisis de sensibilidad y correr nuevamente modelos con las 50 y 25 variables de entrada que más aportan a la caracterización de la variabilidad de la red neuronal.

5.4.2. Resultados Modelo Compuesto

De acuerdo a los parámetros y decisiones descritas anteriormente, se obtuvieron nuevos modelos para evaluación, 3 por cada conjunto de variables de entrada:

Conjunto 1: 352 variables de entrada correspondientes a las variables del reanálisis disponibles para las 16 celdas de estudio.

Conjunto 2: 50 variables de entrada derivadas del análisis de sensibilidad.

Conjunto 3: 25 variables de entrada derivadas del análisis de sensibilidad.

Inicialmente se corrieron los 3 modelos del conjunto 1 de variables de entrada: MLP-1-O-M, MLP-2-O-M Y TLRN-1-O-M para determinar el de mejor desempeño. Los resultados se muestran en el Anexo 10. Se encontró una caída importante en el desempeño de la arquitectura TLRN ($r = 0,078$ MAE = 0,907) frente a los MLP (MLP-1-OM $r = 0,189$ MAE = 0,907; MLP-2-O-M $r = 0,152$ MAE = 0,865) cuando se trabajó con 352 variables de entrada.

Se corrió un análisis de sensibilidad sobre la arquitectura MLP-1-O-M por ser esta la de mejor desempeño y se determinaron las 50 y 25 variables de los conjunto 2 y 3. Los resultados del análisis de sensibilidad se presentan en el Anexo 11 ordenadas de mayor a menor sensibilidad.

Estos conjunto 2 y 3 fueron evaluados en las mismas arquitecturas del conjunto 1 y sus resultados de desempeño se presentan en el Anexo 10. Se observa que el mayor coeficiente de correlación en el grupo de validación lo presenta la arquitectura MLP-2-O-M con 25 variables de entrada, seguida de MLP-1-M con 50 variables y TLRN-1-O-M con 25 variables. En el grupo de prueba el mejor desempeño lo presenta la arquitectura MLP-1-O-M con 352 variables de entrada, seguida por TLRN-1-O-M con 25 variables de entrada y MLP-1-O-M con 50 variables. En el grupo de entrenamiento el mejor ajuste lo presentó TLRN-1-O-M seguida de MLP-2-O-M y MLP-1-O-M, todas con 25 variables de entrada. Hubo una tendencia a ajustar y modelar mejor los datos de las anomalías de temperatura por parte de la arquitectura TLRN al decrecer el número de variables de entrada.

Considerando la cercanía en coeficiente de correlación de la arquitectura TLRN-1-O-M con 25 variables de entrada ($r = 0,248$) frente al modelo de mayor desempeño, MLP-2-O-M ($r = 0,257$), en este mismo grupo y su mejor desempeño en los grupos de entrenamiento y prueba; se considera este modelo como de buen desempeño.

Una comparación del desempeño de los modelos individuales seleccionados para cada estación frente al modelo compuesto seleccionado se presenta en la Tabla 11.

Se observa un mejor desempeño de los modelos individuales sobre todo en el grupo de prueba. El modelo compuesto ajusta de manera equivalente a los individuales en el grupo de entrenamiento pero se empieza a observar un leve distanciamiento en el grupo de validación. Este distanciamiento se acentúa en el grupo de prueba.

TABLA 11: COMPARATIVO DESEMPEÑOS MODELOS INDIVIDUALES Y COMPUESTO

Training										
Performance	YE1		YE2		YE3		YE4		YE5	
	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual
MSE	1,370	0,911	0,809	0,832	0,802	1,105	1,636	0,892	0,818	0,718
MAE	0,963	0,755	0,709	0,722	0,714	0,810	1,053	0,741	0,707	0,670
r	0,341	0,308	0,460	0,450	0,169	0,253	0,350	0,369	0,308	0,445
Cross-validation										
Performance	YE1		YE2		YE3		YE4		YE5	
	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual
MSE	1,202	0,855	0,680	0,710	0,763	0,730	0,873	1,089	1,075	0,944
MAE	0,921	0,749	0,649	66699559,000	0,693	0,682	0,752	0,832	0,809	0,764
r	0,248	0,368	0,322	0,523	0,147	0,196	0,143	0,126	0,430	0,525
Test										
Performance	YE1		YE2		YE3		YE4		YE5	
	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual	Modelo compuesto	Modelo Individual
MSE	1,016	1,118	0,715	0,656	0,796	0,873	0,780	0,910	0,980	0,875
MAE	0,811	0,812	0,678	0,644	0,684	0,723	0,704	0,748	0,781	0,745
r	0,119	0,271	0,194	0,372	0,117	0,101	0,007	0,117	0,322	0,473

5.5. COMPARACIÓN MODELOS RNA CONTRA REANÁLISIS NCEP/NCAR

Las salidas de los modelos RNA por estación y compuesto seleccionados como mejores modelos (ver Tabla 12) fueron comparados con la salida del reanálisis NCEP/NCAR, variable No. 25, Temperatura 2m, de la celda correspondiente a la ubicación de cada estación meteorológica de estudio.

TABLA 12: MEJORES MODELOS SELECCIONADOS

Estación	Arquitectura	No. De variables	Conjunto de datos
E1	TLRN-1-O-M	9VAR	CONJUNTO 3
E2	MLP-1-O-M	30 VAR	CONJUNTO 5
E3	TLRN-1-O-M	12 VAR	CONJUNTO 3

E4	TLRN-1-B-L	25 VAR	CONJUNTO 2
E5	MLP-2-O-M	31 VAR	CONJUNTO 5
Compuesto	MLP-2-O-M	25 VAR	CONJUNTO 3

Se observó en las gráficas de correlación y error medio (ver Figura 5-21) que en todos los casos, excepto para la estación E3 en el modelo compuesto, los modelos por estación o compuesto mejoran la simulación de las anomalías de la temperatura media diaria superficial. Al igual que en los modelos por estación, el modelo compuesto correlaciona mejor las anomalías en las estaciones meteorológicas E2 y E5 ($r^2 = 0,1722$ Vs. $0,0719$ y $r^2 = 0,1594$ Vs. $0,0533$ respectivamente) al compararse con la salida de la variable 25 del reanálisis. El modelo compuesto también presentó una mejora, de menor orden de magnitud, en la simulación de las anomalías de la estación E1 ($r^2 = 0,0760$ Vs. $0,0516$) y E4 ($r^2 = 0,0246$ Vs. $0,0052$). Para el caso de la estación E3 se evidencia que el modelo compuesto no es aplicable pues no es capaz de simular las anomalías. Los modelos por estación presentan tendencias similares al modelo compuesto mejorando importante mente en las estaciones E2 y E5 y en menor orden en las estaciones E1, E3 y E4. El modelo de la estación E3 si presentó una mejor simulación tendencial de las anomalías ($r^2 = 0,0300$ Vs. $0,0148$) aunque de bajo orden de magnitud.

Así mismo los errores medios de los modelos por estación y compuesto fueron menores a los obtenidos con el reanálisis, mostrando la mayor habilidad de los modelos RNA para la simulación de las anomalías de temperatura de las estaciones de estudio.

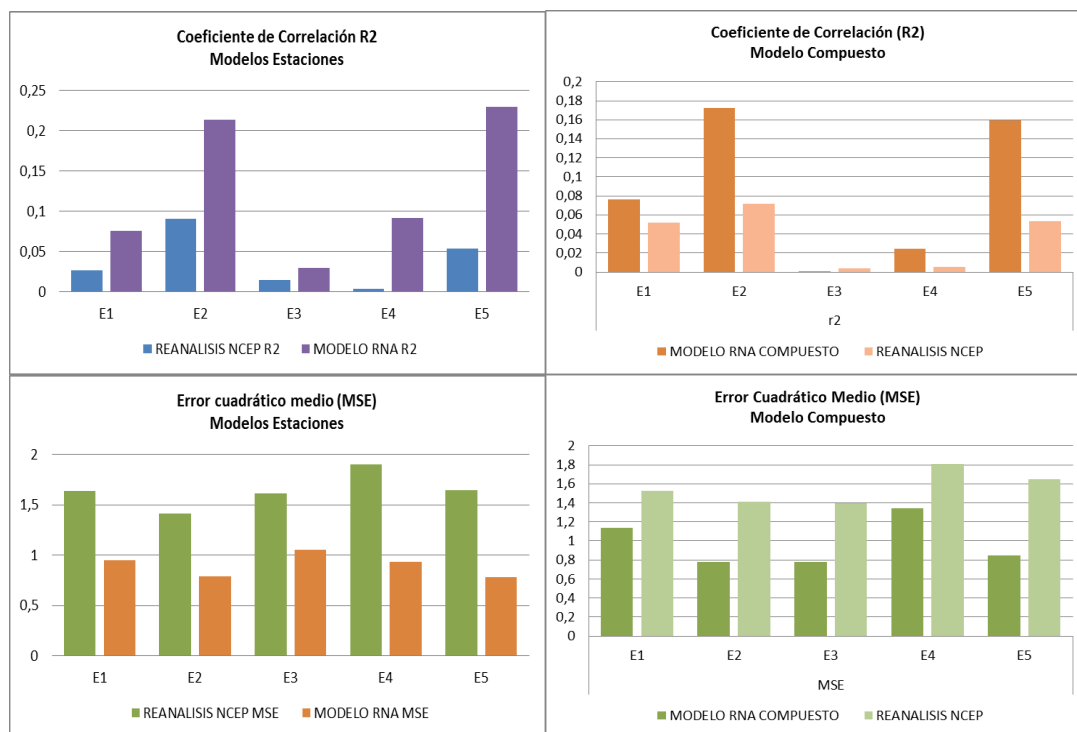
5.6. PRONÓSTICO DE LA TEMPERATURA MEDIA DIARIA

Se planteó utilizar los modelos desarrollados en este trabajo, haciendo un pronóstico de la temperatura media diaria local, para cada estación, en el periodo de tiempo

comprendido entre el 1 de enero de 2001 y el 31 de diciembre de 2100; bajo los escenarios supuestos A2 y A1B del IPCC.

5.6.1. Evaluación de variables de entrada para el Pronóstico

Las variables de cada conjunto de datos, que corresponden a las entradas de los modelos de RNA y a su vez son las variables de salida del reanálisis NCEP/NCAR, fueron correlacionadas con las variables de salida del CGCM 3.1 obtenidas para el escenario A2, determinándose la aproximación de la simulación del reanálisis. Los coeficientes de correlación (r^2) se presentan en la Figura 5-22. Se muestra un gráfico de barras (ver donde se ve claramente que las variables del reanálisis NCEP/NCAR de entrada al modelo compuesto están más correlacionadas con su respectiva variable del CGCM que las variables de entrada de los modelos por estación. Un mayor detalle de los coeficientes de correlación (r y r^2) y el nombre de las variables se presenta en el Anexo 12.



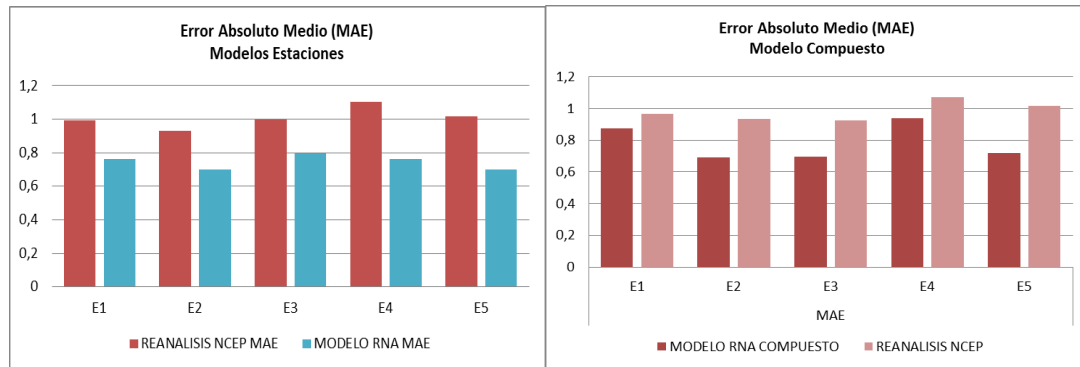


FIGURA 5-21: COMPARACIÓN DESEMPEÑO MODELOS INDIVIDUALES Y COMPUESTO

Se observó que para el modelo RNA de la estación E1, las variables de entrada correspondientes al CONJUNTO 3 derivadas del reanálisis NCEP/NCAR que se encuentran en la franja 24Y, no son simuladas por el CGCM debido al efecto coriolis mencionado en el capítulo 4.3.2. En el caso de esta estación solo se encontraron datos en el CGCM para la variable 500hPa Geopotencial y las 8 variables de entrada restantes no fueron simuladas por el CGCM (ver Tabla 23).

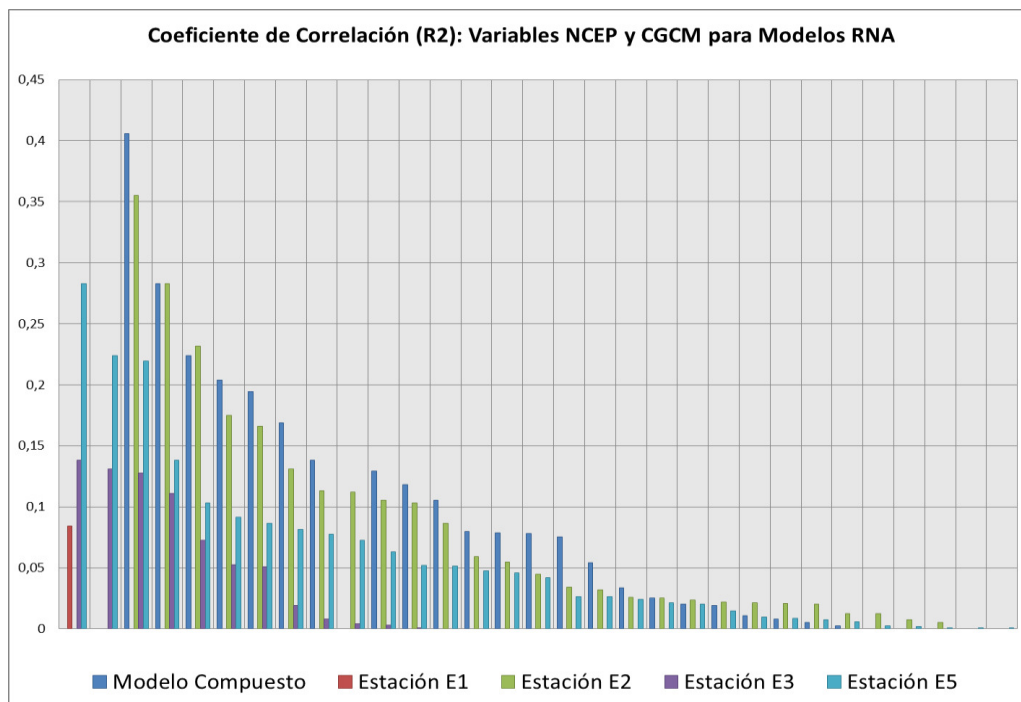


FIGURA 5-22: CORRELACIÓN ENTRE EL REANALISIS NCEP/NCAR Y EL CGCM3.1 PARA LAS VARIABLES ENTRADA

Para la estación E2, igualmente que en el caso 1, 2 de las variables de entrada del CONJUNTO 5, correspondientes a variables del reanálisis NCEP/NCAR ubicadas en la franja 24Y, no se encuentran simuladas en el CGCM. De la misma manera se encontraron 2 de las 25 variables de entrada al modelo compuesto que no fueron simuladas por el CGCM. En la estación E5 todas las variables de entrada se pudieron correlacionar con respectiva variable en el CGCM (Ver Anexo 12).

5.6.2. Generación del pronóstico las temperaturas medias diarias para las estaciones de estudio

Las variables mejor correlacionadas entre el reanálisis NCEP/NCAR y el CGCM son las de entrada al modelo compuesto, es decir, que las variables de entrada al modelo compuesto fueron mejor simuladas por el reanálisis y por tanto en el caso de pronóstico,

con las variables de salida del CGCM para los escenarios A2 y A1B, se espera tener mayor correlación debido a que los modelos fueron desarrollados en base a las variables de salida del reanálisis NCEP/NCAR.

Existen 2 limitaciones del modelo compuesto para el pronóstico: a) los datos de anomalías reportados por este modelo para la estación 3 no son aplicables, b) no hay datos para dos de las variables de entrada.

Para el caso a) los datos de salida para la estación E3 no serán tenidos en cuenta. En el caso b) se decidió rentrenar el modelo RNA con solo las 23 variables de entrada. Los datos de correlación y error medio en los grupos de entrenamiento, validación y prueba se presentan en la siguiente tabla:

TABLA 13: DATOS DE DESEMPEÑO MODELO COMPUESTO CON REDUCCIÓN DE VARIABLES

Performance Metrics										
No. Variables	Model Name	Training			Cross Validation			Testing		
		MSE	r	MAE	MSE	r	MAE	MSE	r	MAE
23	MLP-2-O-M (Multilayer Perceptron)	1,21507242	0,27792482	0,90142105	1,12920205	0,18679944	0,88854103	0,94118976	0,05332485	0,76960893
25	MLP-2-O-M (Multilayer Perceptron)	1,18142352	0,30096024	0,8873654	1,08485154	0,25742272	0,87354169	0,94721645	0,07339086	0,77066374

Se observa una pérdida leve de ajuste del modelo al reducir las dos variables de entrada no simuladas por el CGCM (correspondientes a variable 850hPa Componente V en dos celdas diferentes, ver Tabla 27).

Se utilizaron los datos simulados por el CGCM para las 23 variables de entrada para el periodo de pronóstico mencionado anteriormente, en los escenarios A2 y A1B. Con el modelo entrenado y las variables de entrada generadas por el CGCM para cada escenario, se generaron las salidas de anomalías para cada uno de las 5 estaciones meteorológicas de estudio.

En las figuras abajo se muestran los pronósticos de temperatura media diaria en el periodo comprendido entre el 1 de enero de 2001 hasta el 31 de diciembre de 2100, bajo dos escenarios supuestos A2 (azul) y A1B (rojo).

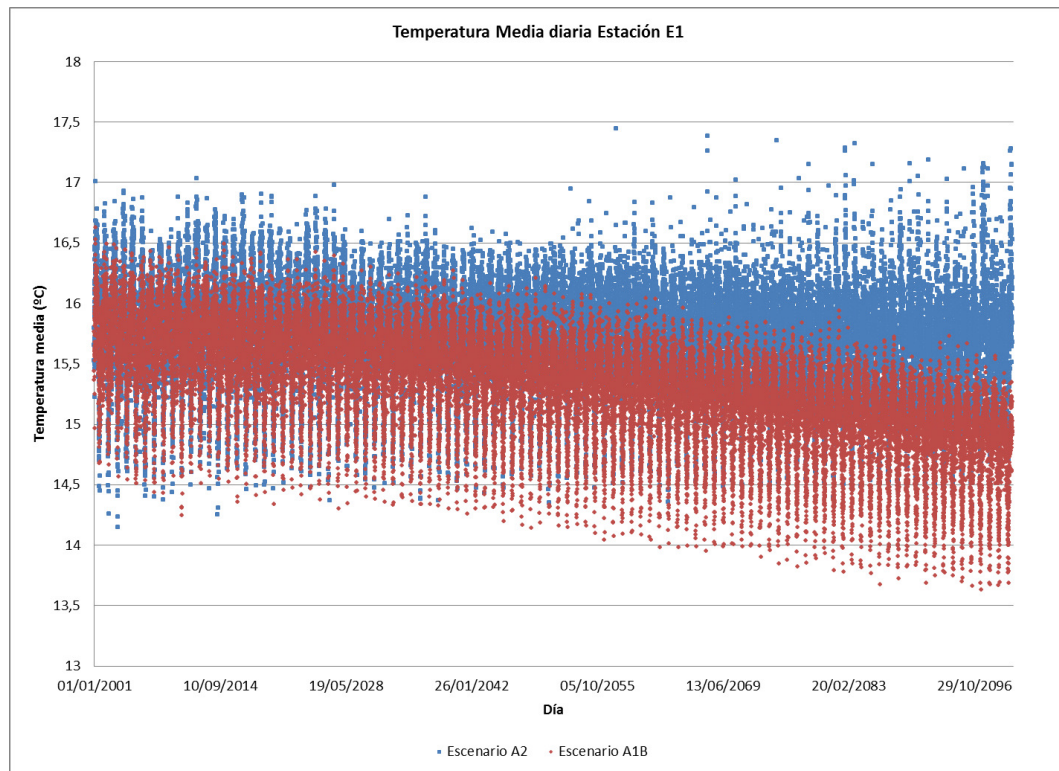


FIGURA 5-23: PRONÓSTICO DE TEMPERATURA MEDIA DIARIA ESTACIÓN E1

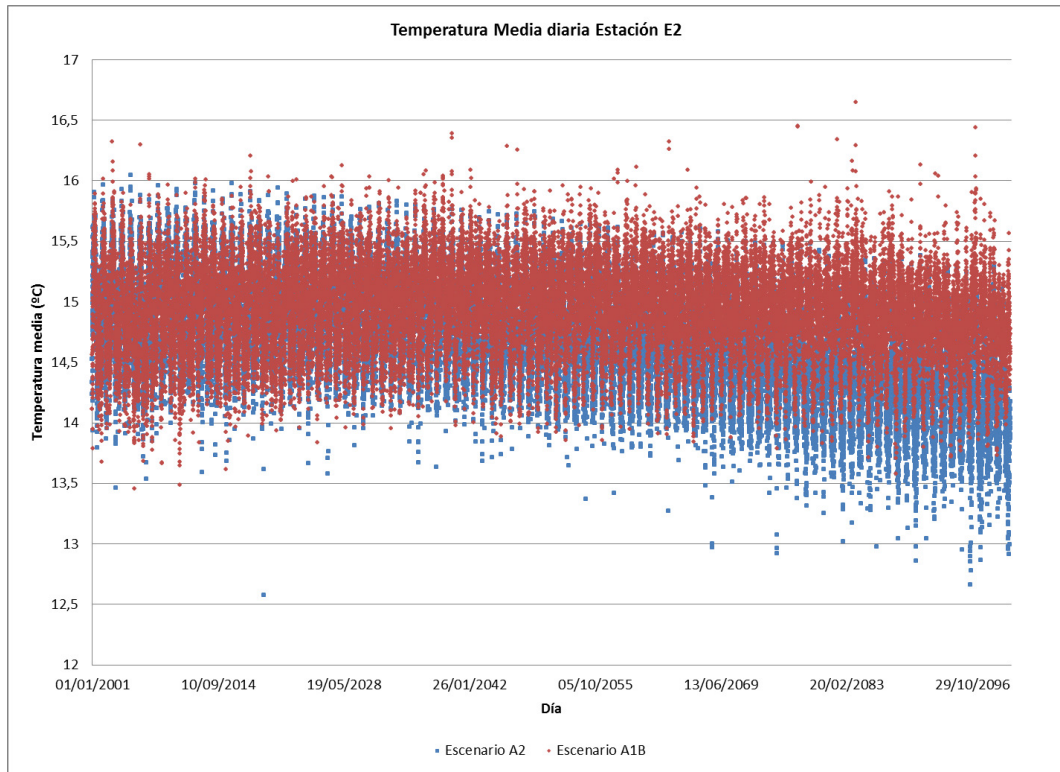


FIGURA 5-24: PRONÓSTICO DE TEMPERATURA MEDIA DIARIA ESTACIÓN E2

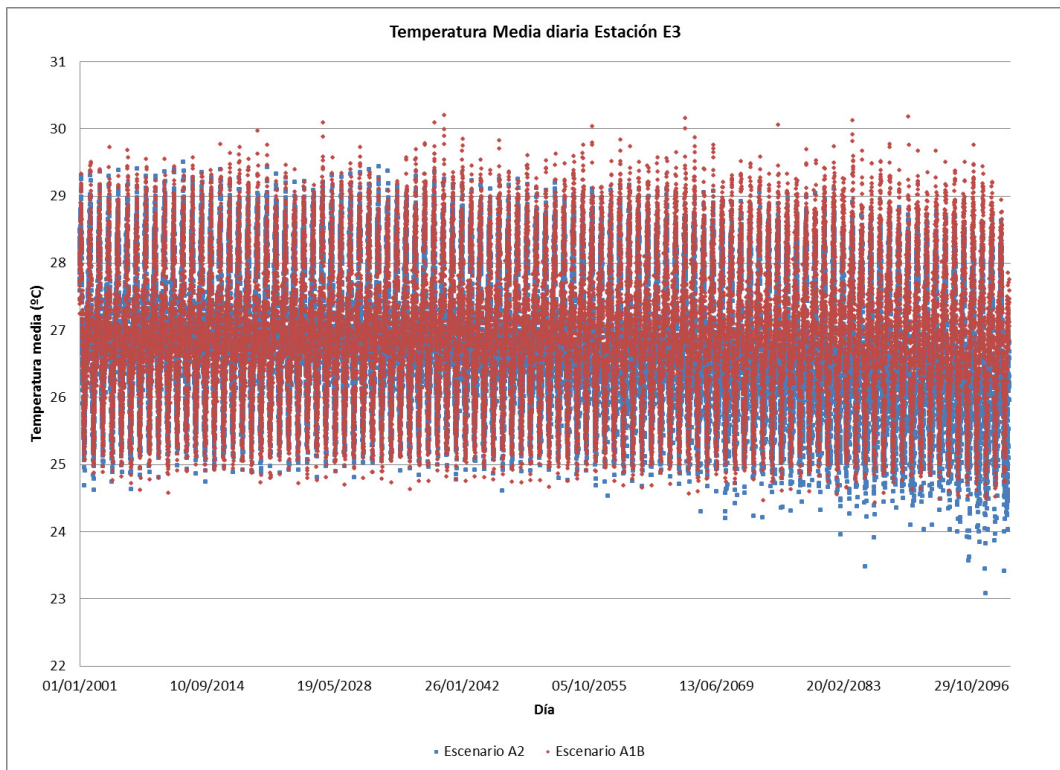


FIGURA 5-25: PRONÓSTICO DE TEMPERATURA MEDIA DIARIA ESTACIÓN E3

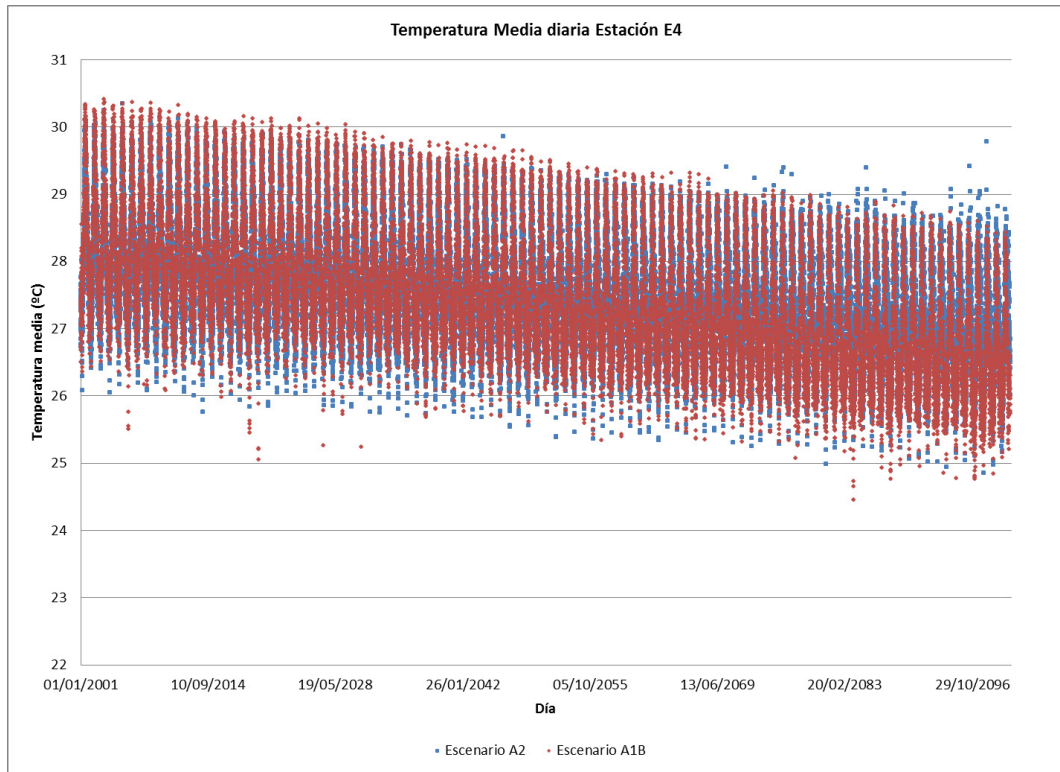


FIGURA 5-26: PRONÓSTICO DE TEMPERATURA MEDIA DIARIA ESTACIÓN E4

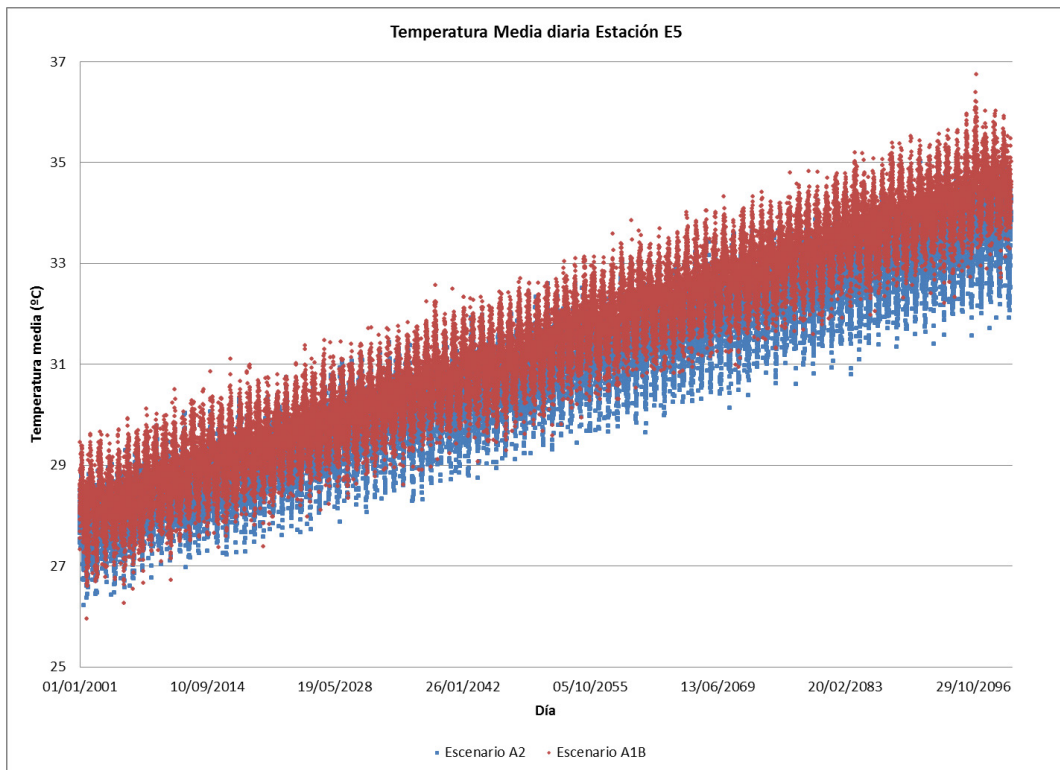


FIGURA 5-27: PRONÓSTICO DE TEMPERATURA MEDIA DIARIA ESTACIÓN E5

6. ANÁLISIS DE RESULTADOS

El trabajo aquí presentado es una herramienta para la reducción de escala de la variable temperatura media diaria del modelo de circulación global canadiense a condiciones de superficie. El desarrollo de esta herramienta contó con limitaciones fuertes principalmente en la discontinuidad y falta de datos históricos de las estaciones meteorológicas de estudio, para lo cual se hizo necesario desarrollar una metodología de reconstrucción de las series de tiempo de forma que se pudieran usar las variables de salida tanto del reanálisis NCEP/NCAR como del modelo climático CGCM. Pero evidentemente las estaciones con menor correlación (estaciones E3 $r=0,216$ y E4 $r=0,117$) o menor ajuste del modelo con la realidad, fueron aquellas donde la falta de datos en la serie fue mayor (32,5% y 27,9% respectivamente). Así mismo se encontraron las mejores correlaciones cuando el número de datos faltantes es menor al 20% (estaciones E2 $r=0,400$ y E5 $r=0,497$).

Diferentes autores (Buda & Jarynowski, 2010) (Cohen, 1998) han propuesto bases para la interpretación del coeficiente de correlación, sin embargo, esas guías son en cierta manera arbitrarias y en muchos casos no se deberían seguir estrictamente (Cohen, 1998). La interpretación del coeficiente de correlación depende del contexto y el propósito del trabajo. Es decir, si se intentara verificar una ley física usando instrumentación de alta calidad un coeficiente de 0,9 podría interpretarse como bajo pero en el estudio de ciencias sociales, donde hay una serie de factores no medibles que contribuyen al fenómeno, un resultado del mismo orden puede ser muy alto. Así mismo, para este trabajo, y aceptando que en muchas publicaciones se acuerda que coeficientes de correlación del orden de 0,1 a 0,5 indican una correlación débil entre la(s) variable(s) de entrada y salida; en este campo de estudio, climatológico, donde los fenómenos son altamente complejos y los factores que

intervienen son múltiples se acepta la Tabla 14: Interpretación Coeficientes de Correlación Tabla 14.

TABLA 14: INTERPRETACIÓN COEFICIENTES DE CORRELACIÓN (COHEN, 1998)

Correlation	Negativa	Positiva
Ninguna	-0,09 a 0,0	0,0 a 0,09
Baja	-0,3 a -0,1	0,1 a 0,3
Moderada	-0,5 a -0,3	0,3 a 0,5
Fuerte	-1,0 a -0,5	0,5 a 1,0

La relevancia de las RNA desarrolladas como herramienta para el pronóstico de la temperatura media diaria a escala superficial, radica en su capacidad para mejorar la correlación frente al reanálisis NCEP/NCAR como se muestra en la Figura 6-1. Se observó en todos los modelos propuestos en este trabajo una mejora en la correlación de las salidas con los datos reales de las estaciones meteorológicas superando en más del 40% la correlación actual que ofrecen el modelo climático CGCM3.1 y el Reanálisis NCEP/NCAR. Es evidente que debido a que la resolución espacial a la cual los procesos atmosféricos pueden ser simulados, en términos computacionales, aun es limitada a $3,75^\circ \times 3,75^\circ$; la información de la temperatura superficial de los modelos climáticos difiere de las realidades naturales.

Adicional se obtuvo una mejorara en la correlación pasando de un ajuste bajo del reanálisis a uno moderado en tres de los casos de modelos de RNA, como se muestra en la Figura 6-2. Para los casos de las estaciones E3 y E4 se mejoró el ajuste en un 78,37% y 185,46% respectivamente (tomando como base el coeficiente de correlación de Spearman que no se afecta por outliers) aunque no se alcanzaron correlaciones moderadas y son por tanto los modelos más débiles para la predicción de la temperatura media diaria superficial.

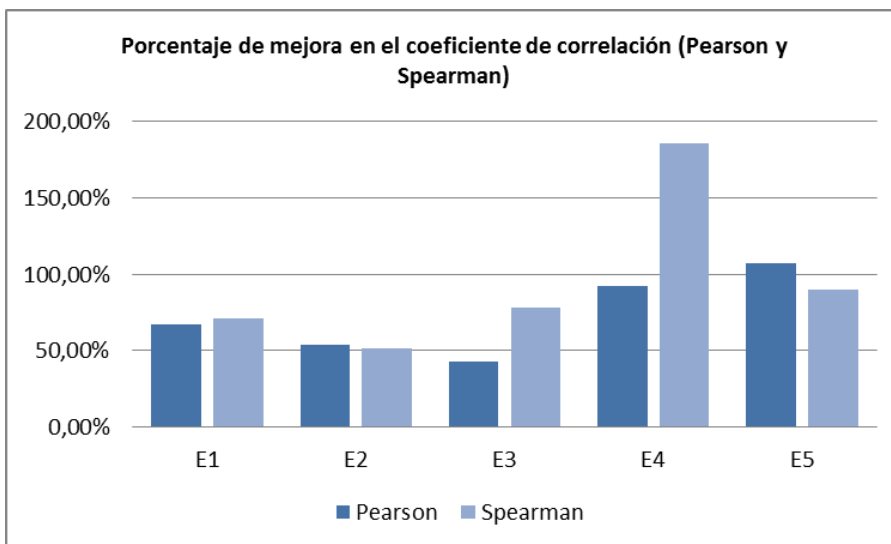


FIGURA 6-1: MEJORA EN LA CORRELACION DE LOS MODELOS RNA FRENTE AL REANÁLISIS

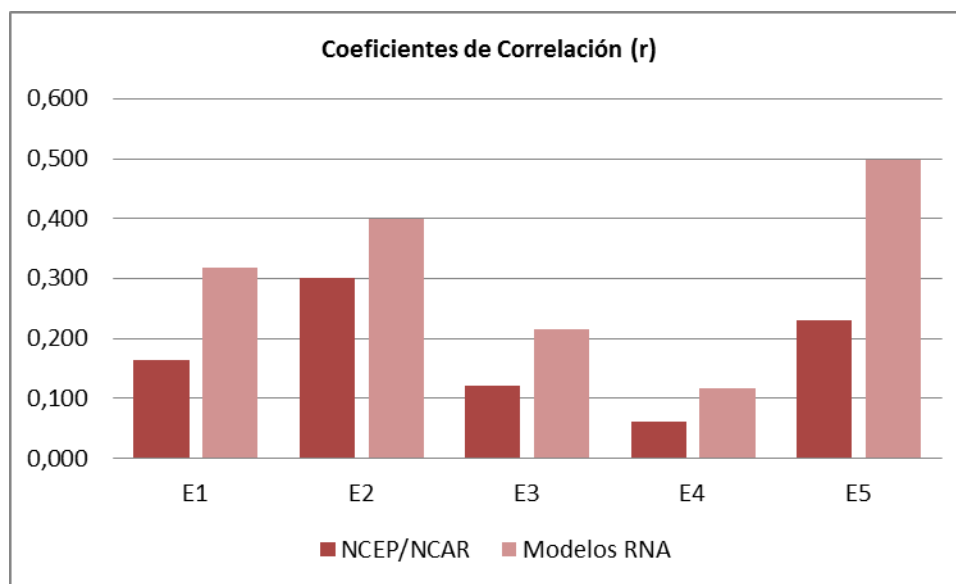


FIGURA 6-2: COEFICIENTES DE CORRELACIÓN (R) MODELOS RNA Y REANÁLISIS

Los modelos de RNA, tienen una relación directa con la capacidad de simulación del reanálisis. Se evidenció que estos modelos decaen en su capacidad de simular la realidad de las anomalías de la temperatura cuando la simulación del reanálisis es

deficiente. Es decir, en los casos en que la correlación de la variable No. 25 (Temperatura a 2 metros) del reanálisis contra las anomalías fue baja (Ej.: estaciones E3 y E4), en esos mismos casos el desempeño de los modelos RNA fue inferior al encontrado en otras estaciones. Caso contrario, para los modelos de mejor desempeño (Estaciones E2 y E5) también el reanálisis tuvo mayor correlación. La baja correlación del reanálisis frente a las anomalías de estudio puede deberse a la incertidumbre propia del modelo o a errores (bias) en la reconstrucción de las series de tiempo pues las estaciones con mayor porcentaje de datos faltantes fueron las que presentaron menor correlación. No obstante la estación E5, a la que le faltaba el 17,31% de los datos, logró un desempeño similar a la estación E2 (datos faltantes 4,87%) y mayor al de la estación E1 (datos faltantes 21,70%).

No se observaron características de la orografía que afecten la capacidad de los modelos RNA para simular los datos de anomalías. Los modelos de las estaciones E2 y E5, que tuvieron los mejores desempeños para la simulación de datos de anomalías de temperatura media diaria, se encuentran en dos zonas climáticas con diferencias fuertes en cuanto a altura sobre el nivel del mar y localización geográfica en el país (interna y costera respectivamente).

Adicional a la comparación del desempeño de las RNA frente a los datos del Reanálisis NCEP/NCAR se implementaron modelos de regresión lineal simple y múltiple para correlacionar los datos reales con los datos de salida del reanálisis. Esta metodología se usó como estándar de comparación para definir si los modelos RNA presentaban mejora. Los resultados mostraron que en ninguno de los casos la metodología de regresión lineal supera los mejores modelos de red neuronal e inclusive se observó que una RNA que utilizó como variables de entrada, las mismas variables desprendidas de la regresión Stepwise, supera en desempeño esa metodología. Como se menciona en el capítulo 2.1 los

procesos físicos que describen el clima son altamente no lineales por lo que se demuestra que una metodología de redes neuronales es más adecuada para la modelación o reducción de escala de variables climáticas.

En este trabajo, para el desarrollo de las arquitecturas de redes neuronales, se utilizó como función de transferencia la tangente hiperbólica que junto con la función logística (sigmoïdal) son de uso frecuente. La razón principales para el uso de esa función de transferencia es que se trabajó con las anomalías de la temperatura, una variable continua que puede ser negativa o positiva y que fue normalizadas en cero y estandarizadas a 1, lo que permite estar dentro del rango de la función de la tangente hiperbólica (-1,1). Posteriormente se evaluó el impacto del desempeño de la RNA al cambiar la función de transferencia por una sigmoïdal en las capas oculta y de salida. Se encontró que el uso de una función de transferencia sigmoïdal en la capa de salida ocasionaba un descenso en la correlación de la salida del modelo. La combinación de funciones sigmoïdales en capa de salida y oculta no mostró ninguna mejora y solo en las estaciones 1 y 5 se encontró que el uso de la función sigmoïdal en la capa oculta mejoraba el desempeño de la red. En ninguno de los casos se observó una reducción de la velocidad de entrenamiento por el uso de funciones sigmoïdales en la capa oculta como se reporta por Hsieh (Hsieh, 2009)

Para cada una de las estaciones de estudio se evaluaron 100 modelos: 1 regresión lineal simple, 2 regresiones lineales múltiples, 2 regresiones Stepwise y 95 modelos de red neuronal con diversas variables de entrada, algoritmos de aprendizaje y métodos de presentación de datos de datos de retropropagación. Los 100 modelos fueron organizados de manera descendente, en el grupo de validación, desde el mayor hasta el menor coeficiente de correlación. Se evaluaron los 10 mejores modelos encontrándose que las arquitecturas de tipo Perceptrón Multicapa (MLP) con una o dos capas ocultas, las redes

dinámicas con sistemas de memoria gamma (TLRN), las redes recurrentes (RN) y Generalized Feed Forward (GFF) son las arquitecturas más apropiadas para captar la complejidad de los datos de anomalías de temperatura. Así mismo se evidenció que las redes con función de base radial (RBF) y las dinámicas con memoria TDNN tuvieron el menor desempeño al compararse con las demás arquitecturas. Basado en estos hallazgos se podría pensar en profundizar en el tipo de conexiones recurrentes entre elementos de procesamiento y cómo las diferentes profundidades de memoria para cada entrada afectan el desempeño del MLP para lograr hacer mejores optimizaciones de los modelos aquí propuestos.

Los modelos RNA seleccionados para la simulación y pronóstico de las anomalías de temperatura en las estaciones E2 y E5 mostraron el mejor desempeño. Sus similitudes están en su arquitectura (MLP), algoritmo de aprendizaje (Momentum), método de presentación de datos (On-line) y conjunto de entrada de datos correspondiente a variables de las celdas circundantes seleccionadas por la técnica de eliminación/incorporación progresiva (Stepwise). En cuanto a sus diferencias, en el caso de la estación E5 el MLP cuenta con 2 capas ocultas y su entrenamiento se hizo con 11 años menos de datos (4.015 datos menos).

El modelo compuesto de mejor desempeño se logró con una arquitectura perceptrón multicapa con 2 capas ocultas al igual que en el mejor modelo de la estación 5. Estos dos modelos se construyeron con solo 5.110 datos (otras estaciones hasta con 9.125 datos), debido a la disponibilidad..

Según las variables de entrada de los modelos RNA individuales la estación E1 ubicada a 2.085 m.s.n.m. en un clima de páramo húmedo las anomalías están mayormente correlacionadas con el flujo y la dirección del viento. En la estación E2 ubicada 2.580

m.s.n.m. en un clima frío y lluvioso las anomalías están más relacionadas con la humedad específica a 1000, 850 y 500 hPa, la dirección del viento y la componente V del viento. Para la estación E3 ubicada a 150 m.s.n.m. en un clima cálido y húmedo las variables de entrada fueron en general relacionadas con el viento y vorticidad. Esta estación se encuentra afectada por la ZCIT por lo cual la dinámica climática varía haciendo más difícil la simulación de las variables climáticas en los modelos globales. La estación E5 las anomalías están influenciadas por la humedad específica, la presión sobre el nivel medio del mar, las componentes U y V del viento y la velocidad del viento. De aquí se observa que para las zonas bajas y cálidas las anomalías están influenciadas por la presión media del mar que no aparece en las zonas altas. En general zonas altas y bajas están afectadas por la humedad específica y componentes del viento. Los modelos de mejor desempeño, estaciones E2 y E5 muestran influencia de las celdas circundantes. La estación E2 mayoritariamente de las variables provenientes del nororiente y la estación E5 del sur. Las variables de la franja Y=24 en general son poco usadas por los modelos RNA, esto puede deberse a la dificultad para la simulación en esta zona debido al efecto Coriolis.

En el modelo compuesto en vez de una técnica de eliminación/incorporación progresiva se efectuó un análisis de sensibilidad donde se encontró que las anomalías son sensibles a la temperatura superficial a 2m provenientes de la franja norte Y=21, la componente V del viento de la zona sur (Y=24), la variable geopotencial a 500hPa en toda la región, la humedad específica en todas las alturas y dirección del viento. No mostró sensibilidad frente a la divergencia o vorticidad y muy poca relación con la velocidad del viento o componente U.

Entendiendo las limitaciones, alcances y mejoras que los modelos RNA obtuvieron, se realizaron los pronósticos de la temperatura media diaria en los 5 puntos de estudio. Se

deben entender como pronósticos posibles bajo unos ciertos escenarios eventuales y estas anomalías tendrán una posible correlación moderada en las estaciones E2, E5 y E1 y una baja en la estación E3 según lo que se reportó en los resultados. Se encontraron limitaciones para el pronóstico de temperatura bajo los escenarios supuestos A1B y A2 del IPCC debido a la baja capacidad del reanálisis de representar las variables climáticas del modelo CGCM3.1. Las variables de entrada a los mejores modelos individuales por estación tienen menor correlación con sus respectivas variables en el CGCM3.1 que el modelo compuesto propuesto. Fue necesario pronosticar con el modelo compuesto el cual, aunque tiene menor desempeño en la simulación de anomalías de temperatura en las 5 estaciones meteorológicas de estudio, sus variables de entrada están mejor simuladas por el reanálisis que fue la herramienta de construcción del modelo. Es decir, se decidió darle peso a las limitaciones que las variables de entrada representaban para los modelos RNA estudiados. Esto permite pensar que el pronóstico, que se hace a partir de las variables de salida del CGCM3.1 para los escenarios futuros A1B y A2, pierde menos información en el modelo compuesto y por tanto los pronósticos son más confiables.

Se encontró una diferencia para la estación E1 en que mientras en el escenario A2 la temperatura hacia futuro se mantiene relativamente constante, bajo el escenario A1B a finales de siglo se comienza a notar un descenso de la temperatura media diaria. En la estación E2, también en altura, se observa que ambos escenarios muestran un descenso en la temperatura para finales de siglo, particularmente bajo el escenario A2 mientras que en el A1B se nota un incremento inicial que luego vuelve a caer. En la estación E3 se nota una gran variabilidad de la temperatura haciendo difícil concluir sobre la tendencia de los datos, aunque un descenso tendencial se observa bajo el escenario A1B a finales de siglo. Para la estación E4 la tendencia es un descenso de la temperatura de casi 1,5°C en el periodo de

100 años bajo los dos escenarios. En la estación E5 la tendencia es a un incremento pronunciado de 4°C de la temperatura en los próximos 100 años en cualquiera de los dos escenarios.

Los pronósticos de las estaciones meteorológicas E1, E2 y E4 bajo el escenario A1B muestran un descenso en la temperatura que puede ser explicado por el decrecimiento de la población mundial y el desarrollo de energías alternativas que plantea este escenario y que no es captado por la estación E5. Bajo el escenario A2 se esperaría un calentamiento sostenido o estable que está mejor pronosticado por las estaciones E1 y E5. El pronóstico de temperatura media diaria de los modelos aquí propuestos está fuertemente influenciado por la componente tendencial de la serie de tiempo las cuales fueron prácticamente estables (pendiente aprox. 0) para las estaciones E2, E3 y E4, negativa para la estación E1 y positiva para la estación E5.

7. CONCLUSIONES Y PERSPECTIVAS

En los anteriores capítulos de este trabajo se desarrollaron modelos de redes neuronales artificiales capaces de mejorar la correlación con las observaciones frente a las simulaciones del reanálisis NCEP/NCAR para la temperatura media diaria superficial. Los modelos RNA propuestos en este trabajo mostraron mejor capacidad tanto para captar la complejidad de las anomalías de las temperaturas medias diarias como la habilidad para simular datos más confiables de la realidad. Aunque persisten muchas limitaciones en los modelos RNA (Ej.: entrenamiento basado en series de tiempo reconstruidas, variables de entrada dependientes de la capacidad de simulación del reanálisis, baja capacidad del reanálisis para simular las variables del CGCM3.1 y bajas correlaciones de las simulaciones RNA con la realidad local) se mostró una mejora versus las herramientas actuales

(CGCM3.1 y reanálisis NCEP/NCAR) lo que permite hacer pronósticos de escenarios futuros con mayor confiabilidad.

Los modelos RNA mostraron habilidad para aprender las características no lineales y complejas que rigen las anomalías de la temperatura en 5 estaciones meteorológicas superando así modelos de regresión lineal simples o múltiples. De esta forma los modelos propuestos cierran la brecha de resolución entre los modelos globales y regionales a modelos de superficie. De todas formas, esta técnica de interpolación no logra compensar la pérdida del reanálisis de la variabilidad del clima. Para superar estas diferencias se debe contar con un Reanálisis de mejor resolución espacial, dado que la correlación entre las temperaturas del Reanálisis y las observaciones del IDEAM es pobre. Igualmente se presume que los modelos propuestos están afectados por la poca calidad de los datos observados (con muchos periodos discontinuos).

Se estudiaron 5 estaciones meteorológicas colombianas ubicadas en diferentes celdas del reanálisis NCEP/NCAR. Se evaluó la correlación y el error de las simulaciones del reanálisis NCEP/NCAR, de modelos de correlación lineal (simple, múltiple y Stepwise) y de modelos RNA contra las anomalías de la temperatura media diaria reportadas para cada una de las estaciones por el IDEAM. Se encontró, en todos los casos, modelos de RNA con mayor correlación y menor error con los datos reales, con lo que se demuestra que esta técnica de reducción de escala es apropiada para mejorar la simulación de la celda del reanálisis independientemente de la posición.

Las 10 mejores arquitecturas de cada modelo fueron evaluadas en cuanto a sus características encontrándose que los perceptrones multicapa son las arquitecturas más apropiadas para el aprendizaje de esta variable climática. Así mismo la presentación de

datos On-Line y algoritmo de aprendizaje Momentum mostraron mayores velocidades de aprendizaje y mejor ajuste de los modelos.

Se desarrolló un modelo compuesto basado en los resultados de los modelos individuales por estación con resultados de correlación similares a los obtenidos por separado y con la ventaja de ser una sola arquitectura.

Se realizó un pronóstico de la temperatura media diaria superficial en 5 puntos específicos de geografía colombiana, correspondientes a sus diferentes zonas climáticas, para el periodo para el periodo 2010 a 2100 bajo dos escenarios supuestos encontrándose coherencia con la descripción de los escenarios según el IPCC.

Se determinó que la construcción de modelos individuales es más apropiada que un modelo compuesto para la simulación de datos de anomalías. Los modelos individuales por estación presentaron mayores correlaciones y menor error frente a los datos reales.

Basado en las observaciones anteriores, parece que el desempeño de los modelos de reducción de escala propuestos está más relacionado con la capacidad del reanálisis para simular las variables climáticas de la región que con la orografía de la zona o la cantidad de datos de entrenamiento. Claramente existe un número crítico de datos requerido para que la red alcance a aprender las características de los datos.

La metodología y modelos propuestos son susceptibles de optimización y mejora a través del estudio de nuevas arquitecturas, técnicas más sofisticadas (Ej.: SupportVector Machines ó máquinas de soporte vectorial (Zeng, Hsieh et al. 2011) o algoritmos genéticos de optimización. También son susceptibles de mejora las metodologías de adecuación de memoria que mostraron ser más eficientes en cuanto su capacidad de entrenamiento con menor número de variables. Este trabajo muestra la habilidad de un modelo RNA acoplado a un modelo climático global para la reducción de escala de variables climáticas y abre una

ventana a la investigación en el campo para optimizar los modelos y lograr modelos de superficie bajo condiciones climáticas y geográficas colombianas, donde aún es un campo inexplorado.

Estudios encaminados a mejorar y optimizar los modelos propuestos pueden enfocarse a trabajar con periodos de estudio mayores (Ej: temperaturas mensuales), preprocesamiento de datos (Ej: escalar variables a magnitudes equivalentes), neuronas con memorias (Ej: Gamma o de retraso) o optimización por otros tipos de error. En todo caso en este trabajo se propuso la minimización y evaluación del desempeño de la red con el error absoluto medio según las recomendaciones encontradas en la bibliografía. (Willmott y otros, 2005).

El trabajo aquí presentado demostró la capacidad de las RNA para la reducción de escala de variables climáticas con capacidad de estrechar la brecha en resolución entre los modelos globales y de superficie con una demanda computacional muy inferior. Esto abre campo a futuras investigaciones en campos como el estudio de la variación de los pisos térmicos, afectación de fauna y flora en las zonas climáticas colombianas, deshielo de picos y nevados, afectación de cosechas y cultivos entre otras. Una de herramienta de este tipo es susceptible de convertirse en base para el diseño de políticas gubernamentales, elaboración del CONPES o como modelo para la conservación y preservación del medio ambiente por parte de entidades como el IDEAM.

Se desarrolló un modelo que demuestra la habilidad de las RNA y se abre una puerta a futuras investigaciones para mejorar los ajustes con la realidad a través de búsqueda de nuevas arquitecturas, métodos de aprendizaje, funciones de transferencia, capas de memorias para creación de redes dinámicas, entre otras.

BIBLIOGRAFÍA

1. Abdel-Aal, R. E. (2004). "Hourly temperature forecasting using abductive networks." Engineering Applications of Artificial Intelligence **17**(5): 543-556.
2. Aksornsingchai, P. and C. Srinilta (2011). Statistical downscaling for rainfall and temperature prediction in Thailand.
3. Aminian, A. (2010). "Prediction of temperature elevation for seawater in multi-stage flash desalination plants using radial basis function neural network." Chemical Engineering Journal **162**(2): 552-556.
4. Anderson, J. A. and E. Rosenfeld (1987). Neurocomputing. Cambridge, MIT press.
5. Balaghi, R., B. Tychon, H. Eerens and M. Jlibene (2008). "Empirical regression models using NDVI, rainfall and temperature data for the early prediction of wheat grain yields in Morocco." International Journal of Applied Earth Observation and Geoinformation **10**(4): 438-452.
6. Białobrzewski, I. (2008). "Neural modeling of relative air humidity." Computers and Electronics in Agriculture **60**(1): 1-7.
7. Bishop, C. M. (1995). Neural Networks for Pattern Recognition. New York, Oxford University Press Inc.
8. Bradley, J. B. (1995). "Neural networks: A comprehensive foundation: S. HAYKIN. New York: Macmillan College (IEEE Press Book) (1994). v + 696 pp. ISBN 0-02-352761-7." Information Processing & Management **31**(5): 786.
9. Bronstert, A., D. Niehoff and G. Burger (2002). "Effects of climate and land-use change on storm runoff generation: present knowledge and modelling capabilities." Hydrological Processes **16**(2): 509-529.
10. Cannon, A. J. (2011). "Quantile regression neural networks: Implementation in R and application to precipitation downscaling." Computers and Geosciences **37**(9): 1277-1284.
11. Cao, Q., B. T. Ewing and M. A. Thompson (2012). "Forecasting wind speed with recurrent neural networks." European Journal of Operational Research **221**(1): 148-154.
12. Cavazos, T. (1999). "Large-scale circulation anomalies conducive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas." Journal of Climate **12**(5 II): 1506-1523.
13. Chattopadhyay, S., D. Jhajharia and G. Chattopadhyay (2011). "Univariate modelling of monthly maximum temperature time series over northeast India: Neural network versus Yule-Walker equation based approach." Meteorological Applications **18**(1): 70-82.
14. Chronopoulos, K., A. Kamoutsis, A. Matsoukis and E. Manoli (2012). "An artificial neural network model application for the estimation of thermal comfort conditions in mountainous regions, Greece." Atmosfera **25**(2): 171-181.
15. Compostela, B. (2003). Análisis de Regresión Lineal. Madrid, Universidad Complutense de Madrid.
16. Coulibaly, P. (2010). "Reservoir Computing approach to Great Lakes water level forecasting." Journal of Hydrology **381**(1-2): 76-88.
17. Coulibaly, P. and N. D. Evora (2007). "Comparison of neural network methods for infilling missing daily weather records." Journal of Hydrology **341**(1-2): 27-41.
18. Crane, R. G. and B. C. Hewitson (1998). "Doubled CO₂ precipitation changes for the Susquehanna basin: Down-scaling from the GENESIS general circulation model." International Journal of Climatology **18**(1): 65-76.
19. Daly, C., R. P. Neilson and D. L. Phillips (1994). "A statistical-topographic model for mapping climatological precipitation over mountainous terrain." Journal of Applied Meteorology **33**(2): 140-158.
20. de Bruin, H. A. R., I. F. Trigo, M. A. Jitan, N. T. Enku, C. van der Tol and A. S. M. Gieske (2010). "Reference crop evapotranspiration derived from geo-stationary satellite imagery: a

- case study for the Fogera flood plain, NW-Ethiopia and the Jordan Valley, Jordan." Hydrology and Earth System Sciences **14**(11): 2219-2228.
21. De, S. S., B. K. De, G. Chattopadhyay, S. Paul, D. K. Haldar and D. K. Chakrabarty (2011). "Identification of the best architecture of a multilayer perceptron in modeling daily total ozone concentration over Kolkata, India." Acta Geophysica **59**(2): 361-376.
 22. de Vries, B. and J. C. Principe (1990). A theory for neural networks with time delays. Proceedings of the 1990 conference on Advances in neural information processing systems 3, Denve, Colorado, United States, Morgan Kaufmann Publishers Inc. .
 23. DeVisser, M. H., J. P. Messina, N. J. Moore, D. P. Lusch and J. Maitima (2010). "A dynamic species distribution model of *Glossina* subgenus *Morsitans*: The identification of tsetse reservoirs and refugia." Ecosphere **1**(1): art6.
 24. Do Hoai, N., K. Udo and A. Mano (2011). "Downscaling global weather forecast outputs using ANN for flood prediction." Journal of Applied Mathematics **2011**.
 25. Doherty, S. J., S. Bojinski, A. Henderson-Sellers, K. Noone, D. Goodrich, N. L. Bindoff, J. A. Church, K. A. Hibbard, T. R. Karl, L. Kajefez-Bogataj, A. H. Lynch, D. E. Parker, I. C. Prentice, V. Ramaswamy, R. W. Saunders, M. S. Smith, K. Steffen, T. F. Stocker, P. W. Thorne, K. E. Trenberth, M. M. Verstraete and F. W. Zwiers (2009). "LESSONS LEARNED FROM IPCC AR4 Scientific Developments Needed To Understand, Predict, And Respond To Climate Change." Bulletin of the American Meteorological Society **90**(4): 497-+.
 26. Eberhart, R. C. and Y. Shi (2007). chapter two - Computational intelligence. Computational Intelligence. Burlington, Morgan Kaufmann: 17-38.
 27. Elman, J. L. (1990). "Finding structure in time." Cognitive Science: 552-558.
 28. Eslava, J. (1993). Apuntes de climatología y diversidad climática. Bogotá, Universidad Nacional de Colombia.
 29. Estrada-Pena, A., I. G. Horak and T. Petney (2008). "Climate changes and suitability for the ticks *Amblyomma hebraeum* and *Amblyomma variegatum* (Ixodidae) in Zimbabwe (1974-1999)." Veterinary Parasitology **151**(2-4): 256-267.
 30. Fahlman, S. E. (1988). Faster-Learnig variations on back-propagation: An empirical study. Proceedings of the 1988 Connectionist Models Summer School, San Mateo, Morgan Kaufmann.
 31. Flato, G. M. and G. J. Boer (2001). "Warming asymmetry in climate change simulations." Geophysical Research Letters **28**(1): 195-198.
 32. Freeman, J. A. and D. M. Skapura (1991). Neural Networks. Reading MA, Addison-Wesley.
 33. Friedel, M. J. (2012). "Data-driven modeling of surface temperature anomaly and solar activity trends." Environmental Modelling and Software **37**: 217-232.
 34. Fruh, B., J. W. Schipper, A. Pfeiffer and V. Wirth (2006). "A pragmatic approach for downscaling precipitation in alpine-scale complex terrain." Meteorologische Zeitschrift **15**(6): 631-646.
 35. Funk, C., J. Michaelsen, J. Verdin, G. Artan, G. Husak, G. Senay, H. Gadain and T. Magadazire (2003). "The collaborative historical African rainfall model: Description and evaluation." International Journal of Climatology **23**(1): 47-66.
 36. Gaitan Ospina, C. F. (2009). "Determinación de cambios en la variabilidad climática bajo diferentes escenarios de cambio climático. Caso de estudio: Ensenada de Alberni Robertson, Isla de Vancouver." Boletín Científico CIOH No. 27: 57-65.
 37. Giorgi, F. and L. O. Mearns (1991). "Approaches to the simulation of regional climate change: A review." Reviews of Geophysics **29**(2): 191-216.
 38. Giorgi, F. and L. O. Mearns (2003). "Probability of regional climate change calculated using the reliability ensemble averaging (REA) method." Geophys. Res. Lett **30**(12): 1629-1632.

39. Goyal, M. K., D. H. Burn and C. S. P. Ojha (2012). "Evaluation of machine learning tools as a statistical downscaling tool: Temperatures projections for multi-stations for Thames River Basin, Canada." Theoretical and Applied Climatology **108**(3-4): 519-534.
40. Guresen, E. and G. Kayakutlu (2011). "Definition of artificial neural networks with comparison to other networks." Procedia Computer Science **3**(0): 426-433.
41. Gurney, K. (2003). An Introduction to Neural Networks, CRC Press.
42. Hansen, J. E. and S. Lebedeff (1987). "Global trends of measured surface air temperature." J. Geophys. Res.: 13345-13372.
43. Hartell, C. E. J. and G. I. Pearman (2010). "Understanding and responding to the climate change issue: Towards a whole-of-science research agenda." Journal of Management & Organization **16**(1): 16-47.
44. Haykin, S. (1999). Neural Networks: A comprehensive foundation., Prentice Hall.
45. Hewitson, B. C. and R. G. Crane (1996). "Climate downscaling: Techniques and application." Climate Research **7**(2): 85-95.
46. Hidrográficas, C. d. I. O. e. (2012). "Climatología Del Caribe || CIOH.ORG - OCEANOGRAF?A OPERACIONAL." from <http://www.cioh.org.co/meteorologia/Climatologia/ClimatologiaCaribe3.php>.
47. Hung, N. Q., M. S. Babel, S. Weesakul and N. K. Tripathi (2009). "An artificial neural network model for rainfall forecasting in Bangkok, Thailand." Hydrology and Earth System Sciences **13**(8): 1413-1425.
48. Huth, R., Kyselý, Jan, Dubrovský and Martin (2003). "Simulation of Surface Air Temperature by GCMs, Statistical Downscaling and Weather Generator: Higher-Order Statistical Moments." Studia Geophysica et Geodaetica **47**(1): 203-216.
49. Jones, R. G., M. Noguer, D. C. Hassell, D. Hudson, S. S. Wilson, G. J. Jenkins and J. F. B. M. O. H. C. E. U. K. Mitchell (2004). Generating high resolution climate change scenarios using PRECIS. R. Taylor: 1-40.
50. Jordan, M. I. (1997). Chapter 25 Serial order: A parallel distributed processing approach. Advances in Psychology. W. D. a. V. P. D. John, North-Holland. **Volume 121**: 471-495.
51. Joubert, A. M. and P. D. Tyson (1996). "Equilibrium and fully coupled GCM simulations of future Southern African climates." South African Journal of Science 471-484.
52. Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzak, W. Higgins, J. Janowiak, K. C. Mo and C. Ropelewski (1996). "The NCAR/NCEP 40 year reanalysis project." Bulletin of the American Meteorological Society: 437-471.
53. Khan, M. S., P. Coulibaly and Y. Dibike (2006). "Uncertainty analysis of statistical downscaling methods." Journal of Hydrology **319**(1-4): 357-382.
54. Khashei, M. and M. Bijari (2012). "A new class of hybrid models for time series forecasting." Expert Systems with Applications **39**(4): 4344-4357.
55. Kim, J. W., J. T. Chang, N. L. Baker, D. S. Wilks and W. L. Gates (1984). "The statistical problem of climate inversion: determination of the relationship between local and large-scale climate." Monthly Weather Review **112**(10): 2069-2077.
56. Kohonen, T. (1988). Self-Organization and Associative Memory. New York, Springer-Verlag.
57. Kottek, M., J. Grieser, C. Beck, B. Rudolf and F. Rubel (2006). "World Map of the Köppen-Geiger climate classification updated." Meteorol. Z. **15**: 259-263.
58. Krasnopolsky, V. M. (2007). "Neural network emulations for complex multidimensional geophysical mappings: Applications of neural network techniques to atmospheric and oceanic satellite retrievals and numerical modeling." Reviews of Geophysics **45**(3).
59. Lapedes, A. S. and R. M. Farber (1987). "Nonlinear signal processing using neural networks: prediction and system modeling." Los Alamos Report: LA-UR 87-2662.

60. Lecun, Y., P. Y. Simard and B. Pearlmutter (1992). "Automati Learning Rate Maximization by On-line Estimation of the Hessian's Eigenvectors." Advances in Neural information Processing Systems: 156-163.
61. Li, G. and J. Shi (2010). "On comparing three artificial neural networks for wind speed forecasting." Applied Energy **87**(7): 2313-2320.
62. Lippmann, R. P. (1987). "INTRODUCTION TO COMPUTING WITH NEURAL NETS." IEEE ASSP magazine **4**(2): 4-22.
63. Lippmann, R. P. (1989). "Pattern classification using neurla networks." IEEE Communication Magazine: 47-64.
64. Lozán, J. L., H. Grassl and P. Hupfer (2001). Climate of the 21st century: changes and risks : scientific facts, Wissenschaftliche Auswertungen.
65. Makhoul, J. (1991). Pattern recognition properties of neural networks.
66. Manabe, Syukuro and T. Richard (1967). "Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity." Journal of Atmospheric Sciences **24**(3): 241-259.
67. Maurer, E. P. and H. G. Hidalgo (2008). "Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods." Hydrology and Earth System Sciences **12**(2): 551-563.
68. Mesa, O., G. Poveda and L. Carvajal (1997). Introducción al clima de Colombia. Medellín, Universidad Nacional de Colombia.
69. Miksovsky, J. and A. Raidl (2005). "Testing the performance of three nonlinear methods of time series analysis for prediction and downscaling of European daily temperatures." Nonlinear Processes in Geophysics **12**(6): 979-991.
70. Moore, N. and J. Messina (2010). "A Landscape and Climate Data Logistic Model of Tsetse Distribution in Kenya." Plos One **5**(7).
71. Muller, B., J. Reinhardt and M. T. Strickland (1995). Neural Networks an Introduction. Verlag Berlin Heidelberg, Springer.
72. Oja, E. (1982). "A simplified neuron model as a principal component analyzer." Journal of Mathematical Biology: 927-935.
73. Oliveira e Silva, T. (1995). "Laguerre Filters – An Introduction." Revista do Detua: 237-248.
74. Paniagua-Tineo, A., S. Salcedo-Sanz, C. Casanova-Mateo, E. G. Ortiz-García, M. A. Cony and E. Hernández-Martín (2011). "Prediction of daily maximum temperature using a support vector regression algorithm." Renewable Energy **36**(11): 3054-3060.
75. Piper, E. L., M. A. Smit, K. J. Boote and J. W. Jones (1996). "The role of daily minimum temperature in modulating the development rate to flowering in soybean." Field Crops Research **47**(2-3): 211-220.
76. Principe, J. C., N. R. Euliano and W. C. Lefebvre (2000). Neural and Adaptative systems: Fundamentals through simulations. New York, Wiley.
77. Randall, D. A., S. Wood, R. Bony, T. Colman, J. Fichefet, V. Fyfe, A. Kattsov, J. Pitman, J. Shukla, R. J. Srinivasan, A. Stouffer, Sumi and K. E. Taylor (2007). Climate Models and Their Evaluation. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, Cambridge University Press.
78. Research, N. s. O. o. O. a. A. (2012). "3-dimensional grid." Retrieved 29/09, 2012, from http://www.research.noaa.gov/climate/images/modeling_grid.png.
79. Research, N. s. O. o. O. a. A. and O. W. Team. (2005, 2005-09-30). "OAR Climate Research Programs." from http://www.research.noaa.gov/climate/t_modeling.html.

80. Rojas, R. (1996). Neural Networks: A Systematic Introduction. Berlin, Springer.
81. Ruano, A. E., E. M. Crispim, E. Z. E. Conceicao and M. Lucio (2006). "Prediction of building's temperature using neural networks models." Energy and Buildings **38**(6): 682-694.
82. Rubinstein, K. G. and A. M. Sterin (2002). "Comparison of reanalysis products and aerological data." Izvestiya Atmospheric and Oceanic Physics **38**(3): 264-276.
83. Rumelhart, D. E., G. E. Hinton and R. J. Williams (1986). "Learning representations by back-propagating errors." Nature **323**(6088): 533-536.
84. Rummukainen, M. (2010). "State-of-the-art with regional climate models." Wiley Interdisciplinary Reviews: Climate Change **1**(1): 82-96.
85. Sanger, T. D. (1989). "Optimal Unsupervised Learning in a Single-Layer Linear Feedforward Neural Network." Neural Networks: 459-473.
86. Sauter, T., B. Weitzenkamp and C. Schneider (2010). "Spatio-temporal prediction of snow cover in the Black Forest mountain range using remote sensing and a recurrent neural network." International Journal of Climatology **30**(15): 2330-2341.
87. Savage, M. J., C. S. Everson and B. R. Metelerkamp (2009). "Bowen ratio evaporation measurement in a remote montane grassland: Data integrity and fluxes." Journal of Hydrology **376**(1-2): 249-260.
88. Semenov, M. A., R. J. Brooks, E. M. Barrow and C. W. Richardson (1998). "Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates." Climate Research **10**(2): 95-107.
89. Silva, F. M. and L. B. Almeida (1990). Acceleration techniques for the backpropagation algorithm. Neural Networks – EURASIP Workshop, LNCS, Springer-Verlag.
90. Singh, K. P., A. Basant, A. Malik and G. Jain (2009). "Artificial neural network modeling of the river water quality-A case study." Ecological Modelling **220**(6): 888-895.
91. Smith, B. A., G. Hoogenboom and R. W. McClendon (2009). "Artificial neural networks for automated year-round temperature prediction." Computers and Electronics in Agriculture **68**(1): 52-61.
92. Solaiman, T. A., P. Coulibaly and P. Kanaroglou (2008). "Ground-level ozone forecasting using data-driven methods." Air Quality, Atmosphere and Health **1**(4): 179-193.
93. Taghi Sattari, M., K. Yurekli and M. Pal (2012). "Performance evaluation of artificial neural network approaches in forecasting reservoir inflow." Applied Mathematical Modelling **36**(6): 2649-2657.
94. Tolika, K., C. Anagnostopoulou, P. Maheras and M. Vafiadis (2008). "Simulation of future changes in extreme rainfall and temperature conditions over the Greek area: A comparison of two statistical downscaling approaches." Global and Planetary Change **63**(2-3): 132-151.
95. Tomassetti, B., M. Verdecchia and F. Giorgi (2009). "NN5: A neural network based approach for the downscaling of precipitation fields - Model description and preliminary results." Journal of Hydrology **367**(1-2): 14-26.
96. Ustaoglu, B., H. K. Cigizoglu and M. Karaca (2008). "Forecast of daily mean, maximum and minimum temperature time series by three artificial neural network methods." Meteorological Applications **15**(4): 431-445.
97. Vallet-Coulomb, C., D. Legesse, F. Gasse, Y. Travi and T. Chernet (2001). "Lake evaporation estimates in tropical Africa (Lake Ziway, Ethiopia)." Journal of Hydrology **245**(1-4): 1-18.
98. Von Storch, H., E. Zorita and U. Cubasch (1993). "Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime." Journal of Climate **6**(6): 1161-1171.
99. Vries, B. d. and J. C. Principe (1992). "The gamma model-A new neural model for temporal processing." Neural Networks **5**(4): 565-576.

100. Waibel, A., T. Hanazawa, G. Hinton, K. Shikano and K. J. Lang (1989). "Phoneme recognition using time-delay neural networks." IEEE Transactions on Acoustics, Speech, and Signal Processing **37**(3): 328-339.
101. Weart, S. R. (2004). The Discovery of Global Warming, Harvard University Press.
102. Wetterhall, F., S. Halldin and C. Y. Xu (2005). "Statistical precipitation downscaling in central Sweden with the analogue method." Journal of Hydrology **306**(1-4): 174-190.
103. Zemouri, R., R. Gouriveau and N. Zerhouni (2010). "Defining and applying prediction performance metrics on a recurrent NARX time series model." Neurocomputing **73**(13-15): 2506-2521.
104. Zeng, Z., W. Hsieh, A. Shabbar and W. Burrows (2011). "Seasonal prediction of winter extreme precipitation over CANada by support vector regression." Hydrology and Earth System Sciences **15**(1): 65-74.
105. Zhao, S. H., Y. H. Yang, G. Y. Qiu, Q. M. Qin, Y. J. Yao, Y. J. Xiong and C. Q. Li (2010). "Remote detection of bare soil moisture using a surface-temperature-based soil evaporation transfer coefficient." International Journal of Applied Earth Observation and Geoinformation **12**(5): 351-358.
106. Zorita, E. and H. Von Storch (1999). "The analog method as a simple statistical downscaling technique: Comparison with more complicated methods." Journal of Climate **12**(8 PART 2): 2474-2489.

ANEXO 1: DATOS ESTACIONES METEOROLÓGICAS DE ESTUDIO

TABLA 15: DATOS ESTACIONES METEOROLÓGICAS DE ESTUDIO

Estación	Zona Climática	Nombre	Código IDEAM	Ubicación			Grilla NCEP	Periodo Estudio		Datos faltantes	
				Lat	Long	Elevación m.s.n.m.		Inicio	Fin	#	%
E1	Af	STA LETICIA	21055030	0214N	7609W	2085	77X24Y	01/01/1978	30/09/2010	2595	21,70%
E2	Cfc	ISLA DEL SANTUARIO	24015120	0528N	7344W	2580	77X23Y	01/01/1978	30/09/2009	565	4,87%
E3	Am	MARGARITA LAS HDA	35125010	0420N	7209W	150	78X23Y	01/01/1978	31/10/2010	3897	32,50%
E4	BSh	MANAURE	15075030	1146N	7228W	1	78X21Y	01/01/1984	31/03/2010	2674	27,89%
E5	Aw	APTO RAFAEL BARVO	25025080	0920N	7516W	166	77X22Y	01/01/1989	30/09/2010	1375	17,31%

En la tabla se presentan datos de la codificación de las estaciones para el proyecto, el nombre según el IDEAM, la zona climática correspondiente según la clasificación de Köppen y Geiger, ubicación de la estación, la celda del reanálisis donde se ubica la estación, periodos de evaluación según la disponibilidad de datos y la cantidad de información faltante.

ANEXO 2: PRUEBAS ESTADÍSTICAS ALTERACION SERIES RECONSTRUIDAS

ESTACIÓN E1

Prueba de Kruskal-Wallis:

K (Valor observado)	0,281
K (Valor crítico)	3,841
GDL	1
p-valor (bilateral)	0,596
Alfa	0,05

Se ha utilizado una aproximación para calcular el p-valor.

Interpretación de la prueba:

H0: Las muestras vienen de la misma población.

Ha: Las muestras no vienen de la misma población.

Como el p-valor calculado es mayor que el nivel de significación $\alpha=0,05$, no se puede rechazar la hipótesis nula H0.

El riesgo de rechazar la hipótesis nula H0 cuando es verdadera es de 59,63%.

Prueba de Kolmogorov-Smirnov sobre dos muestras / Prueba bilateral:

D	0,013
p-valor	0,342
alfa	0,05

Se ha utilizado una aproximación para calcular el p-valor.

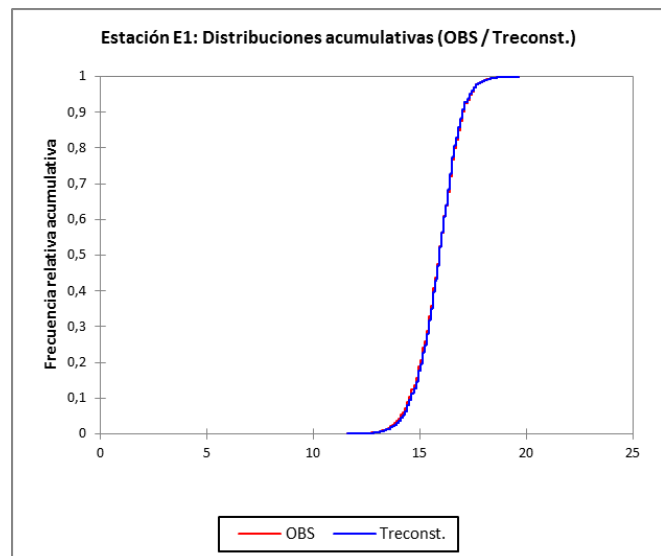
Interpretación de la prueba:

H0: La distribución de las dos muestras es la misma.

Ha: Las distribuciones de las dos muestras son diferentes.

Como el p-valor calculado es mayor que el nivel de significación $\alpha=0,05$, no se puede rechazar la hipótesis nula H0.

El riesgo de rechazar la hipótesis nula H0 cuando es verdadera es de 34,21%.



ESTACIÓN E2

Prueba de Kruskal-Wallis:

K (Valor observado)	0,351
K (Valor crítico)	3,841
GDL	1
p-valor (bilateral)	0,553
Alfa	0,05

Se ha utilizado una aproximación para calcular el p-valor.

Interpretación de la prueba:

H0: Las muestras vienen de la misma población.

Ha: Las muestras no vienen de la misma población.

Como el p-valor calculado es mayor que el nivel de significación $\alpha=0,05$, no se puede rechazar la hipótesis nula H0.

El riesgo de rechazar la hipótesis nula H0 cuando es verdadera es de 55,34%.

Prueba de Kolmogorov-Smirnov sobre dos muestras / Prueba bilateral:

D	0,006
p-valor	0,991
Alfa	0,05

Se ha utilizado una aproximación para calcular el p-valor.

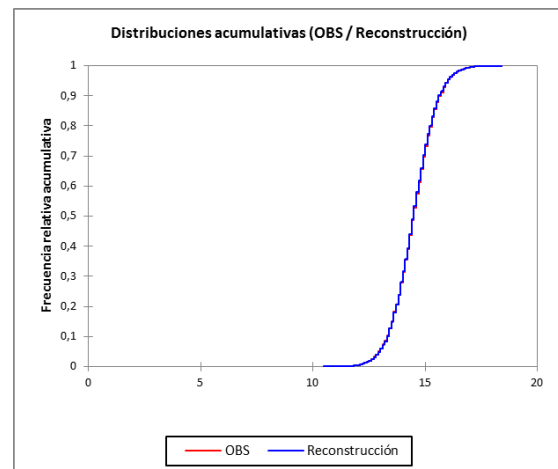
Interpretación de la prueba:

H0: La distribución de las dos muestras es la misma.

Ha: Las distribuciones de las dos muestras son diferentes.

Como el p-valor calculado es mayor que el nivel de significación $\alpha=0,05$, no se puede rechazar la hipótesis nula H0.

El riesgo de rechazar la hipótesis nula H0 cuando es verdadera es de 99,08%.



ESTACIÓN E3

Prueba de Kruskal-Wallis:

K (Valor observado)	0,448
K (Valor crítico)	3,841
GDL	1
p-valor (bilateral)	0,503
Alfa	0,05

Se ha utilizado una aproximación para calcular el p-valor.

Interpretación de la prueba:

H0: Las muestras vienen de la misma población.

Ha: Las muestras no vienen de la misma población.

Como el p-valor calculado es mayor que el nivel de significación $\alpha=0,05$, no se puede rechazar la hipótesis nula H0.

El riesgo de rechazar la hipótesis nula H0 cuando es verdadera es de 50,35%.

Prueba de Kolmogorov-Smirnov sobre dos muestras / Prueba bilateral:

D	0,011
p-valor	0,623
Alfa	0,05

Se ha utilizado una aproximación para calcular el p-valor.

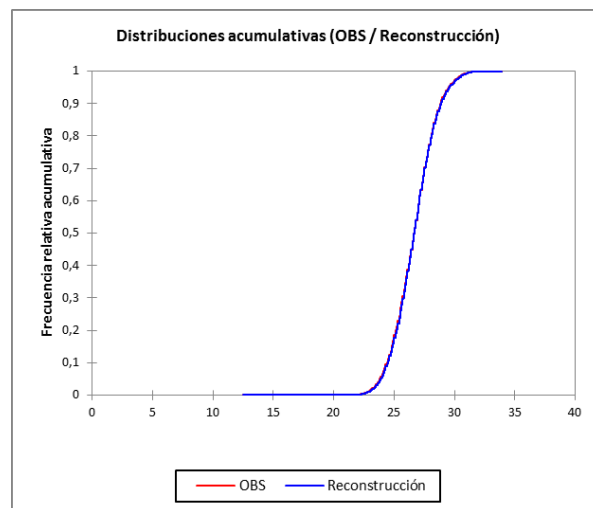
Interpretación de la prueba:

H0: La distribución de las dos muestras es la misma.

Ha: Las distribuciones de las dos muestras son diferentes.

Como el p-valor calculado es mayor que el nivel de significación $\alpha=0,05$, no se puede rechazar la hipótesis nula H0.

El riesgo de rechazar la hipótesis nula H0 cuando es verdadera es de 62,29%.



ESTACIÓN E4

Prueba de Kruskal-Wallis:

K (Valor observado)	21,946
K (Valor crítico)	3,841
GDL	1
p-valor (bilateral)	< 0,0001
Alfa	0,05

Se ha utilizado una aproximación para calcular el p-valor.

Interpretación de la prueba:

H0: Las muestras vienen de la misma población.

Ha: Las muestras no vienen de la misma población.

Como el p-valor computado es menor que el nivel de significación $\alpha=0,05$, se debe rechazar la hipótesis nula H0, y aceptar la hipótesis alternativa Ha.

El riesgo de rechazar la hipótesis nula H0 cuando es verdadera es menor que 0,01%.

Prueba de Kolmogorov-Smirnov sobre dos muestras / Prueba bilateral:

D	0,038
p-valor	< 0,0001
Alfa	0,05

Se ha utilizado una aproximación para calcular el p-valor.

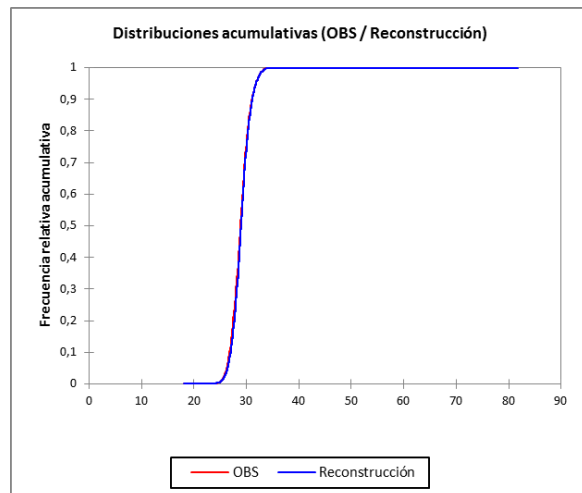
Interpretación de la prueba:

H0: La distribución de las dos muestras es la misma.

Ha: Las distribuciones de las dos muestras son diferentes.

Como el p-valor computado es menor que el nivel de significación $\alpha=0,05$, se debe rechazar la hipótesis nula H0, y aceptar la hipótesis alternativa Ha.

El riesgo de rechazar la hipótesis nula H0 cuando es verdadera es menor que 0,01%.



ESTACIÓN E5

Prueba de Kruskal-Wallis:

K (Valor observado)	0,347
K (Valor crítico)	3,841
GDL	1
p-valor (bilateral)	0,556
Alfa	0,05

Se ha utilizado una aproximación para calcular el p-valor.

Interpretación de la prueba:

H0: Las muestras vienen de la misma población.

Ha: Las muestras no vienen de la misma población.

Como el p-valor calculado es mayor que el nivel de significación $\alpha=0,05$, no se puede rechazar la hipótesis nula H0.

El riesgo de rechazar la hipótesis nula H0 cuando es verdadera es de 55,56%.

Prueba de Kolmogorov-Smirnov sobre dos muestras / Prueba bilateral:

D	0,012
p-valor	0,524
Alfa	0,05

Se ha utilizado una aproximación para calcular el p-valor.

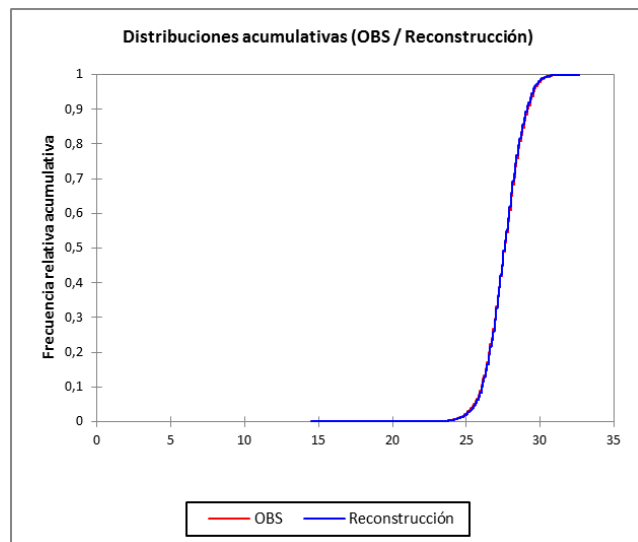
Interpretación de la prueba:

H0: La distribución de las dos muestras es la misma.

Ha: Las distribuciones de las dos muestras son diferentes.

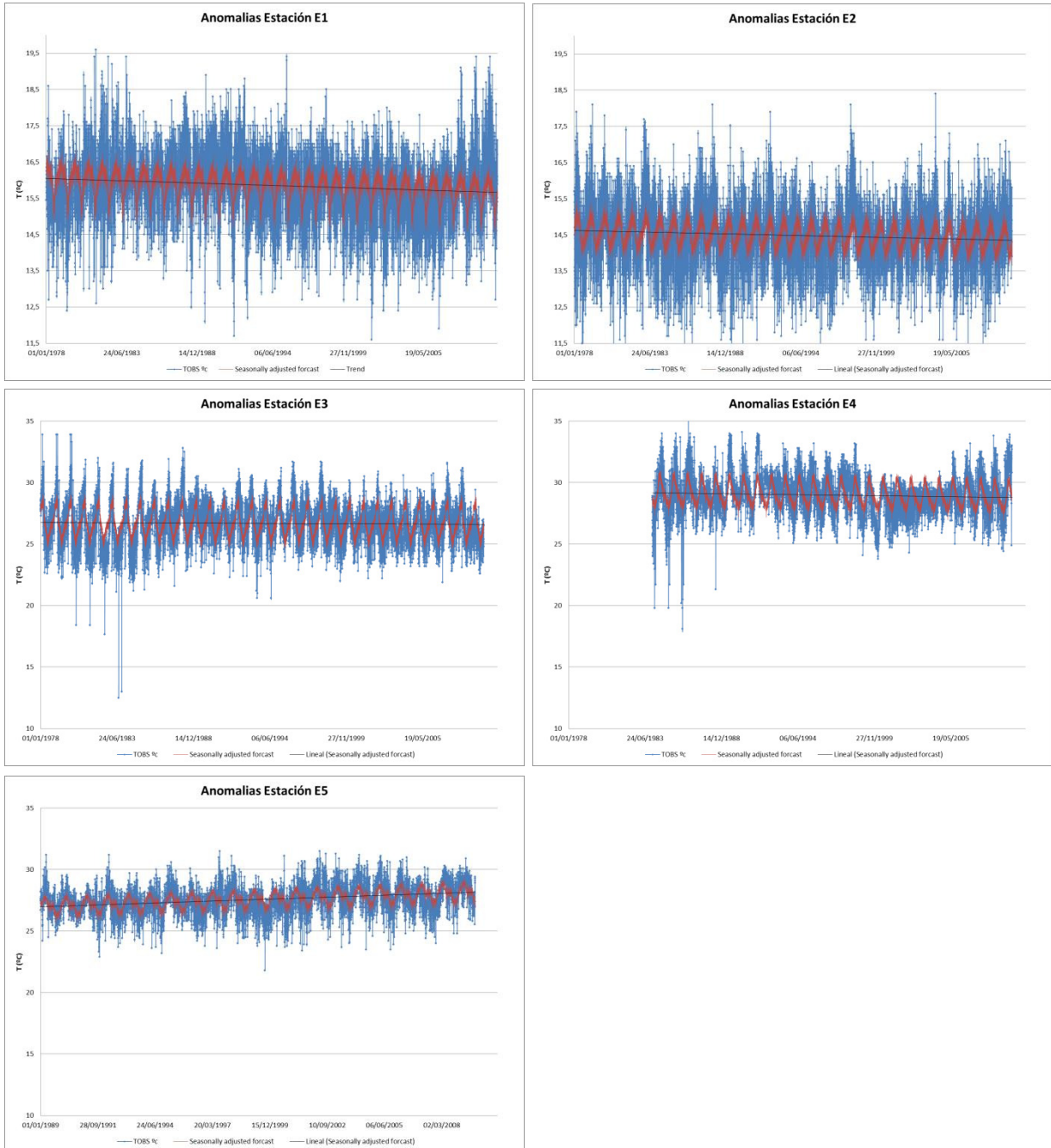
Como el p-valor calculado es mayor que el nivel de significación $\alpha=0,05$, no se puede rechazar la hipótesis nula H0.

El riesgo de rechazar la hipótesis nula H0 cuando es verdadera es de 52,43%.



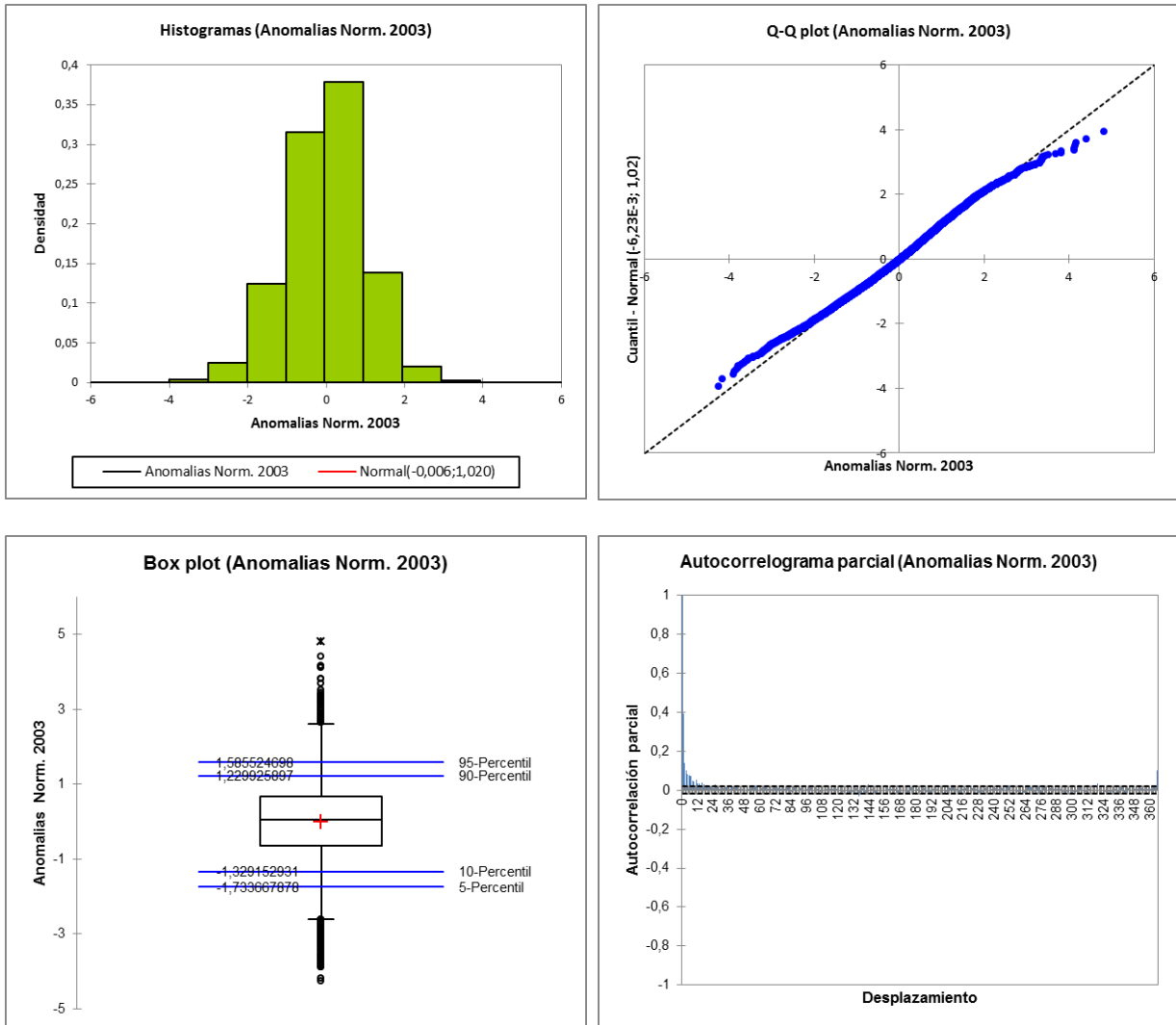
ANEXO 3: GRÁFICAS SEPARACIÓN TENDENCIA, ESTACIONALIDAD Y ANOMALÍAS ESTACIONES METEOROLÓGICAS

Las siguientes graficas muestran en negro la tendencia de la temperatura, en rojo la estacionalidad según se calculó en capítulo 5.2.2. y en azul los datos de temperatura según las series de tiempo reconstruidas.



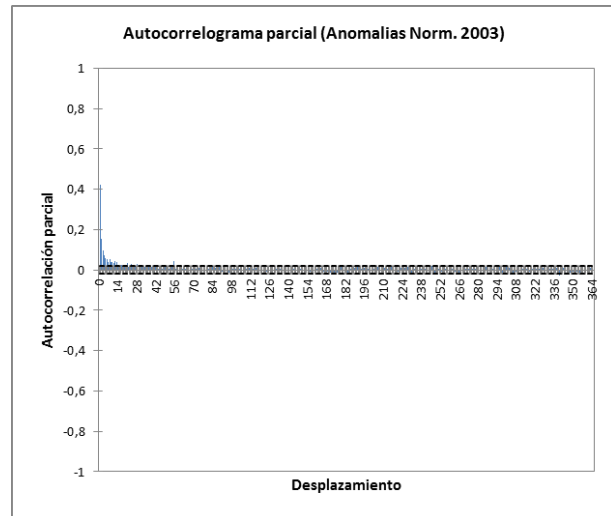
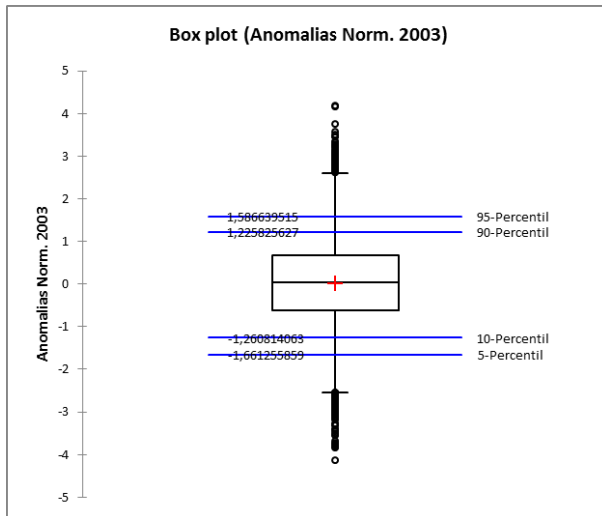
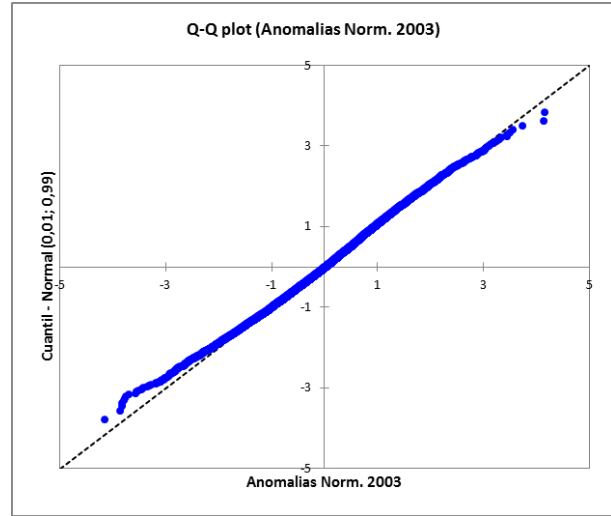
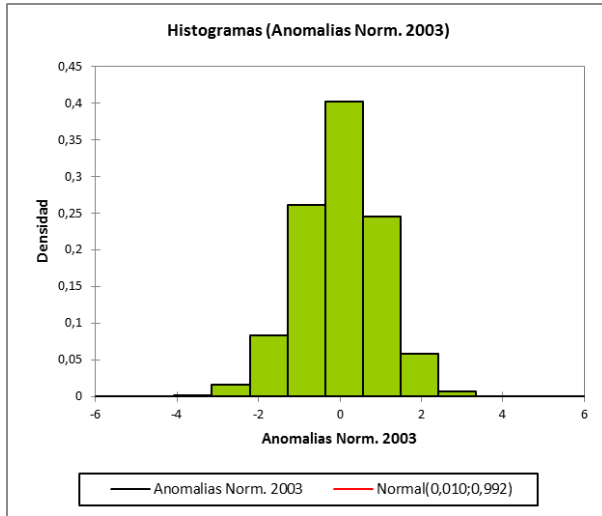
ANEXO 4: ANALISIS DESCRIPTIVO DE LAS ANOMALÍAS DE LA ESTACIONE METEOROLÓGICAS

Estación E1:



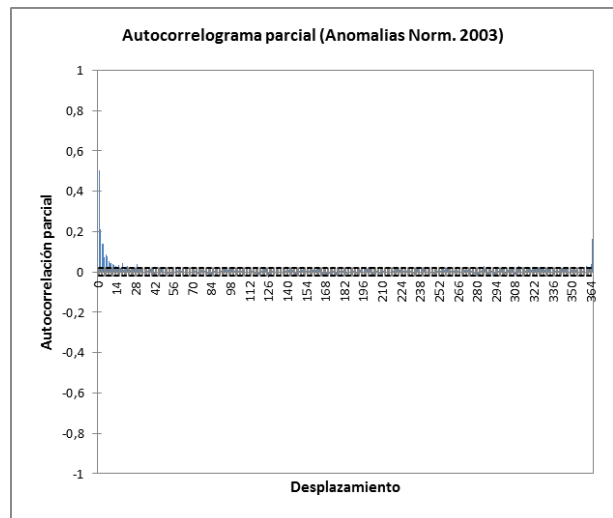
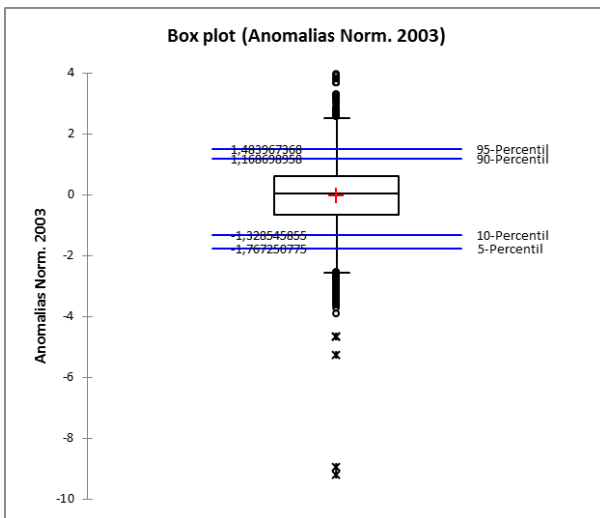
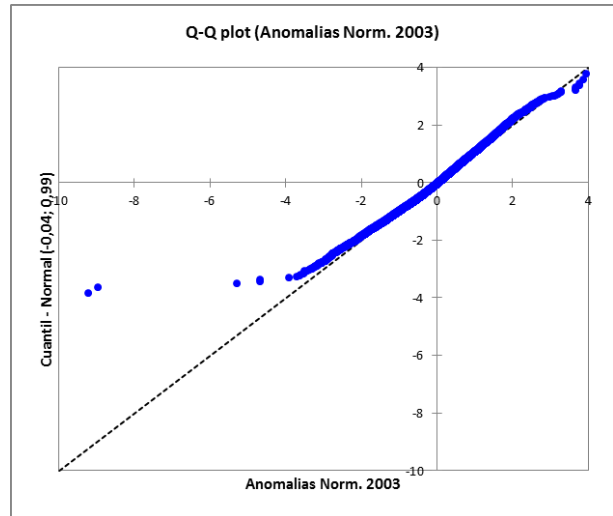
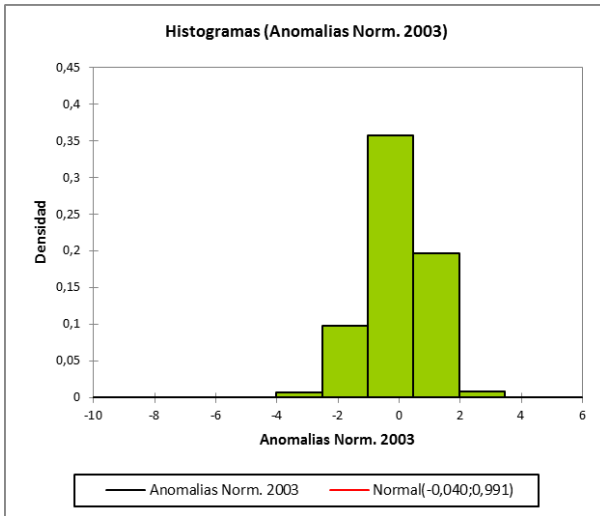
De izquierda a derecha de arriba a abajo: a) Histograma de los datos de anomalías de la temperatura, b) Gráfica cuantil-cuantil para verificación de la normalidad, c) Grafica boxplot para identificación de outliers y d) Grafica de autocorrelación de los datos

Estación E2:



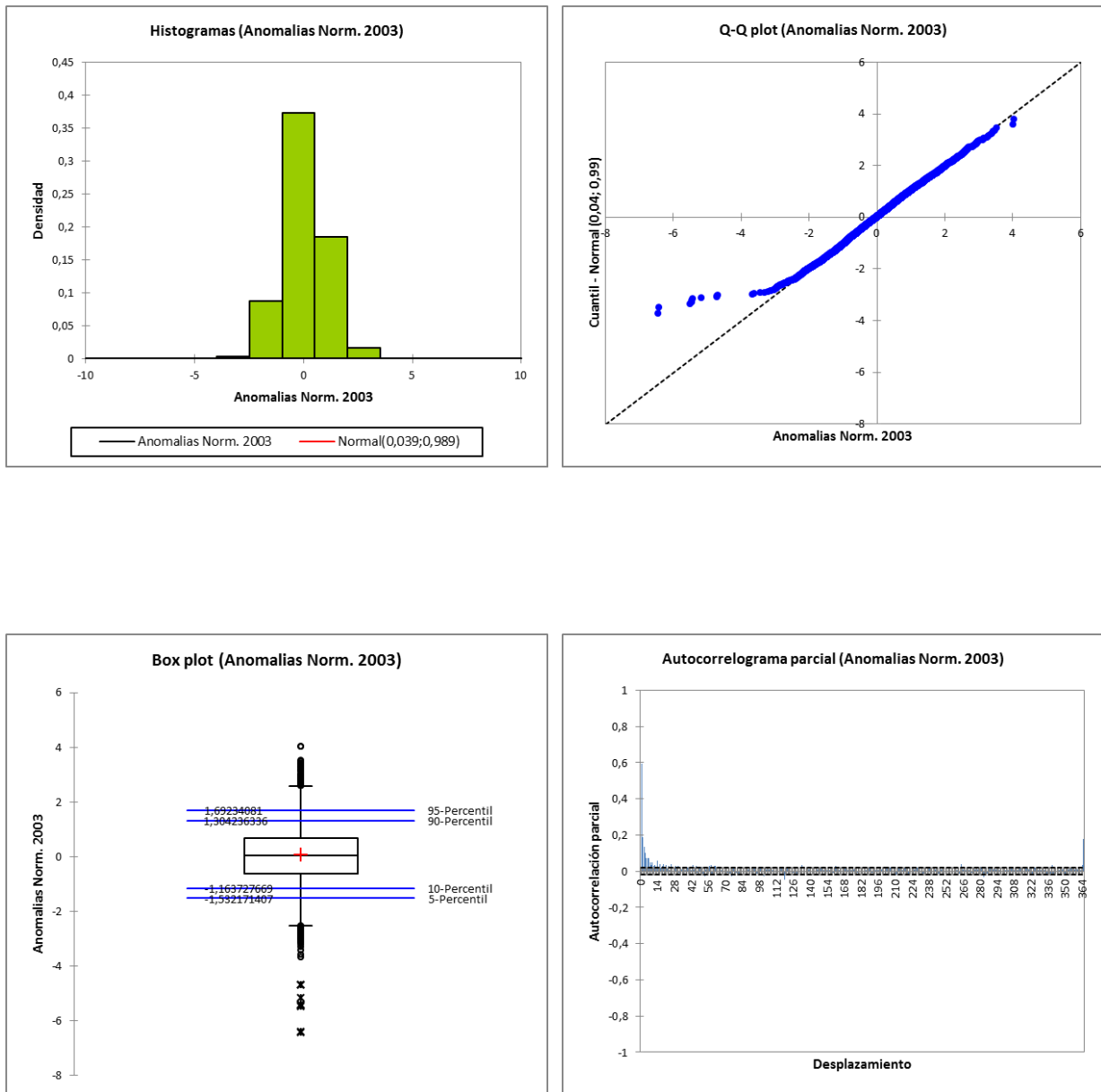
De izquierda a derecha de arriba a abajo: a) Histograma de los datos de anomalías de la temperatura, b) Gráfica cuantil-cuantil para verificación de la normalidad, c) Grafica boxplot para identificación de outliers y d) Grafica de autocorrelación de los datos

Estación E3



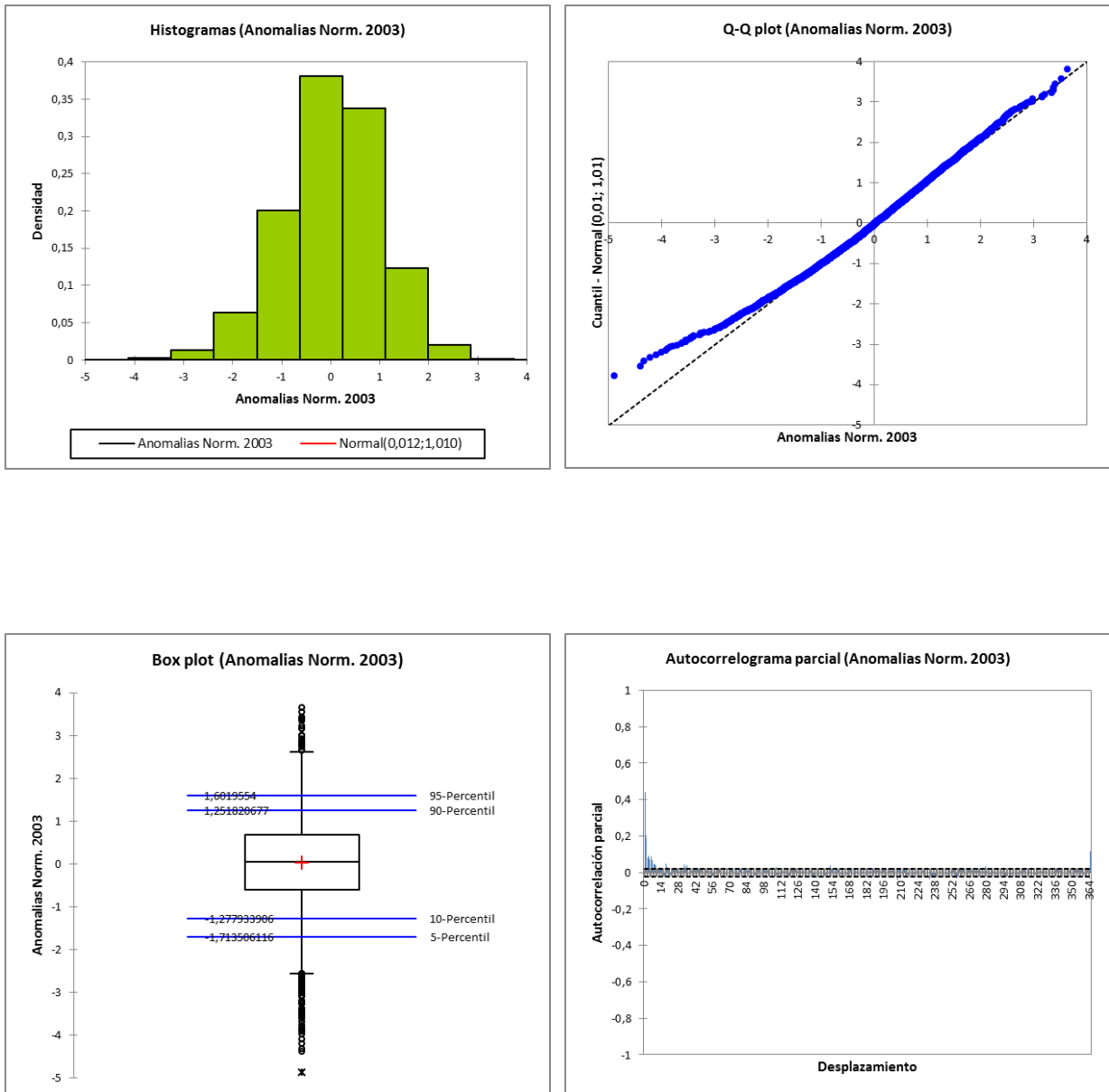
De izquierda a derecha de arriba a abajo: a) Histograma de los datos de anomalías de la temperatura, b) Gráfica cuantil-cuantil para verificación de la normalidad, c) Grafica boxplot para identificación de outliers y d) Grafica de autocorrelación de los datos

Estación E4:



De izquierda a derecha de arriba a abajo: a) Histograma de los datos de anomalías de la temperatura, b) Gráfica cuantil-cuantil para verificación de la normalidad, c) Grafica boxplot para identificación de outliers y d) Grafica de autocorrelación de los datos

Estación E5:

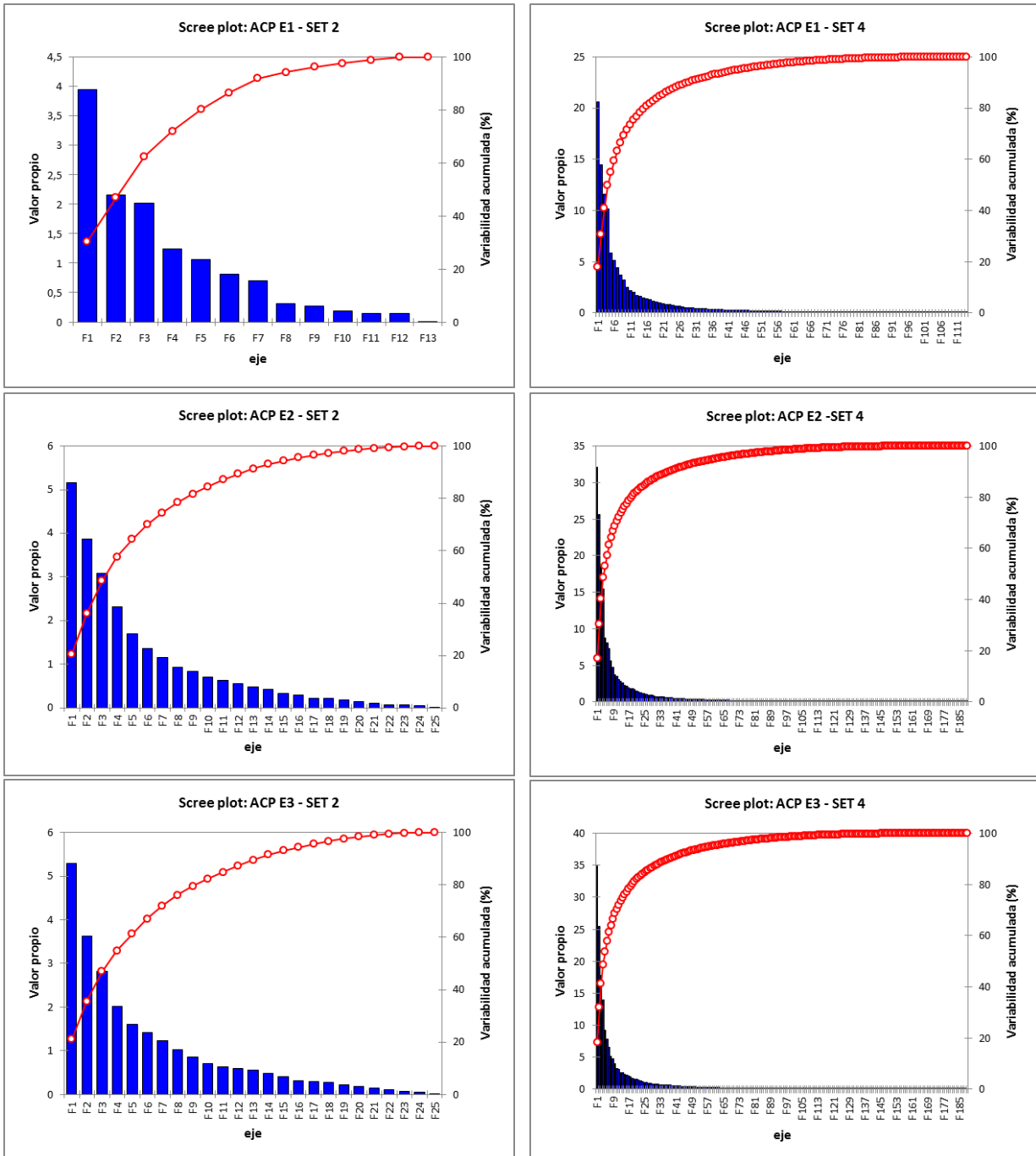


De izquierda a derecha de arriba a abajo: a) Histograma de los datos de anomalías de la temperatura, b) Gráfica cuantil-cuantil para verificación de la normalidad, c) Grafica boxplot para identificación de outliers y d) Grafica de autocorrelación de los datos

ANEXO 5: ANÁLISIS DE COMPONENTES PRINCIPALES ESTACIONES

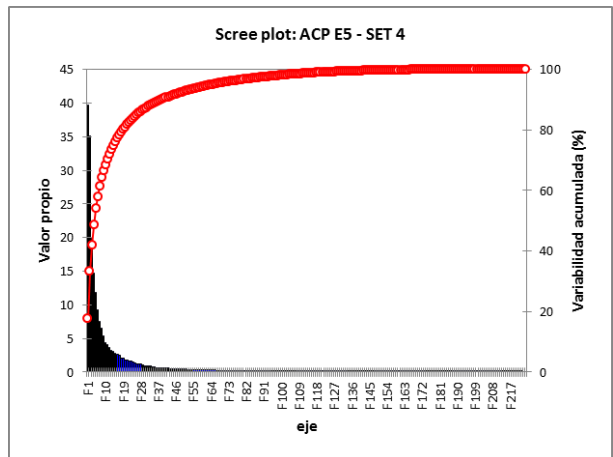
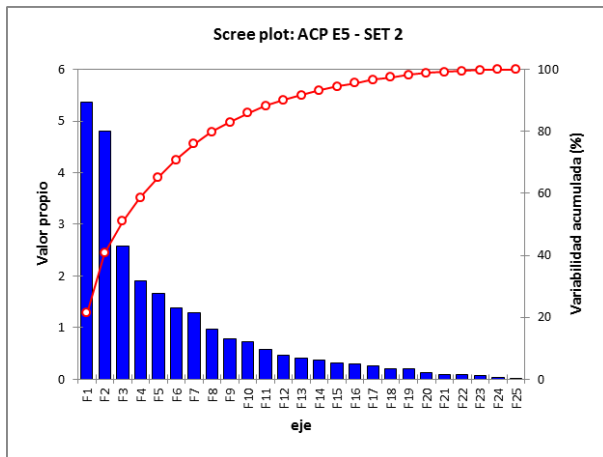
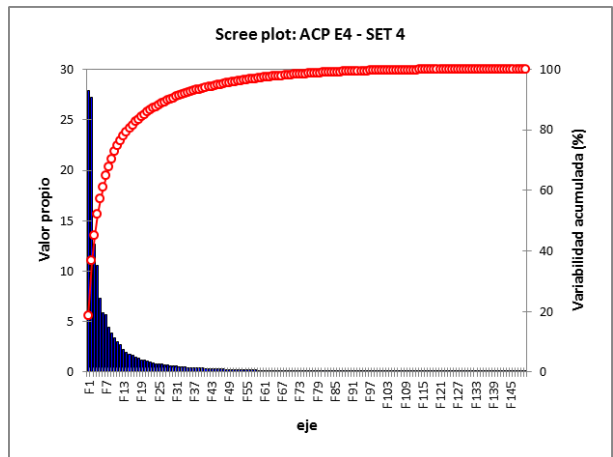
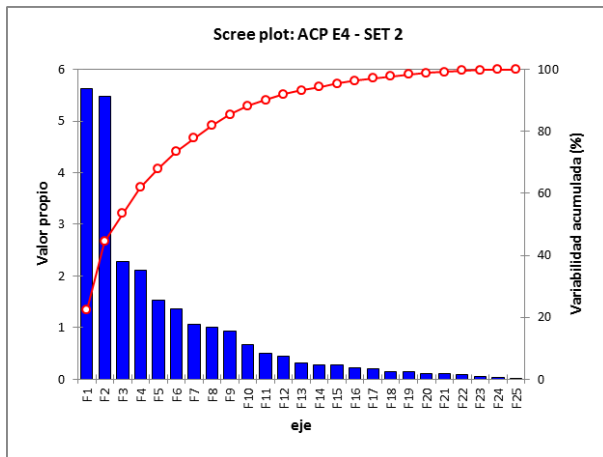
ESTACIONES E1, E2 y E3:

En la columna izquierda se presenta el Análisis de Componentes Principales aplicado sobre el conjunto de datos 2 de cada estación. En la columna derecha se presenta el ACP aplicado sobre el conjunto de datos 4 de cada estación:



ESTACIONES E4 y E5

En la columna izquierda se presenta el Análisis de Componentes Principales aplicado sobre el conjunto de datos 2 de cada estación. En la columna derecha se presenta el ACP aplicado sobre el conjunto de datos 4 de cada estación:



ANEXO 6: REGRESION LINEAL STEPWISE

TABLA 16: ÚLTIMA LÍNEA DE ENTRADA/SALIDA DE LA REGRESIÓN STEPWISE

Estación	SET	No. de variables	Variables	Variable IN/OUT	Estado	MEC	R ²	R ² ajustado	Cp de Mallows	AIC de Akaike	SBC de Schwarz	PC de Amemiya
1	2	9	X66 / X67 / X69 / X70 / X71 / X72 / X73 / X74 / X75	X70	IN	0,939	0,050	0,049	7,738	-393,363	-325,714	0,953
1	4	21	X8 / X20 / X23 / X24 / X28 / X40 / X60 / X61 / X63 / X71 / X77 / X85 / X87 / X88 / X101 / X104 / X108 / X109 / X110 / X112 / X113	X111	OUT	0,900	0,091	0,088	68,319	-653,672	-504,842	0,915
2	2	10	X105 / X109 / X112 / X114 / X116 / X119 / X120 / X122 / X123 / X124	X107	OUT	0,841	0,194	0,193	25,664	-1100,076	-1025,662	0,809
2	4	30	X6 / X7 / X16 / X17 / X23 / X24 / X27 / X53 / X75 / X81 / X92 / X94 / X97 / X99 / X111 / X122 / X123 / X147 / X154 / X162 / X165 / X171 / X174 / X177 / X186 / X188 / X197 / X199 / X204 / X223	X171	IN	0,803	0,232	0,229	78,666	-1374,340	-1164,626	0,775
3	2	12	X176 / X178 / X180 / X182 / X183 / X185 / X186 / X190 / X193 / X194 / X195 / X197	X200	OUT	1,022	0,075	0,073	14,348	154,237	242,182	0,929
3	4	38	X10 / X14 / X17 / X24 / X30 / X33 / X34 / X35 / X41 / X42 / X47 / X50 / X53 / X60 / X67 / X73 / X91 / X97 / X98 / X99 / X103 / X106 / X116 / X122 / X126 / X128 / X134 / X153 / X162 / X163 / X169 / X175 / X176 / X178 / X180 / X184 / X191 / X193	X98	IN	0,965	0,130	0,125	78,372	-191,467	72,368	0,880
4	2	11	X101 / X109 / X110 / X112 / X113 / X115 / X117 / X118 / X121 / X123 / X125	X109	IN	0,976	0,033	0,031	12,210	-108,601	-30,569	0,972
4	4	18	X12 / X23 / X42 / X43 / X48 / X54 / X68 / X77 / X79 / X89 / X97 / X100 / X106 / X107 / X113 / X125 / X132 / X141	X106	IN	0,957	0,052	0,049	92,671	-195,420	-71,869	0,955
5	2	7	X101 / X102 / X108 / X112 / X122 / X123 / X124	X122	IN	0,734	0,140	0,139	20,079	-1136,488	-1086,770	0,863
5	4	31	X26 / X41 / X65 / X66 / X71 / X72 / X74 / X75 / X76 / X77 / X80 / X92 / X100 / X108 / X112 / X124 / X128 / X140 / X147 / X166 / X177 / X178 / X179 / X195 / X201 / X202 / X204 / X205 / X213 / X221 / X223	X41	IN	0,676	0,213	0,207	60,288	-1416,592	-1217,720	0,800

La tabla muestra la última línea de la Eliminación/incorporación progresiva (Stepwise Regression). Aparecen las variables derivadas del conjunto de datos (set) 2 o 4 según el caso que fueron seleccionadas para el modelo de regresión lineal y posteriores modelos comparativos RNA. Adicional se presentan los datos estadísticos de la regresión.

ANEXO 7: ARQUITECTURAS DE MODELOS EVALUADOS EN EL PROYECTO

No. Modelo	Nombre del modelo	No. Modelo	Nombre del modelo	No. Modelo	Nombre del modelo	No. Modelo	Nombre del modelo
1	RL-T (Linear Regression)	26	TDNN-1-O-M (Time-Delay Network)	51	LR-0-B-L (Linear Regression)	76	LR-0-B-M (Linear Regression)
2	RL25 (Linear Regression)	27	TDNN-1-B-M (Time-Delay Network)	52	MLP-1-B-L (Multilayer Perceptron)	77	LR-0-B-L (Linear Regression)
3	SW25 (Stepwise)	28	RN-1-O-M (Recurrent Network)	53	RBF-1-B-L (Radial Basis Function)	78	MLP-1-B-L (Multilayer Perceptron)
4	RL225 (Linear Regression)	29	RN-1-B-M (Recurrent Network)	54	GFF-1-B-L (Generalized Feedforward)	79	RBF-1-B-L (Radial Basis Function)
5	SW225 (Stepwise)	30	TLRN-1-O-M (Time-Lag Recurrent Network)	55	MLPPCA-1-B-L (MLP with PCA)	80	GFF-1-B-L (Generalized Feedforward)
6	MLP-1-O-M (Multilayer Perceptron)	31	TLRN-1-B-M (Time-Lag Recurrent Network)	56	TDNN-1-B-L (Time-Delay Network)	81	MLPPCA-1-B-L (MLP with PCA)
7	LR-0-B-M (Linear Regression)	32	MLP-1-O-M (Multilayer Perceptron)	57	TLRN-1-B-L (Time-Lag Recurrent Network)	82	TDNN-1-B-L (Time-Delay Network)
8	LR-0-B-L (Linear Regression)	33	LR-0-B-M (Linear Regression)	58	RN-1-B-L (Recurrent Network)	83	TLRN-1-B-L (Time-Lag Recurrent Network)
9	MLP-1-B-L (Multilayer Perceptron)	34	MLP-1-B-M (Multilayer Perceptron)	59	MLP-2-B-L (Multilayer Perceptron)	84	RN-1-B-L (Recurrent Network)
10	RBF-1-B-L (Radial Basis Function)	35	MLP-2-O-M (Multilayer Perceptron)	60	MLP-1-B-M (Multilayer Perceptron)	85	MLP-2-B-L (Multilayer Perceptron)
11	GFF-1-B-L (Generalized Feedforward)	36	MLP-2-B-M (Multilayer Perceptron)	61	MLP-2-O-M (Multilayer Perceptron)	86	MLP-1-B-M (Multilayer Perceptron)
12	MLPPCA-1-B-L (MLP with PCA)	37	MLPPCA-1-O-M (MLP with PCA)	62	MLP-2-B-M (Multilayer Perceptron)	87	MLP-2-O-M (Multilayer Perceptron)
13	TDNN-1-B-L (Time-Delay Network)	38	MLPPCA-1-B-M (MLP with PCA)	63	MLPPCA-1-O-M (MLP with PCA)	88	MLP-2-B-M (Multilayer Perceptron)
14	TLRN-1-B-L (Time-Lag Recurrent Network)	39	GFF-1-O-M (Generalized Feedforward)	64	MLPPCA-1-B-M (MLP with PCA)	89	MLPPCA-1-O-M (MLP with PCA)
15	RN-1-B-L (Recurrent Network)	40	GFF-1-B-M (Generalized Feedforward)	65	GFF-1-O-M (Generalized Feedforward)	90	MLPPCA-1-B-M (MLP with PCA)
16	MLP-2-B-L (Multilayer Perceptron)	41	RBF-1-O-M (Radial Basis Function)	66	GFF-1-B-M (Generalized Feedforward)	91	GFF-1-O-M (Generalized Feedforward)
17	MLP-1-B-M (Multilayer Perceptron)	42	RBF-1-B-M (Radial Basis Function)	67	RBF-1-O-M (Radial Basis Function)	92	GFF-1-B-M (Generalized Feedforward)
18	MLP-2-O-M (Multilayer Perceptron)	43	TDNN-1-O-M (Time-Delay Network)	68	RBF-1-B-M (Radial Basis Function)	93	RBF-1-O-M (Radial Basis Function)
19	MLP-2-B-M (Multilayer Perceptron)	44	TDNN-1-B-M (Time-Delay Network)	69	TDNN-1-O-M (Time-Delay Network)	94	RBF-1-B-M (Radial Basis Function)
20	MLPPCA-1-O-M (MLP with PCA)	45	RN-1-O-M (Recurrent Network)	70	TDNN-1-B-M (Time-Delay Network)	95	TDNN-1-O-M (Time-Delay Network)
21	MLPPCA-1-B-M (MLP with PCA)	46	RN-1-B-M (Recurrent Network)	71	RN-1-O-M (Recurrent Network)	96	TDNN-1-B-M (Time-Delay Network)
22	GFF-1-O-M (Generalized Feedforward)	47	TLRN-1-O-M (Time-Lag Recurrent Network)	72	RN-1-B-M (Recurrent Network)	97	RN-1-O-M (Recurrent Network)
23	GFF-1-B-M (Generalized Feedforward)	48	TLRN-1-B-M (Time-Lag Recurrent Network)	73	TLRN-1-O-M (Time-Lag Recurrent Network)	98	RN-1-B-M (Recurrent Network)
24	RBF-1-O-M (Radial Basis Function)	49	MLP-1-O-M (Multilayer Perceptron)	74	TLRN-1-B-M (Time-Lag Recurrent Network)	99	TLRN-1-O-M (Time-Lag Recurrent Network)
25	RBF-1-B-M (Radial Basis Function)	50	LR-0-B-M (Linear Regression)	75	MLP-1-O-M (Multilayer Perceptron)	100	TLRN-1-B-M (Time-Lag Recurrent Network)

ANEXO 8: 10 MEJORES MODELOS POR ESTACIÓN

En las siguientes tablas se muestran los 10 mejores modelos para cada estación. La tabla contiene información del puesto del modelo, el número del modelo que corresponde al orden cronológico en que se desarrolló, la topología de la red y los resultados de correlación y error medio en los grupos de entrenamiento, validación y prueba.

TABLA 17: MEJORES MODELOS ESTACIÓN E1

Puesto	Modelo	Variables	Model Name	Training			Cross Validation			Testing		
				MSE	r	MAE	MSE	r	MAE	MSE	r	MAE
1	30	13	TLRN-1-O-M (Time-Lag Recurrent Network)	0,94941968	0,24468512	0,7579074	0,92327985	0,31173	0,76626416	1,11891254	0,28997526	0,81216752
2	73	9	TLRN-1-O-M (Time-Lag Recurrent Network)	0,93318334	0,26190068	0,75266111	0,94080289	0,30757987	0,76862032	1,1060513	0,29935675	0,81079566
3	99	21	TLRN-1-O-M (Time-Lag Recurrent Network)	0,89905615	0,32264224	0,74569935	0,92784316	0,3041005	0,770984	1,09710311	0,28930295	0,8074919
4	97	21	RN-1-O-M (Recurrent Network)	1,02586577	0,21796604	0,79078748	0,99808403	0,29987559	0,79239697	1,41747863	0,16063561	0,91896554
5	32	114	MLP-1-O-M (Multilayer Perceptron)	1,16364299	0,23819224	0,84819826	1,09869477	0,28699233	0,83510506	1,49455656	0,26310201	0,96135889
6	95	21	TDNN-1-O-M (Time-Delay Network)	1,07164674	0,25576375	0,81512788	1,03800669	0,28175383	0,8169322	1,30590646	0,28378633	0,90153706
7	71	9	RN-1-O-M (Recurrent Network)	0,98617112	0,19501294	0,7750218	0,94029565	0,27869924	0,77240453	1,12523518	0,22712345	0,8203334
8	69	9	TDNN-1-O-M (Time-Delay Network)	1,01182023	0,20798185	0,79269184	0,94234013	0,27502908	0,77788275	1,04869387	0,28643252	0,79455832
9	75	21	MLP-1-O-M (Multilayer Perceptron)	1,06520534	0,23332219	0,80830324	1,00449367	0,27207672	0,79997036	1,30168873	0,24673043	0,88661947
10	18	13	MLP-2-O-M (Multilayer Perceptron)	0,93448504	0,24782317	0,75216926	0,94365526	0,27199894	0,77306997	1,09433889	0,28216384	0,80409331

TABLA 18: MEJORES MODELOS ESTACION E2

Puesto	Modelo	Variables	Model Name	Training			Cross Validation			Testing		
				MSE	r	MAE	MSE	r	MAE	MSE	r	MAE
1	75	30	MLP-1-O-M (Multilayer Perceptron)	0,83235668	0,44969479	0,72204051	0,71039154	0,52275131	0,66699559	0,65564165	0,37175982	0,64429104
2	77	30	LR-0-B-L (Linear Regression)	0,79914092	0,48208515	0,70957948	0,7207934	0,51840377	0,67272852	0,6706299	0,37044996	0,65238918
3	5	30	SW225 (Stepwise)	0,79914092	0,48208515	0,70957972	0,72079361	0,51840352	0,67272867	0,67063	0,37045227	0,6523885
4	78	30	MLP-1-B-L (Multilayer Perceptron)	0,78017369	0,50071691	0,7014809	0,72429247	0,51216845	0,67741758	0,68462273	0,35960135	0,65825971
5	80	30	GFF-1-B-L (Generalized Feedforward)	0,80636135	0,47596214	0,71251932	0,73259647	0,51128169	0,67765447	0,66771513	0,36615955	0,65041354
6	4	189	RL225 (Linear Regression)	0,77342652	0,50705561	0,6980317	0,73488442	0,5106996	0,68237496	0,6798031	0,37500625	0,65216073
7	16	25	MLP-2-B-L (Multilayer Perceptron)	0,81387236	0,46991357	0,70916192	0,72348948	0,50796444	0,67262346	0,67782445	0,34417183	0,64886457
8	86	30	MLP-1-B-M (Multilayer Perceptron)	0,78927518	0,49194595	0,70521137	0,73581535	0,50611711	0,68022351	0,6768549	0,36555589	0,6549602
9	11	25	GFF-1-B-L (Generalized Feedforward)	0,80827532	0,47456675	0,70712495	0,72862282	0,50502972	0,67208608	0,67615307	0,35542213	0,64860075
10	17	25	MLP-1-B-M (Multilayer Perceptron)	0,80907433	0,47222031	0,70770019	0,72899251	0,50387288	0,67559087	0,69345178	0,33781777	0,65667677

TABLA 19: MEJORES MODELOS ESTACION E3

Puesto	Modelo	Variables	Model Name	Training			Cross Validation			Testing		
				MSE	r	MAE	MSE	r	MAE	MSE	r	MAE
1	75	38	MLP-1-O-M (Multilayer Perceptron)	1,09786355	0,22177449	0,80390967	0,72623466	0,20980439	0,67829809	0,93403552	0,07199884	0,75792498
2	91	38	GFF-1-O-M (Generalized Feedforward)	1,12427413	0,19440054	0,81247156	0,75729019	0,17120126	0,69277804	1,02330084	0,01971104	0,78841049
3	78	38	MLP-1-B-L (Multilayer Perceptron)	1,07954054	0,2032823	0,80206016	0,75528903	0,17086022	0,69544222	0,92011977	0,0580253	0,75468576
4	73	12	TLRN-1-O-M (Time-Lag Recurrent Network)	1,094803	0,25912432	0,80689911	0,74176145	0,17032488	0,6878333	0,89052282	0,10849594	0,73704992
5	97	38	RN-1-O-M (Recurrent Network)	1,05522595	0,2653392	0,79285248	0,76809229	0,16962089	0,69733394	0,974615	0,02109677	0,77096939
6	35	189	MLP-2-O-M (Multilayer Perceptron)	1,07057077	0,24872566	0,7995409	0,74866052	0,14230569	0,69277666	0,92053624	0,01156888	0,74292647
7	89	19	MLPPCA-1-O-M (MLP with PCA)	1,12092605	0,15988845	0,81437755	0,74542153	0,13018057	0,68511001	0,91309693	-0,00241521	0,74399735
8	14	25	TLRN-1-B-L (Time-Lag Recurrent Network)	0,89070031	0,43869372	0,73848964	0,93692021	0,12729613	0,77795837	0,99634042	0,02275384	0,79685038
9	18	25	MLP-2-O-M (Multilayer Perceptron)	1,07542536	0,24982618	0,79694136	0,75115623	0,12151166	0,69228829	0,9226681	-0,02946855	0,74185853
10	99	38	TLRN-1-O-M (Time-Lag Recurrent Network)	1,13992512	0,12655593	0,82496166	0,7587338	0,11475192	0,68788128	0,89229101	0,08107787	0,73841735

TABLA 20: MEJORES MODELOS ESTACION E4

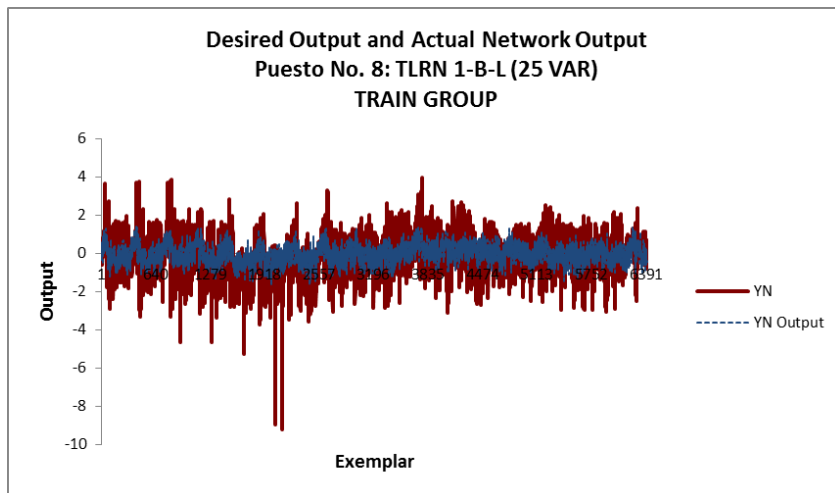
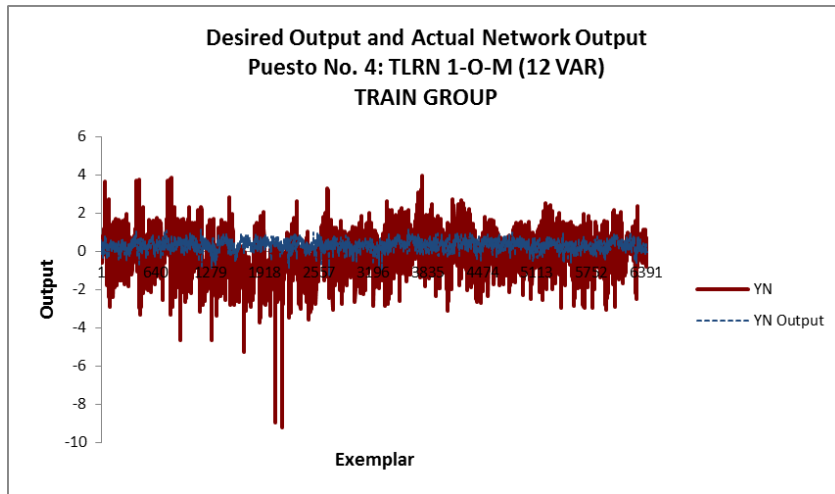
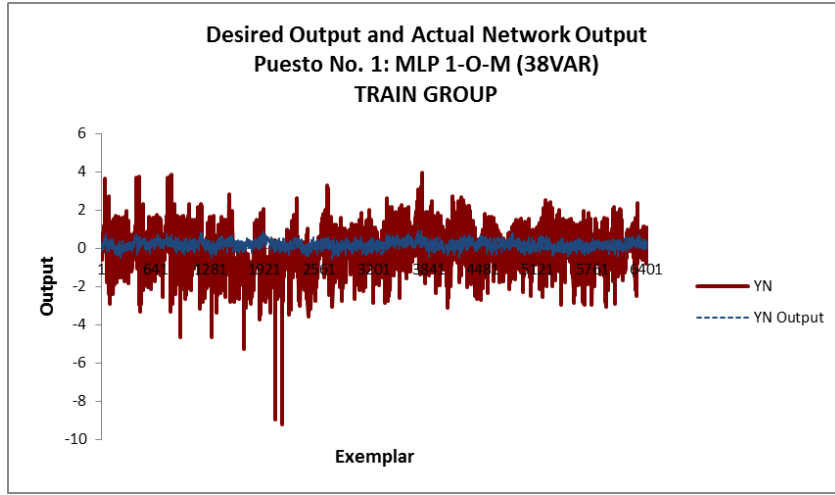
Puesto	Variables	Model Name	Training			Cross Validation			Testing		
			MSE	r	MAE	MSE	r	MAE	MSE	r	MAE
1	11	TDNN-1-O-M (Time-Delay Network)	1,13046376	0,17988047	0,82521731	0,83233325	0,38956901	0,72577163	1,05439584	-0,22820954	0,82510219
2	18	MLP-2-B-L (Multilayer Perceptron)	1,49525402	0,17749169	0,97296377	0,967421	0,36391733	0,80315271	0,85277066	-0,15491996	0,73750936
3	75	MLPPCA-1-O-M (MLP with PCA)	1,07837064	0,12772731	0,80941973	0,8628629	0,3609197	0,73982263	0,98139138	-0,28166901	0,79860684
4	25	RN-1-O-M (Recurrent Network)	1,12266176	0,21285239	0,8249895	0,92917161	0,349514	0,76552244	1,08833421	-0,04510162	0,83415962
5	18	TLRN-1-O-M (Time-Lag Recurrent Network)	0,97559902	0,30783023	0,76971009	0,93708655	0,34256721	0,76530636	1,11870326	-0,09251849	0,84903451
6	11	GFF-1-O-M (Generalized Feedforward)	1,02829769	0,05984814	0,79181872	0,968226	0,33719436	0,78060753	1,0456853	-0,19542478	0,82291757
7	11	TLRN-1-O-M (Time-Lag Recurrent Network)	0,97199516	0,25955285	0,76588119	0,98943137	0,32733751	0,78773743	1,22142945	-0,13539957	0,8695775
8	25	TLRN-1-O-M (Time-Lag Recurrent Network)	0,95548294	0,29692306	0,76030275	0,96413633	0,32645327	0,77632088	1,16049361	-0,07093212	0,86224406
9	25	GFF-1-O-M (Generalized Feedforward)	1,0283344	0,11503184	0,7899413	0,90895097	0,31821927	0,76150655	0,9437409	-0,11724557	0,77565111
10	150	MLP-1-O-M (Multilayer Perceptron)	1,14980847	0,08784514	0,83499798	0,89353358	0,31694828	0,7573322	0,95167312	-0,13037537	0,78380221

TABLA 21: MEJORES MODELOS ESTACION E5

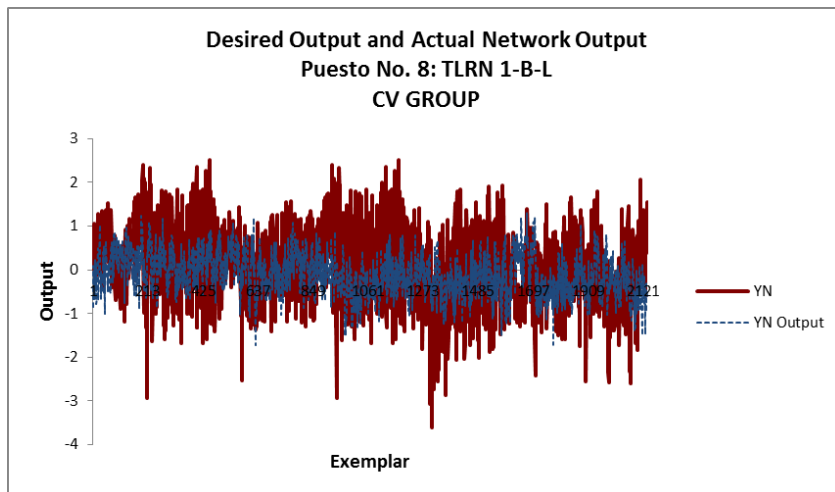
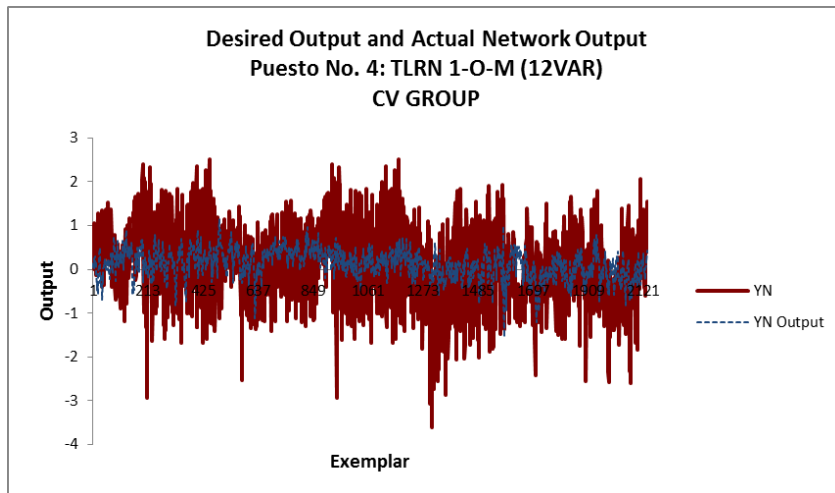
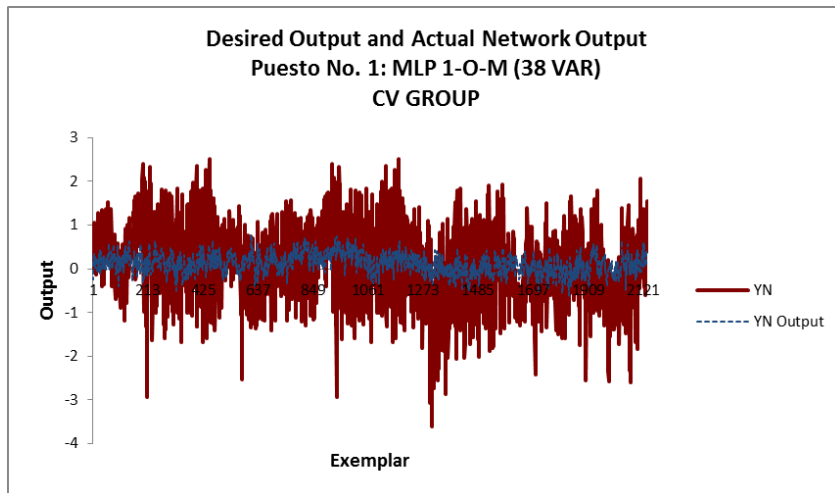
Puesto	Modelo	Variables	Model Name	Training			Cross Validation			Testing		
				MSE	r	MAE	MSE	r	MAE	MSE	r	MAE
1	87	31	MLP-2-O-M (Multilayer Perceptron)	0,71777714	0,44494925	0,66984172	0,94408251	0,52482528	0,76380092	0,87531933	0,47261153	0,74549589
2	61	7	MLP-2-O-M (Multilayer Perceptron)	0,84167086	0,36496123	0,72672569	0,94899975	0,52317342	0,7732559	1,02140953	0,43893144	0,83790753
3	62	7	MLP-2-B-M (Multilayer Perceptron)	0,73227633	0,37560173	0,66358958	1,09553806	0,51838885	0,82672315	0,89090231	0,45465715	0,75063316
4	49	7	MLP-1-O-M (Multilayer Perceptron)	0,83270995	0,37125801	0,72196484	0,97302046	0,51031001	0,78397875	1,0337225	0,43593686	0,84404389
5	52	7	MLP-1-B-L (Multilayer Perceptron)	0,75909197	0,3740305	0,68074937	1,02489286	0,50897711	0,79959987	0,98076864	0,43095362	0,80897555
6	71	7	RN-1-O-M (Recurrent Network)	0,83920498	0,36012636	0,72084596	0,96233763	0,50872818	0,77867473	0,96363666	0,44611526	0,80360194
7	76	31	LR-0-B-M (Linear Regression)	0,70474903	0,41559857	0,65558811	1,009652	0,5072763	0,78803849	0,84564067	0,47001816	0,71715793
8	65	7	GFF-1-O-M (Generalized Feedforward)	0,83755699	0,3719653	0,72345103	0,96641561	0,50679049	0,78103316	1,013256	0,42299687	0,83237789
9	91	31	GFF-1-O-M (Generalized Feedforward)	0,74061721	0,43522281	0,68145138	0,96407957	0,50592622	0,77854763	0,98493361	0,45574603	0,81060067
10	60	7	MLP-1-B-M (Multilayer Perceptron)	0,7277253	0,38176097	0,6619126	1,08487899	0,50460852	0,82008979	0,88568275	0,4507682	0,74466555

ANEXO 9: SERIES DE TIEMPO MODELOS E3

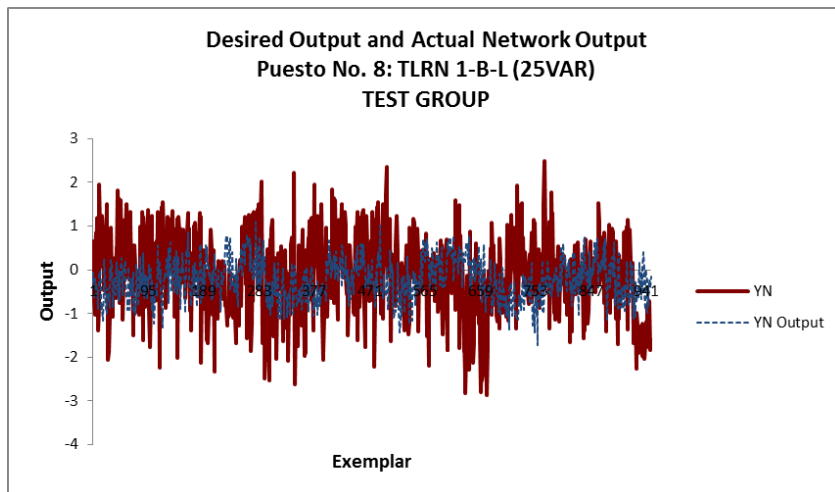
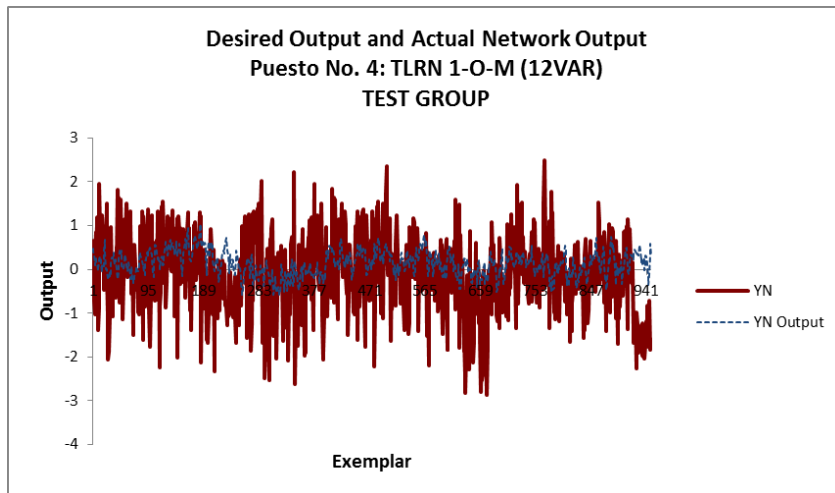
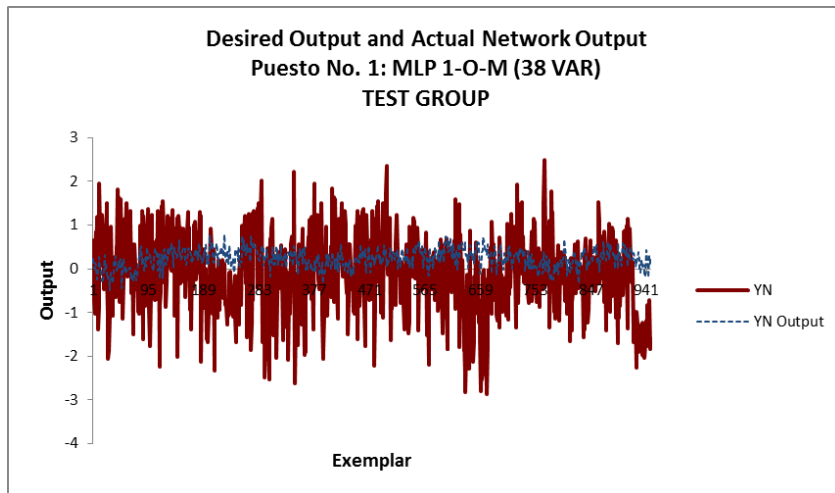
TRAIN GROUP



CV GROUP



TEST GROUP



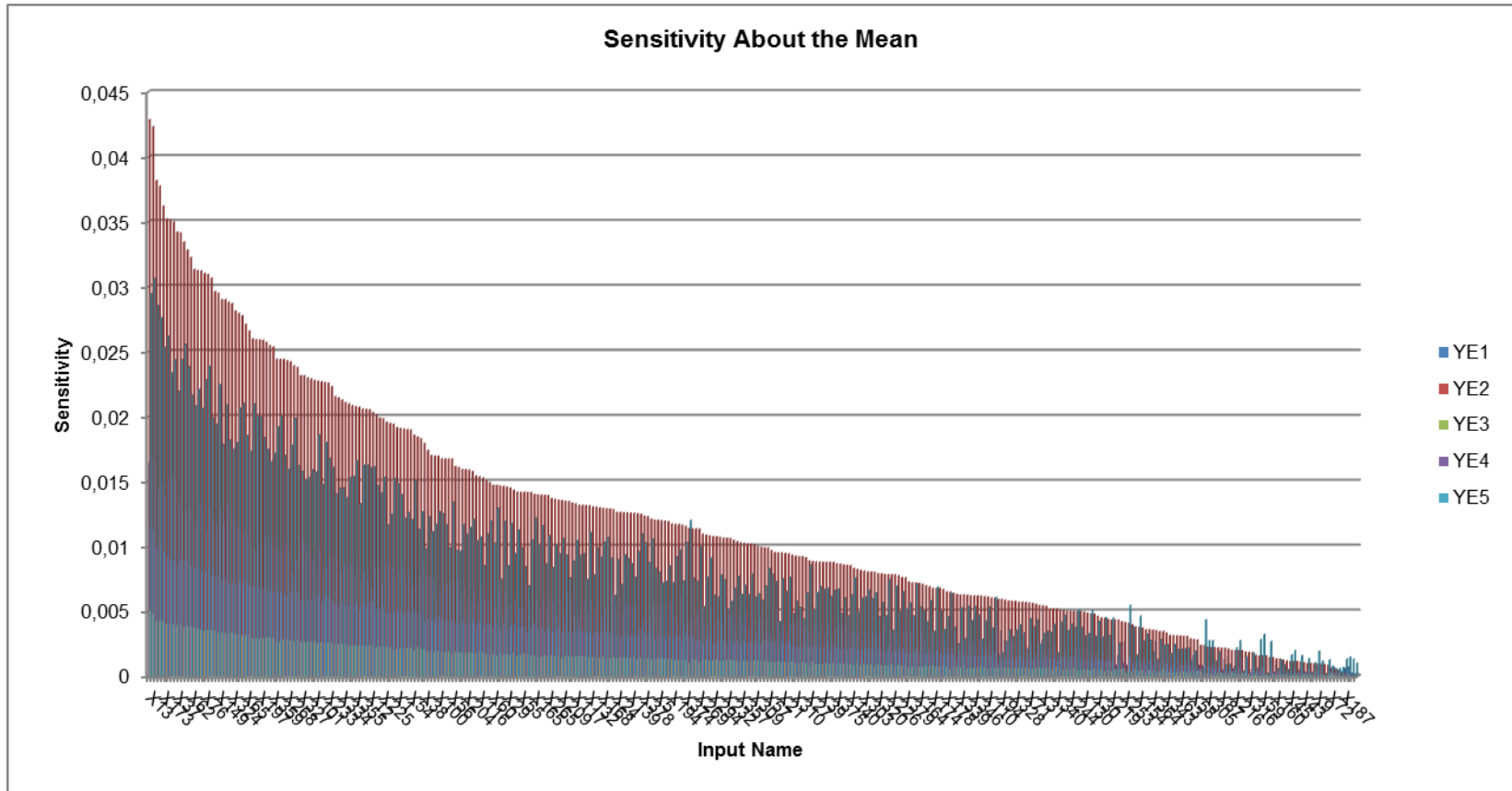
ANEXO 10: MODELOS PARA MODELO COMPUESTO

En las siguientes tablas se muestran los 9 modelos para el modelo compuesto. La tabla contiene información del puesto del modelo, el número variables de entrada, la topología de la red y los resultados de correlación y error medio en los grupos de entrenamiento, validación y prueba.

TABLA 22: MODELOS DEL MODELO COMPUESTO

Modelo	Variables	Model Name	Training			Cross Validation			Testing		
			MSE	r	MAE	MSE	r	MAE	MSE	r	MAE
1	352	MLP-1-O-M (Multilayer Perceptron)	1,4318297	0,23012938	0,98818685	1,15984413	0,18897189	0,90698214	0,95552324	0,17403304	0,78787304
2	352	MLP-2-O-M (Multilayer Perceptron)	1,22881225	0,2225597	0,906363	1,06669198	0,15242927	0,86523726	0,93689017	0,0918318	0,77932444
3	352	TLRN-1-O-M (Time-Lag Recurrent Network)	1,4276677	0,13001732	0,98403921	1,16950457	0,07818989	0,90646355	0,99180039	0,08834436	0,80681193
4	50	MLP-1-O-M (Multilayer Perceptron)	1,22820272	0,27455979	0,89490392	1,11846299	0,25530678	0,877648	0,98641213	0,09818146	0,79778098
5	50	MLP-2-O-M (Multilayer Perceptron)	1,07328508	0,2961678	0,8389112	1,02274506	0,20800933	0,83160648	0,94533843	0,08229813	0,76756787
6	50	TLRN-1-O-M (Time-Lag Recurrent Network)	1,21642403	0,2770498	0,90302564	1,1142834	0,14269072	0,88276954	0,94615551	0,05050334	0,77540886
7	25	MLP-1-O-M (Multilayer Perceptron)	1,22213738	0,30075795	0,89714266	1,14918083	0,21529062	0,89480534	0,99909112	0,04765252	0,79221492
8	25	MLP-2-O-M (Multilayer Perceptron)	1,18142352	0,30096024	0,8873654	1,08485154	0,25742272	0,87354169	0,94721645	0,07339086	0,77066374
9	25	TLRN-1-O-M (Time-Lag Recurrent Network)	1,37002321	0,34114773	0,96303913	1,20187816	0,24784877	0,9210284	1,01578193	0,11873146	0,81125361

ANEXO 11: ANALISIS DE SENSIBILIDAD MODELO COMPUESTO



La gráfica muestra la sensibilidad de la variable por estación ordenadas en orden descendente. A mayor altura mayor sensibilidad de la correspondiente variable.

ANEXO 12: CORRELACIÓN DE VARIABLES NCEP Y CGCM PARA LAS VARIABLES DE ENTRADA DE LOS MODELOS RNA

Las siguientes tablas muestran las variables de entrada a los mejores modelos por estación y el modelo compuesto. Se muestra información de la identificación de la variable, los datos de correlación con el CGCM3.1. y la variable climática que representa. En los casos que aparece NaN, se entiende que la variable no es simulada por el CGCM3.1.

TABLA 23: CORRELACIÓN VARIABLES NCEP VS. CGCM ENTRADAS E1

Variable No.	r	r2	Variable
187	0,29	0,0841	500hPa Geopotencial
178	NaN	NaN	1000hPa Componente U
179	NaN	NaN	1000hPa Componente V
181	NaN	NaN	1000hPa Dirección del viento
182	NaN	NaN	1000hPa Divergencia
183	NaN	NaN	500hPa Velocidad del viento
184	NaN	NaN	500hPa Componente U
185	NaN	NaN	500hPa Componente V
186	NaN	NaN	500hPa Vorticidad

TABLA 24: CORRELACIÓN VARIABLES NCEP VS. CGCM ENTRADAS E2

Variable No.	r	r2	Variable
78	NaN	NaN	1000hPa Componente U
279	NaN	NaN	1000hPa Componente V
249	0,5961	0,3554	1000hPa humedad específica
149	0,5316	0,2826	1000hPa humedad específica
124	0,4815	0,2318	1000hPa humedad específica
147	0,4183	0,175	500hPa Humedad específica
49	0,4076	0,1661	1000hPa humedad específica
272	0,3617	0,1308	500hPa Humedad específica
274	0,3361	0,113	1000hPa humedad específica
122	0,3352	0,1123	500 Humedad específica
237	0,325	0,1056	500hPa Geopotencial
172	0,3209	0,1029	500hPa Humedad específica

Variable No.	r	r2	Variable
41	0,2944	0,08667	850hPa Componente U
246	0,2429	0,059	850hPa Divergencia
197	0,2345	0,055	500hPa Humedad específica
131	0,2116	0,0448	1000hPa Dirección del viento
42	0,1848	0,0342	850hPa Componente V
119	0,1787	0,0319	850hPa Geopotencial
142	0,1605	0,0258	850 Componente V
117	0,1586	0,02515	850hPa Componente V
144	0,1536	0,0236	850hPa Geopotencial
53	0,1476	0,0218	1000hPa Componente U
240	0,1464	0,0214	850hPa Velocidad del viento
48	0,1447	0,0209	850hPa humedad específica
252	0,1426	0,0203	1000hPa Velocidad del viento
100	0,1113	0,0124	Temperatura 2m
31	0,1108	0,0123	1000hPa Dirección del viento
263	0,0852	0,0073	500hPa Dirección del viento
173	-0,0734	0,0054	850 Humedad específica
298	-0,0513	0,0026	850 Humedad específica
32	0,026	0,0006	1000hPa Divergencia
229	-0,016	0,0002	1000hPa Componente V
161	0,0092	8,40E-05	500hPa Vorticidad
261	-0,005	2,40E-05	500hPa Vorticidad

TABLA 25: CORRELACIÓN VARIABLES NCEP VS. CGCM ENTRADAS E3

Variable No.	r	r2	Variable
251	0,3719	0,1383	Presión nivel medio del mar
272	0,3617	0,1308	500hPa Humedad específica
253	0,3576	0,1279	1000hPa Componente U
269	0,3327	0,1107	850hPa Geopotencial
255	-0,269	0,0724	1000hPa Vorticidad
265	0,229	0,0525	850hPa Velocidad del viento
257	0,2252	0,0507	1000hPa Divergencia
258	0,1388	0,0193	500hPa Velocidad del viento
260	0,0903	0,0081	500hPa Componente V
270	0,0651	0,0042	850hPa Dirección del viento
268	-0,0531	0,0028	850hPa Vorticidad
261	-0,005	2,49E-05	500hPa Vorticidad

TABLA 26: CORRELACIÓN VARIABLES NCEP VS. CGCM ENTRADAS E5

Variable No.	r	r2	Variable
149	0,5316	0,2826	1000hPa humedad especifica
125	0,4731	0,2238	Temperatura 2m
227	0,4685	0,2195	1000hPa Velocidad del viento
251	0,3719	0,1383	Presión nivel medio del mar
172	0,3209	0,1029	500hPa Humedad especifica
137	0,3023	0,0914	500hPa Geopotencial
41	0,2944	0,0867	850hPa Componente U
273	-0,2851	0,0813	850hPa humedad especifica
66	0,2789	0,0778	850hPa Componente U
255	-0,2691	0,0724	1000hPa Vorticidad
228	0,2508	0,0629	1000hPa Componente U
271	0,2282	0,0521	850hPa Divergencia
71	0,2262	0,0512	850hPa Divergencia
165	0,2183	0,0477	850hPa Velocidad del viento
65	0,2139	0,0458	850hPa Velocidad del viento
101	0,2055	0,0422	Presión nivel medio del mar
102	0,1631	0,0266	1000hPa Velocidad del viento
216	0,1626	0,0264	850hPa Componente U
254	0,2332	0,0244	1000hPa Componente V
117	0,1586	0,0215	850hPa Componente V
252	0,1426	0,0203	1000hPa Velocidad del viento
153	0,1203	0,0145	1000hPa Componente U
75	-0,0987	0,0097	Temperatura 2m
133	0,0922	0,0085	500hPa Velocidad del viento
263	0,0852	0,0073	500hPa Dirección del viento
74	0,0754	0,0057	1000hPa humedad especifica
72	0,3371	0,0023	500hPa Humedad especifica
105	-0,0428	0,0018	1000hPa Vorticidad
229	-0,016	0,0003	1000hPa Componente V
245	-0,0058	3,41E-05	850hPa Dirección del viento
26	-0,0009	7,83E-07	Presión nivel medio del mar

TABLA 27: CORRELACIÓN VARIABLES NCEP VS. CGCM ENTRADAS MODELO COMPUESTO

Celda	Variable No.	r	r2	Variable
77X24Y	X192	NaN	NaN	850hPa Comp. V
79X24Y	X392	NaN	NaN	850hPa Comp. V

Celda	Variable No.	r	r2	Variable
79X21Y	X325	0,63707056	0,40585889	T 2m
77X22Y	X149	0,53163519	0,28263598	1000hPa Hum. Espc.
77X21Y	X125	0,47311202	0,22383498	T 2m
76X21Y	X25	0,45177915	0,2041044	T 2m
79X21Y	X324	0,44102541	0,19450341	1000hPa Hum. Espc.
78X23Y	X275	0,41061706	0,16860637	T 2m
78X23Y	X251	0,37189112	0,13830301	P
79X22Y	X347	0,35932521	0,12911461	500hPa Hum. Esp.
79X21Y	X301	0,34375587	0,1181681	P
78X22Y	X237	0,32495687	0,10559697	500hPa Geop.
79X24Y	X387	0,28286841	0,08001454	500hPa Geop.
76X23Y	X62	0,28012926	0,0784724	500hPa Geop.
76X22Y	X37	0,27984988	0,07831596	500hPa Geop.
77X23Y	X166	0,27473536	0,07547952	850hPa Comp. U
78X23Y	X254	0,23318332	0,05437446	1000hPa Comp. V
76X24Y	X99	0,18405061	0,03387463	1000hPa Hum. Espc.
76X21Y	X13	0,15820978	0,02503034	500hPa Dir. Viento
78X23Y	X252	0,14259574	0,02033355	1000hPa VV
76X24Y	X76	0,13913685	0,01935906	P
79X22Y	X338	0,10476761	0,01097625	500hPa Dir. Viento
79X23Y	X363	0,09009133	0,00811645	500hPa Dir. Viento
77X23Y	X173	-0,0733566	0,00538119	850 Hum. Esp.
76X23Y	X73	0,04841153	0,00234368	850 Hum. Esp.