



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA CIVIL

**USO DE REDES NEURONALES ARTIFICIALES
CALIBRADAS EN EL PERIODO HISTÓRICO PARA EL
PRONÓSTICO DE CAUDALES DE DESHIELO
PROYECTADOS EN EL PERÍODO 2020-2050 EN LA
CUENCA DEL RÍO MAIPO EN EL MANZANO**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL

FRANCISCO JAVIER MIRANDA ARAYA

PROFESORA GUÍA:

XIMENA VARGAS MESA

MIEMBROS DE LA COMISIÓN:

MIGUEL LAGOS ZÚÑIGA

PABLO MENDOZA ZÚÑIGA

SANTIAGO DE CHILE

2019

RESUMEN DE LA MEMORIA PARA OPTAR AL

TÍTULO DE: Ingeniero Civil Hidráulico
Sanitario y Ambiental
POR: Francisco Javier Miranda Araya
FECHA: 15/05/19
PROFESORA GUÍA: Ximena Vargas M

USO DE REDES NEURONALES ARTIFICIALES CALIBRADAS EN EL PERIODO HISTÓRICO PARA EL PRONÓSTICO DE CAUDALES DE DESHIELO PROYECTADOS EN EL PERÍODO 2020-2050 EN LA CUENCA DEL RÍO MAIPO EN EL MANZANO

La variabilidad hidrológica es un factor importante en varios ámbitos del desarrollo de países, la falta de manejo y planificación futura de este recurso puede tener repercusiones graves como escasez, déficit de energía, etc. Dicho esto, es importante conocer el comportamiento de la disponibilidad del recurso hídrico en el período de deshielo con el fin de evitar impactos negativos en su gestión, más aún si se sabe que este ya no es totalmente estacionario debido al cambio climático. En este contexto, surge la idea de analizar si un modelo estocástico de pronóstico de variables hidrológicas calibrado a base de ejemplos pasados es posible de ser usada en el futuro.

Utilizando Redes Neuronales Artificiales (desde ahora RNA) se realiza un pronóstico de caudales para la temporada de deshielo que se calibra en base a información histórica para validar su uso en el periodo 2020-2050 en la cuenca definida por la estación fluviométrica Maipo en el Manzano. Se trabaja a escala semanal, mensual y semestral para obtener los caudales medios y volumen total del período de deshielo (octubre a marzo). Para la calibración y validación de la RNA se utiliza información hidrometeorológica generada en el proyecto “Actualización del Balance Hídrico Nacional” que desde el año 2017 se lleva a cabo en el Departamento de Ingeniería Civil de la Universidad de Chile. En particular, se utilizaron las forzantes climáticas de los modelos climáticos MIROC-ESM, CSSM4, CSIRO e IPSL y los resultados de caudales obtenidos para el período futuro que se simulan en dicho estudio usando el modelo hidrológico VIC.

La calibración de las RNA resultó en valores de NSE y KGE mayores a 0,9 y NMSE menores a 0,1, además se eligieron las estructuras más parsimoniosas según el criterio de AKAIKE para cada paso temporal donde las variables más relevantes fueron el caudal (Q) y equivalente de agua en nieve (SWE). El pronóstico mostró un 90% de las temporadas con error bajo el 20% (DGA, 2018) si el tiempo de adelanto es una temporada completa (modelo 1) y 100% de las temporadas con error bajo el 20% si el tiempo de adelanto es menor (modelo 2), lo que es considerado aceptable por la DGA. Además, se encuentra una disminución importante de la incertidumbre al analizar los resultados del modelo 2. Lo anterior indica que si es factible el uso de RNA calibradas en el periodo histórico para futuros no estacionarios en la cuenca de estudio.

“And everything under the sun is in
tune but the sun is eclipsed by the
moon”

Pink Floyd

Agradecimientos

Quisiera agradecer en primer lugar a familia por haberme apoyado en este proceso, a mis padres Julio y Lidia por permitirme estudiar esta carrera sin contraer ninguna deuda.

A mi profesora guía Ximena Vargas por darme la oportunidad realizar este interesante trabajo de memoria. A los profesores Miguel y Pablo por su disposición, valiosos comentarios y sugerencias, sin duda fueron un gran aporte en este trabajo. No puedo dejar de agradecer a la gente del 3er piso de civil por facilitarme la información para realizar esta memoria en las fechas más complicadas.

A la gente de PROYES por darme trabajo y facilidades para poder titularme.

Finalmente quiero agradecer a mis compañeros que hicieron mi paso por esta carrera más placentera.

Tabla de Contenido

Introducción.....	1
1.1.- Motivación	1
1.2.- Objetivos.....	2
1.3.- Organización del Documento	2
2.- Revisión Bibliográfica.....	3
2.1.- Actualización del balance hídrico nacional (2017).....	3
2.2.- Pronósticos de Caudales de Deshielo: Estado del Arte.....	5
2.3.- Cambio Climático	7
2.3.1.- Primeros Acercamientos	7
2.3.2.- Modelos de circulación global	9
2.4.- Redes Neuronales Artificiales.....	11
2.4.1.- Neurona.....	12
2.4.2.- Arquitectura de las redes neuronales artificiales	14
2.4.3.- Ventajas de las RNA.....	15
2.4.4.- Aprendizaje o Entrenamiento de RNA	16
2.4.5.- Redes neuronales secuenciales	17
2.5.- Modelo Variable Infiltration Capacity (VIC)	18
2.6.- Análisis de Componentes Principales	22
2.7.- Métricas de Evaluación	24
2.7.1.- Error porcentual.....	24
2.7.2.- Normalized Mean Squared Error (NMSE)	25
2.7.3.- Nash-Sutcliffe Model Efficiency (NSE).....	25
2.7.4.- Kling-Gupta Efficiency.....	25
2.7.5.- Criterio de información de Akaike	26
2.8.- Validación cruzada.....	26
3.- Metodología.....	27
4.- Zona de Estudio	29
5.- Resultados.....	35
5.1.- Análisis de las variables del modelo VIC.....	35
5.2.- Determinación de los Predictores.....	40
5.2.1.- Datos de adicionales para el Entrenamiento	41

5.3.-	Pronóstico del periodo de deshielo 2020-2050	44
5.3.1.-	Bases del Pronóstico y definición de Modelos de pronóstico	44
5.3.2.-	Pronóstico Semanal.....	45
5.3.3.-	Pronóstico Mensual	52
5.3.4.-	Pronóstico Semestral.....	59
5.4.-	Desempeño de las RNA por GCM.....	62
5.4.1.-	Nivel Semanal	62
5.4.2.-	Nivel Mensual	65
5.4.3.-	Nivel Semestral	67
5.5.-	Bandas de Incertidumbre	67
6.-	Análisis de los resultados	69
7.-	Conclusiones	71
8.-	Bibliografía	72
9.-	Anexos.....	75
Anexo A.	Análisis de las forzantes y Variables VIC.....	75
Anexo B.	Ranking de los conjuntos de variables aleatorias	83
Anexo C.	Valores conocidos de variables para el pronóstico.....	86
Anexo D.	Resultados del pronóstico	87
Anexo E.	Desempeño del Pronóstico	90
Anexo F.	Errores del Pronóstico por temporada y modelo.	97
Anexo G.	Matrices de Correlación	109
Anexo H.	Anexo Digital	114
Apéndice A.	Algoritmos de Entrenamiento de RNA	115
Apéndice B.	MATLAB	121

Índice de Tablas

Tabla 2.1: Promedios temporales sobre la cuenca del río Maipo para la precipitación, escorrentía y evapotranspiración considerando 2 periodos de análisis (DGA, 2017).	4
Tabla 2.2: Estudios que han utilizado métodos estadísticos para pronosticar caudales (Easey et al., 2006).....	5
Tabla 2.3: Estudios que han utilizado modelos dinámicos para el pronóstico de caudales (Easey et al., 2006).....	5
Tabla 2.4: resultado del pronóstico de caudales de deshielo por parte de la DGA (DGA,2018)	7
Tabla 4.1: Delimitación de la cuenca alta del río Maipo (DGA)	29
Tabla 4.2: Estaciones Fluviométricas y régimen hidrológico	32
Tabla 4.3: Centrales Hidroeléctricas en la Cuenca alta del Río Maipo (DGA,2004).....	34
Tabla 5.1: Estadísticos Variables Mensuales	37
Tabla 5.2: PCA mensual	37
Tabla 5.3: Número de neuronas en cada paso temporal	40
Tabla 5.4: Set de datos para el pronóstico 2020-2050	41
Tabla 5.5: Estadísticos de las Series Históricas de Caudales medio diarios Cuenca Maipo en el Manzano periodo 1980-2015.....	42
Tabla 5.6: Variables RNA semanal modelo 1	45
Tabla 5.7: Variables RNA semanal modelo 2	45
Tabla 5.8: Resumen de temporada con error bajo el 20% en el pronóstico semanal modelo 1	46
Tabla 5.9:Resumen de temporada con error bajo el 20% en el pronóstico semanal modelo 2	46
Tabla 5.10: Indicadores de desempeño entrenamiento RNA Semanal	49
Tabla 5.11: Variables RNA Mensual escenario 1	52
Tabla 5.12: Variables RNA mensual modelo 2	52
Tabla 5.13: Porcentaje de temporadas con error menor al 20% modelo 1	52
Tabla 5.14:Porcentaje de temporadas con error menor al 20% modelo 2	53
Tabla 5.15: Indicadores de desempeño del entrenamiento RNA Mensual modelo1	56
Tabla 5.16: Arquitectura RNA Semestral	59
Tabla 5.17:Porcentaje de temporadas con error bajo el 20%	59
Tabla 5.18: Indicadores de Desempeño de entrenamiento semestral	61
Tabla 5.19: Estructuras RNA según modelo.	67
Tabla 9.1: Estadísticos Variables periodo de deshielo nivel Semanal.....	76
Tabla 9.2: PAC semanal.....	76
Tabla 9.3: Estadísticos Variable Semestral	80
Tabla 9.4: PCA Semestral	80
Tabla 9.5: Conjuntos de predictores RNA semanal.....	83
Tabla 9.6: Conjunto de predictores RNA mensual	84
Tabla 9.7: Conjuntos de predictores RNA semestral.....	85

Tabla 9.8: Valores conocidos Pp pronóstico mensual	86
Tabla 9.9: Valores Conocidos SWE pronóstico Mensual	86
Tabla 9.10: Temperaturas medias Históricas para el pronóstico mensual.	86
Tabla 9.11: Pesos Capa De Entrada RNA Mensual modelo 1	87
Tabla 9.12: Pesos y Bias Capa Oculta modelo 1	87
Tabla 9.13: Pesos Capa De Entrada RNA Mensual modelo 2	88
Tabla 9.14: Pesos de la capa de entrada RNA semanal. Modelo 1	88
Tabla 9.15: Pesos y biases neuronas capa oculta y capa de salida. RNA semanal. Modelo 1	88
Tabla 9.16: Pesos de la capa de entrada RNA semanal modelo 2	89
Tabla 9.17: Pesos y biases neuronas capa oculta y capa de salida modelo 2	89
Tabla 9.18: Pesos y Bias Capa Oculta modelo 2	89
Tabla 9.19: Pesos Capa de Entrada calibración semestral	89
Tabla 9.20: Pesos Capa oculta y salida calibración semestral	89
Tabla 9.21: Frecuencia de errores de pronóstico semanales para las proyecciones de cada GCM con RNA Modelo 1	92
Tabla 9.22: Frecuencia de Errores semanales modelo 2	93
Tabla 9.23: Frecuencia de Errores por temporada Pronóstico Semanal modelo 1	93
Tabla 9.24: Frecuencia de errores pronóstico semanal modelo 2	94
Tabla 9.25: Frecuencia de errores mensuales modelo 1. RNA calibrada con... ..	94
Tabla 9.26: Frecuencia de Errores mensuales modelo 2	95
Tabla 9.27: Frecuencias de error por temporada pronóstico mensual modelo 1	95
Tabla 9.28: Frecuencia de errores pronóstico mensual modelo 2	96
Tabla 9.29: Frecuencia de errores por temporada pronóstico semestral	96
Tabla 9.30: Error por temporada pronóstico semanal GCM CCSM4 escenario 1	97
Tabla 9.31: Error por temporada pronóstico semanal GCM CSIRO escenario 1	98
Tabla 9.32: Error por temporada pronóstico semanal GCM IPSL escenario 1	99
Tabla 9.33: Error por temporada pronóstico semanal GCM MIROC escenario 1	100
Tabla 9.34: Errores por temporada pronóstico Mensual GCM CCSM4 escenario 1	101
Tabla 9.35: Errores por temporada pronóstico Mensual GCM CSIRO escenario 1	102
Tabla 9.36: Errores por temporada pronóstico Mensual GCM IPSL escenario 1	103
Tabla 9.37: Errores por temporada pronóstico Mensual GCM MIROC escenario 1	104
Tabla 9.38: Error total por temporada pronóstico semestral GCM CCSM4	105
Tabla 9.39: Error total por temporada pronóstico semestral GCM CSIRO	106
Tabla 9.40: Error total por temporada pronóstico semestral GCM IPSL	107
Tabla 9.41: Error total por temporada pronóstico semestral GCM MIROC	108
Tabla 9.42: Correlaciones cruzadas Semestrales de todas las variables.	113
Tabla 9.43: Funciones de transferencia RNA (Artigas,2008)	122

Índice de Figuras

Figura 2.1: Resultados y comparación del Balance Hídrico Nacional en la precipitación (DGA, 2017).....	3
Figura 2.2: Resultados y comparación del balance hídrico de la temperatura media anual (DGA,2017).....	4
Figura 2.3: Efecto Invernadero (Fuente: www.ipcc.ch).....	8
Figura 2.4: Cambio en la temperatura media global. (Fuente: IPCC Quinto Informe)	9
Figura 2.5: Grilla GCM (http://www.ipcc-data.org/guidelines/pages/gcm_guide.html).....	10
Figura 2.6: Comparación entre distintas técnicas de escalamiento (DGA,2017).....	11
Figura 2.7: Neurona y sus partes (Fuente: http://www.innovacd.net/zona-cerebro/la-neurona)	12
Figura 2.8: Esquema RNA (Fuente: http://www.redes-neuronales.com.es)	13
Figura 2.9: RNA (Fuente: https://artfromcode.wordpress.com/2017/04/18/red-neuronal-en-python-con-numpy-parte-1/)	14
Figura 2.10: Arquitectura RNA a) prealimentadas, b) competitivas, c) recurrentes (García, et al ,2002).....	14
Figura 2.11: Actualización de pesos RNA usando aprendizaje supervisado	17
Figura 2.12: RNA secuencial.	17
Figura 2.13: Representación de celda modelo VIC (Gao et al,2009)	19
Figura 2.14: Aproximación de Penman Montheith (FAO,2006).....	20
Figura 2.15: Tipo de suelo según texturas.....	21
Figura 2.16: Distribución de suelos en VIC	21
Figura 2.17: Simulación de Nieve en VIC	22
Figura 2.18: Ejemplo de grafico de sedimentación.	24
Figura 3.1: Diagrama de flujo de la Metodología.....	28
Figura 4.1: Cuenca Alta del Río Maipo (Fuente: Elaboración propia)	30
Figura 4.2: Caudales Medios Anuales Cuenca Maipo en el Manzano Periodo 1970 - 201630	
Figura 4.3: Curva Hipsométrica Cuenca Alta Del Maipo (ASTER GDEM NASA: 20170814125817_550532624).....	31
Figura 4.4: Coeficientes de Pardé de las subcuencas del Rio Maipo en el Manzano 1980-2015	32
Figura 4.5: Curvas de variación estacional cuenca Maipo en el Manzano período 1980 a 2016	33
Figura 4.6: Secciones de Río Maipo	33
Figura 5.1: Caudal medio mensual	35
Figura 5.2: Precipitación Mensual	35
Figura 5.3: Evaporación Mensual.....	36
Figura 5.4: Temperatura Media Mensual.....	36
Figura 5.5: SWE medio Mensual.....	36
Figura 5.6: Serie histórica de caudales de deshielo 1985-2015.....	37

Figura 5.7: Porcentaje de Varianza explicada por cada componente principal nivel mensual	38
Figura 5.8: PACF Q medio mensual	38
Figura 5.9: Correlación cruzada Q-T	38
Figura 5.10: Correlación Q-E	39
Figura 5.11: Correlación Q-P	39
Figura 5.12: Correlación Q-SWE	39
Figura 5.13: Caudales medios semanales en Maipo en El Manzano, simulados por forzantes históricas y por forzantes obtenidas a base de modelos GCM. Periodo histórico 1985-2015	42
Figura 5.14: CVE Periodo 1985-2015 Cuenca Maipo en el Manzano	42
Figura 5.15: Curva de Duración periodo de deshielo 1985-2015 Cuenca Maipo en el Manzano	43
Figura 5.16: Regresión del entrenamiento, prueba y validación RNA semanal modelo 1 ..	47
Figura 5.17: Regresión del entrenamiento, prueba y validación RNA semanal modelo 2 ..	48
Figura 5.18: Histograma de Errores RNA semanal modelo 1	48
Figura 5.19: Histograma de Errores RNA semanal modelo 2	49
Figura 5.20: Proyección a nivel semanal GCM CCSM4 M1	49
Figura 5.21: Proyección semanal GCM CSIRO M1	50
Figura 5.22: Proyección semanal GCM IPSL M1	50
Figura 5.23: Proyección semanal GCM MIROC M1	50
Figura 5.24: Pronóstico semanal GCM CCSM4 modelo 2	51
Figura 5.25: Pronóstico semanal GCM CSIRO modelo 2	51
Figura 5.26: Pronóstico semanal GCM IPSL modelo 2	51
Figura 5.27: Pronóstico semanal GCM MIROC modelo 2	52
Figura 5.28: Regresión entrenamiento RNA Mensual modelo 1	53
Figura 5.29: Regresión entrenamiento RNA Mensual. Proyecciones GCM CSIRO. modelo 2	54
Figura 5.30: Histograma de errores entrenamiento Mensual modelo 1	55
Figura 5.31: Histograma de errores entrenamiento RNA Mensual modelo 2	55
Figura 5.32: Pronóstico mensual GCM CCSM4 modelo 1	56
Figura 5.33: Pronóstico mensual GCM CSIRO modelo 1	56
Figura 5.34: Pronóstico mensual GCM IPSL modelo 1	57
Figura 5.35: Pronóstico mensual GCM MIROC modelo 1	57
Figura 5.36: Pronóstico mensual GCM CCSM4 modelo 2	57
Figura 5.37: Pronóstico mensual GCM CSIRO modelo 2	58
Figura 5.38: Pronóstico mensual GCM IPSL modelo 2	58
Figura 5.39: Pronóstico mensual GCM MIROC modelo 2	58
Figura 5.40: Regresión del entrenamiento RNA semestral	60
Figura 5.41: Histograma de errores entrenamiento semestral	60
Figura 5.42: Pronóstico semestral GCM CCSM4	61
Figura 5.43: Pronóstico semestral GCM CSIRO	61
Figura 5.44: Pronóstico semestral GCM IPSL	62

Figura 5.45: Pronóstico semestral GCM MIROC	62
Figura 5.46:Error medio porcentual pronóstico semanal M1	63
Figura 5.47:Error medio porcentual pronóstico semanal M2.....	63
Figura 5.48: Histograma de errores por temporada pronóstico semanal modelo 1	64
Figura 5.49: Histograma de errores pronóstico semanal modelo 2	64
Figura 5.50: Error medio porcentual pronóstico mensual M1	65
Figura 5.51:Error medio porcentual pronóstico mensual M2	65
Figura 5.52: Histograma de errores por temporada pronóstico mensual modelo 1	66
Figura 5.53: Histograma de errores pronóstico mensual modelo 2	66
Figura 5.54: Histograma de errores por temporada pronóstico semestral	67
Figura 5.55: Banda de Incertidumbre de los Volúmenes acumulados en la temporada de deshielo 2020-2050.....	68
Figura 5.56:Banda de Incertidumbre de los Volúmenes acumulados en la temporada de deshielo 2020-2050 M2.....	68
Figura 9.1:Caudal Semanal Cuenca Maipo en el Manzano 1980-2015.....	75
Figura 9.2:Precipitación Semanal Maipo en el Manzano 1980-2015.....	75
Figura 9.3:SWE Semanal Cuenca Maipo en el Manzano 1980-2015	75
Figura 9.4: Temperatura Media Semanal Maipo en el Manzano 1980-2015.....	76
Figura 9.5:Evaporación Semanal Maipo en el Manzano 1980-2015.....	76
Figura 9.6:Porcentaje de Varianza explicada por cada componente principal nivel semanal	77
Figura 9.7: PACF Caudal Medio Semanal.....	77
Figura 9.8:Correlación Caudal-Temperatura Semanal	77
Figura 9.9: Correlación cruzada caudal- evaporación nivel semanal	78
Figura 9.10: Correlación caudal-precipitación nivel semanal.....	78
Figura 9.11:Correlacion caudal-swe nivel semanal.....	78
Figura 9.12: Caudales Medios Semestrales.....	79
Figura 9.13: Precipitación Semestral	79
Figura 9.14: Evaporación Semestral	79
Figura 9.15: Temperatura Media semestral.....	79
Figura 9.16: SWE medio semestral	80
Figura 9.17:Porcentaje de Varianza explicada por cada componente principal nivel semestral	81
Figura 9.18: PACF Q medio semestral	81
Figura 9.19: Correlación Q-T Semestral.....	81
Figura 9.20: Correlación Q-E semestral.....	82
Figura 9.21: Correlación Q-P semestral.....	82
Figura 9.22: Correlación Q-SWE semestral.....	82
Figura 9.23: Estructura RNA Semanal modelo 1	87
Figura 9.24: Estructura RNA Semanal modelo 2	87
Figura 9.25: Estructura RNA mensual modelo 1.....	87
Figura 9.26: Estructura RNA mensual modelo 2.....	88
Figura 9.27: Estructura RNA Semestral.....	89

Figura 9.28: Histograma de errores semanales modelo 1. RNA calibrada con GCM CCSM4	90
Figura 9.29: Histograma de errores semanales modelo 2 RNA calibrada con Set HIST	90
Figura 9.30: Histograma de errores mensuales modelo 1	91
Figura 9.31: Histograma de errores mensuales modelo 2	91
Figura 9.32: Matriz de Correlación Q-E semanal.....	109
Figura 9.33: Autocorrelación Q semanal	109
Figura 9.34: Matriz de correlación Q-P semanal.....	109
Figura 9.35: Correlación Q-T Media semanal.....	110
Figura 9.36: Correlación Q-SWE semanal.....	110
Figura 9.37: Autocorrelación Q mensual	111
Figura 9.38: Correlación Q-T mensual	111
Figura 9.39: Correlación Q-SWE mensual	112
Figura 9.40: Correlación Q-PP mensual	112
Figura 9.41: Correlación Q-E Mensual.....	113
Figura 9.42: Perceptrón (Fuente: http://ataspinar.com/2016/12/22/the-perceptron/)	115
Figura 9.43: Neurona Lineal (http://www.hugo-inc.com/RNA/Unidad%202/2.2.5.html)	115

Introducción

1.1.- Motivación

Los recursos hídricos tienen un importante impacto en gran parte de los sectores de desarrollo del país tales como la agricultura, minería, generación energética, etc. (Iniciativa Científica Milenio, 2017). Por ello es importante identificar su comportamiento a corto y largo plazo ante el cambio climático.

En ese sentido, los efectos del cambio climático tales como derretimiento acelerado de glaciares, aumento en la temperatura media, aumento o disminución de la precipitación media según la región de estudio son cada vez más notorios (Lagos et al, 2013). Para aunar esfuerzos en el estudio de este fenómeno se forma en 1988 el IPCC (Panel Intergubernamental de Expertos sobre el Cambio Climático) con el fin de que se “facilitaran evaluaciones integrales del estado de los conocimientos científicos, técnicos y socioeconómicos sobre el cambio climático, sus causas, posibles repercusiones y estrategias de respuesta” (IPCC, 1988). A la fecha el IPCC ha realizado un sinnúmero de publicaciones donde se distinguen 5 informes de evaluación general de la situación los cuales poseen 3 ejes importantes: Base de ciencia física, Impactos de adaptación y vulnerabilidad, Mitigación del cambio climático.

En Chile el impacto del cambio climático se ve reflejado en tendencias a la disminución de la precipitación en zona sur, desertificación y salinización de zonas agrícolas, insuficiencia en la oferta de agua potable y energía y, riesgos forestales entre otros, (CR2, 2018). Más aún, sus efectos amplificaron lo que se denominó la “Mega sequía (2010-2015)” que afectó principalmente a la zona centro-sur del país (Garreaud. et al, 2017).

Muchas son las medidas que se han tomado para poder adaptarse a esta situación. Por ejemplo, organismos del estado como la Dirección General de Aguas (DGA) emiten anualmente un pronóstico para la temporada de deshielos. Sin embargo, no se ha validado la utilización de dichos modelos, en general estadísticos, para un futuro no estacionario. Existen muchos métodos para la generación de pronósticos hidrológicos entre los que se encuentran modelos autorregresivos (AR), de media móvil (MA), ARMA, redes neuronales artificiales (ANN), etc. Es por lo que, en el contexto del manejo y planificación futura del recurso hídrico, ha surgido la idea de investigar la aplicabilidad futura de un modelo de redes neuronales para el pronóstico de caudales de deshielo que se calibre a base de información pasada.

Las redes neuronales artificiales son de gran utilidad debido a que levantan el supuesto de la linealidad entre las variables de entrada y salida del modelo, además de que son capaces de adaptarse a un sinnúmero de condiciones y mejorar su desempeño al obtener más información de las variables de entrada (Gonzalez,2003). Varios autores como Troncoso (1999), Muñoz (2008), Court (2008), Barría (2010) y Artigas (2015) entre otros han desarrollado estos modelos en cuencas chilenas.

1.2.- Objetivos

El objetivo principal de esta memoria es verificar si modelos de redes neuronales para el pronóstico de caudales de deshielo, construidos a base de información hidrometeorológica del período histórico (1980-2015) en la cuenca definida por la estación fluviométrica Maipo en el Manzano, pueden ser utilizados en períodos futuros (2020 -2050) usando forzantes meteorológicas (precipitación, temperatura, etc.) proyectadas por los modelos de circulación global CCSM4, CSIRO , IPSL y MIROC-ESM además de variables hidrológicas (caudales medios mensuales, cobertura nival, humedad del suelo, equivalente en agua de la nieve, etc.) simuladas por el modelo hidrológico VIC.

Los objetivos específicos son:

- Estudiar las variables hidrometeorológicas disponibles en el estudio de la DGA (2018) y analizar su influencia en el comportamiento estacional de los caudales de deshielo, para posteriormente definir cuáles son los mejores predictores (Lags) de las estructuras neuronales.
- Determinar las estructuras de red óptimas que maximizan el desempeño (minimicen el error).
- Obtener los caudales medios y volúmenes a nivel semanal, mensual y semestral mediante la aplicación de la red óptima utilizando las forzantes meteorológicas futuras disponibles (GCM) en el estudio DGA (2017) considerando los modelos climáticos seleccionados para el presente estudio.

1.3.- Organización del Documento

El capítulo 2 corresponde a la revisión bibliográfica donde se encuentra una reseña sobre el cambio climático y el uso de GCM, una revisión acerca de redes neuronales (definición y métodos de entrenamiento) para finalizar con una revisión del método de las componentes principales y métricas de evaluación de pronósticos. Más adelante en el capítulo 3, se muestra la metodología utilizada: análisis de las variables disponibles, construcción de redes neuronales, obtención de los caudales y volúmenes de deshielo.

En el capítulo 4 se presenta la zona de estudio donde se muestran las características de la cuenca de Maipo en el Manzano (Descripción general, infraestructura hidráulica, información hidrológica disponible, comportamiento estacional, etc.)

Luego, en el capítulo 5 se presentan los resultados de los objetivos propuestos en el capítulo 2, es decir, los resultados de los pronósticos a nivel semanal, mensual y semestral. Finalmente, en los capítulos 6 y 7 se presentan el análisis de los resultados y las conclusiones del presente trabajo.

2.-Revisión Bibliográfica

2.1.- Actualización del balance hídrico nacional (2017)

En el año 2017, la DGA en conjunto con especialistas de la Universidad de Chile y la Pontificia Universidad católica de Chile desarrollaron la actualización del balance hídrico nacional (DGA, 2017).

El presente proyecto tiene como objetivo desarrollar una metodología que contemple una modelación hidrológica en base física. Además, se incluyeron los potenciales efectos del cambio climático. Esta metodología se aplicó a cinco cuencas dentro de distintas macrozonas (DGA, 2017). Con motivo de cuantificar eficazmente el balance hídrico se utilizó el modelo VIC.

En la Figura 2.1 se muestran los resultados del balance hídrico al analizar la precipitación media anual (Pma), además se compara con el balance anterior. Los resultados muestran déficits en la zona norte (color amarillo), en algunas partes de la zona centro norte se observan superávits (0 a 50mm de Pma) y en las zonas sur y centro sur se observan déficits (0-50mm).

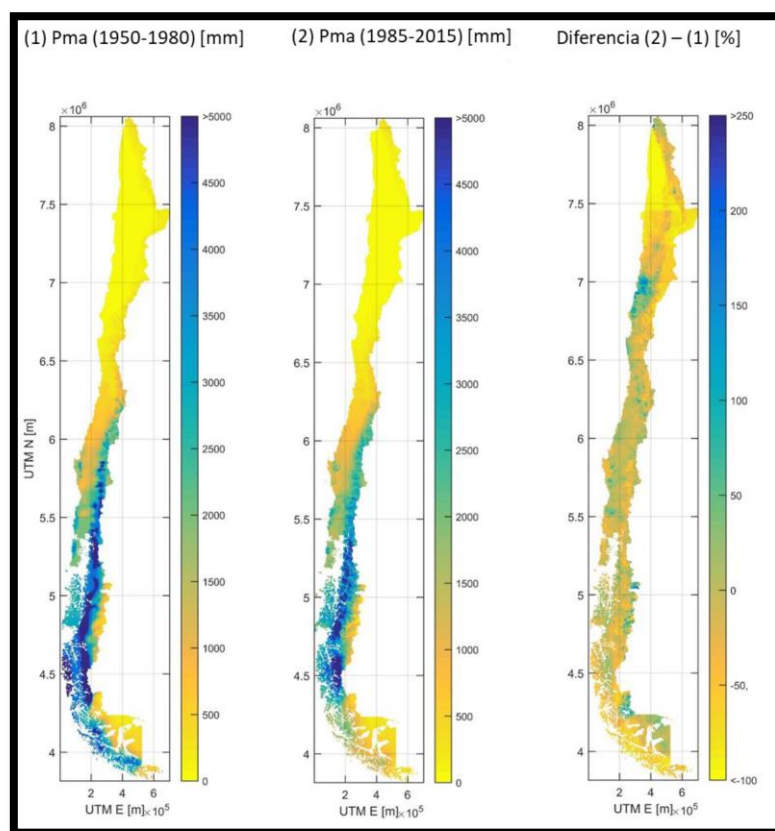


Figura 2.1: Resultados y comparación del Balance Hídrico Nacional en la precipitación (DGA, 2017).

En la Figura 2.2 se muestran los resultados del balance hídrico al analizar la temperatura media anual, además se compara con el balance anterior. Los resultados muestran disminución en la temperatura media anual en la zona norte costera (0-3°C). En las zonas centro y sur se observan diferencias de -1°C a +2°C aproximadamente.

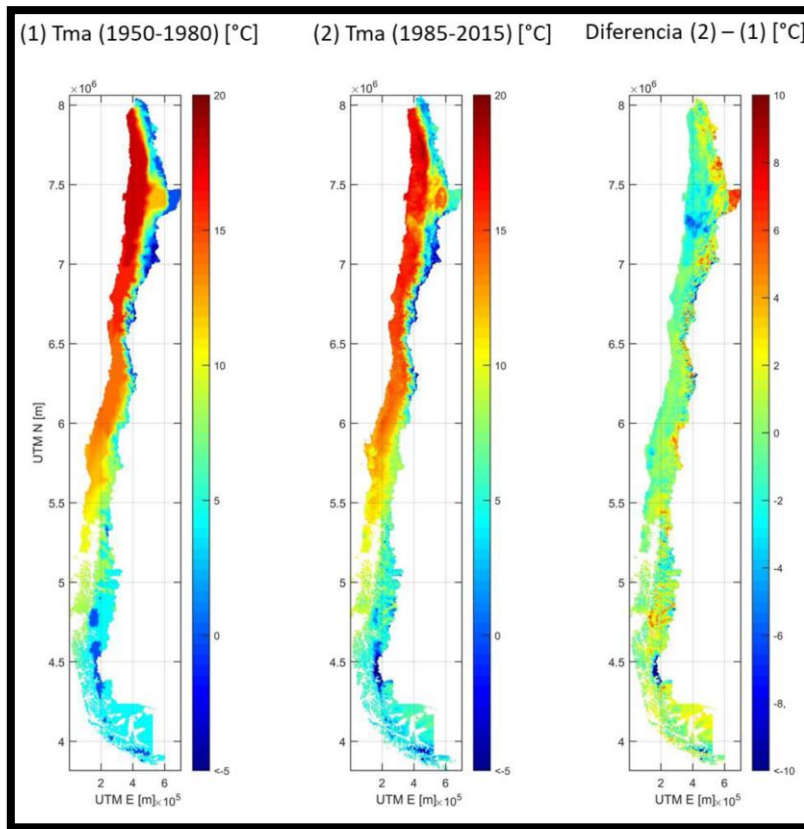


Figura 2.2: Resultados y comparación del balance hídrico de la temperatura media anual (DGA,2017)

Finalmente, en la Tabla 2.1 se muestran los resultados del balance aplicado en la cuenca del Río Maipo que será la cuenca de estudio de este trabajo.

Tabla 2.1: Promedios temporales sobre la cuenca del río Maipo para la precipitación, escorrentía y evapotranspiración considerando 2 periodos de análisis (DGA, 2017).

Cuenca Río Maipo	Precipitación Anual Promedio [mm]	Escorrentía media anual [mm]	Evapotranspiración anual promedio [mm]	Coefficiente de Escorrentía promedio
Periodo (1950-1980)	663	207	360	0,31
Periodo (1985-2015)	603	354	250	0,59
Δ	-60	147	-110	0,27
Δ [%]	-9%	71%	-31%	88%

2.2.- Pronósticos de Caudales de Deshielo: Estado del Arte.

A nivel internacional se ha desarrollado una vasta experiencia en el pronóstico de caudales estacionales, esto debido a la relevancia social que implica tener un adecuado manejo de los recursos hídricos disponibles en una región (Wedgbrow et al., 2002). Básicamente, se han desarrollado dos principales metodologías: métodos estadísticos y métodos dinámicos.

Modelos empíricos han sido desarrollados a partir de información histórica con el fin de encontrar relaciones estadísticas entre uno o más predictores sobre una variable objetivo. Las técnicas de regresión lineal (Tabla 2.2) son una de las más ocupadas (Easey et al., 2006).

Tabla 2.2: Estudios que han utilizado métodos estadísticos para pronosticar caudales (Easey et al., 2006)

Region	Method	Predictand	Most significant predictor(s)	Reference
UK	Stepwise linear regression	Monthly mean river flow	500 hPa Geopotential height, Bering Sea ice concentration and Madeira MSLP	Wilby <i>et al.</i> (2004)
UK	Regression	Summer monthly mean river flows	Winter NAOI	Wilby (2001)
UK	Regression	Regional mean summer river flow	SST and temperature differences between different regions of the North Atlantic	Svensson & Prudhomme (2005)
UK	Expert System	Summer river flow	Geopotential heights, SIC, MSLP, SST, Central England Temperature and Arctic Oscillation Index	Wedgbrow <i>et al.</i> (2005)
Australia	Categorical composites	December–May river flow	ENSO indices	Kiem & Franks (2001)
Australia	Categorical composites	August–February river flow	Southern Oscillation Index	Abawi & Dutta (1998)

Por otra parte, los modelos dinámicos son herramientas confiables para el pronóstico estacional. Algunos modelos de este tipo involucran la integración con GCM (Modelos de circulación general) de los cuales es posible obtener variables buscadas mediante la aplicación de modelos lluvia-escorrentía (Tabla 2.3).

Tabla 2.3: Estudios que han utilizado modelos dinámicos para el pronóstico de caudales (Easey et al., 2006)

Region	GCM model	Downscaling	Reference
Northeastern Brazil	Hadley Centre LDA model	Linear regression	Galvao <i>et al.</i> (1999)
USA	NCEP/NCAR reanalysis	Stepwise, multiple linear regression and RCM (RegCM2)	Wilby <i>et al.</i> (2000)
USA	NCEP Medium Range Forecast	Pacific Northwest National Laboratory RCM	Leung <i>et al.</i> (1999)
South Africa	Centre for Ocean-Land-Atmosphere Studies T30	Canonical correlation model	Landman <i>et al.</i> (2001)
Asia	ERA-15	None	Schar <i>et al.</i> (2004)

A nivel nacional, se han realizado varias investigaciones acerca del pronóstico de caudales de deshielo. Destacan los trabajos de tesis de Gonzales (1970) y Espinoza (1973), en las cuales se desarrollaron modelos estocásticos en bases a registros de precipitación y temperatura. Más recientemente, se han desarrollado modelos de pronóstico utilizando redes neuronales artificiales (RNA) en distintas instancias, como por ejemplo en tiempo real (Gonzalez, 2003), pronóstico de caudales diarios (Barría, 2010), mensuales (Artigas, 2015), semestral, etc.

En adición a lo anterior, la DGA elabora anualmente un pronóstico de volúmenes de deshielo, caudales medios mensuales y máximos instantáneos para la temporada de riego del periodo primavera-verano. Este pronóstico tiene como propósito dar a conocer la situación hidrológica general del país, específicamente en las zonas de riego de las principales cuencas del norte chico, zona central y sur. Estos pronósticos se realizan desde la cuenca del río Copiapó hasta la cuenca del río Ñuble (DGA, 2018).

Las bases del pronóstico son los siguientes (DGA, 2018):

- El período de pronóstico es el comprendido entre los meses de septiembre a marzo, debido a la importancia que éstos tienen en la agricultura de riego y está orientado a estos usuarios.
- Se pronostican los volúmenes para la temporada de deshielo, los que se entregan junto con la probabilidad de excedencia o tipo de año. La distribución mensual de caudales medios puede tener fluctuaciones apreciables en la medida que las variables meteorológicas, tales como precipitación, temperatura, radiación y nubosidad, presenten comportamientos irregulares en el período de deshielo. Con mayor razón pueden presentarse variaciones muy significativas a nivel diario, en relación con el valor medio mensual, por lo que este pronóstico no es estricto para aquellos usuarios cuyas necesidades de caudales diarios son determinantes para su producción. Ante estos requerimientos, los interesados debieran desarrollar sus propias metodologías para satisfacer sus necesidades particulares.
- Los pronósticos se realizan para las cuencas señaladas en el informe, por lo cual los resultados no se aplican a cuencas más pequeñas, intermedias o de secano, que presentan condiciones hidrológicas diferentes. En estos últimos casos habrá que hacer análisis particulares.
- Se consideran como variables independientes los datos registrados hasta el mes de agosto y para primavera-verano se han supuesto precipitaciones promedio. Por lo tanto, los caudales pronosticados pueden resultar distintos de los reales si se producen durante dicho período, precipitaciones muy diferentes a la situación supuesta.

En la Tabla 2.4 se presentan los resultados del último pronóstico de caudales y volúmenes de deshielo (DGA, 2018):

Tabla 2.4: resultado del pronóstico de caudales de deshielo por parte de la DGA (DGA, 2018)

ESTACION	REGION	Sep	Oct	Nov	Dic m ³ /s	Ene	Feb	Mar	Vol. mill-m ³	Vol.medio mill-m ³	Vol / Vol.medio	Prob. Exc. %
Copiapó en Pastillo	III	2.0	1.8	1.7	1.5	1.6	1.5	1.5	30	47	0.65	43
Huasco en Algodones	III	4	4	4	4	4	4	3	70	178	0.40	41
Elqui en Algarrobal	IV	5.0	5.0	5.5	5.0	4.5	4.0	3.5	85	260	0.33	72
Hurtado en San Agustín	IV	1.5	1.3	1.1	1.0	1.0	1.0	0.8	20	71	0.29	87
Grande en las Ramadas	IV	1.7	1.7	1.7	1.5	1.0	1.0	1.0	25	110	0.23	79
Choapa en Cuncumén	IV	4.5	5.0	7.5	5.0	3.5	3.0	2.0	80	262	0.31	90
Aconcagua en Chacabuquito	V	9	12	27	30	25	18	12	350	893	0.39	91
Putando Resguardo Patos	V	2.0	3.5	4.5	3.5	2.0	1.7	1.7	50	210	0.24	87
Juncal en Juncal	V	2.0	3.0	5.0	6.5	7.0	6.0	4.8	90	151	0.60	94
Mapocho en los Almendros	RM	3.0	3.5	3.5	3.0	2.5	2.0	1.5	50	155	0.32	94
Maipo en el Manzano	RM	55	55	90	110	95	75	53	1400	2914	0.48	94
Colorado antes junta Maipo	RM	11	13	25	35	40	27	20	450	816	0.55	95
Cachapoal en Puente Termas	VI	30	50	93	111	95	80	55	1350	2556	0.53	90
Claro en Hacienda Las Nieves	VI	3	5	7	7	5	4	3	90	171	0.53	92
Tinguiririca en B. Briones	VI	18	25	53	55	50	40	25	700	1227	0.57	92
Teno despues de Junta	VII	25	45	66	60	35	20	15	700	1200	0.58	90
Claro en Los Queñes	VII	11	14	16	12	8	6	5	190	324	0.59	91
Maule en Armerillo	VII	220	250	325	235	130	95	77	3500	4598	0.76	82
Ñuble en San Fabián	VIII	150	120	100	60	30	20	15	1300	1664	0.78	75

2.3.- Cambio Climático

En el año 1988 se funda el IPCC que a la fecha ha publicado 6 informes sobre el cambio climático. El IPCC define el cambio climático como “cambios en el estado del clima que pueden ser identificados (ej.: utilizando métodos estadísticos) en la media y/o la variabilidad de sus propiedades y que además persisten por un periodo de tiempo extendido (décadas o más). El cambio climático puede ser producto de procesos internos o por forzantes externas como ciclos solares, erupciones volcánicas y cambios en la composición de la atmosfera producto de la actividad antrópica” (IPCC, 2007, Glosario).

2.3.1.- Primeros Acercamientos

Como primer acercamiento el IPCC (IPCC, 1990) aceptó la existencia de un efecto invernadero natural (Figura 2.3) que hace que la tierra sea más cálida. El aumento de las emisiones antrópicas tiene como consecuencia la producción de gases (anhídrido carbónico, *clorofluorocarbonos* y óxido nitroso) lo cual potencia fuertemente el efecto invernadero y como consecuencia el aumento del vapor de agua.

Las concentraciones atmosféricas de estos gases reaccionan lentamente a las variaciones de sus emisiones. Es decir, si se sigue con el ritmo constante las repercusiones se dejarán sentir durante siglos.

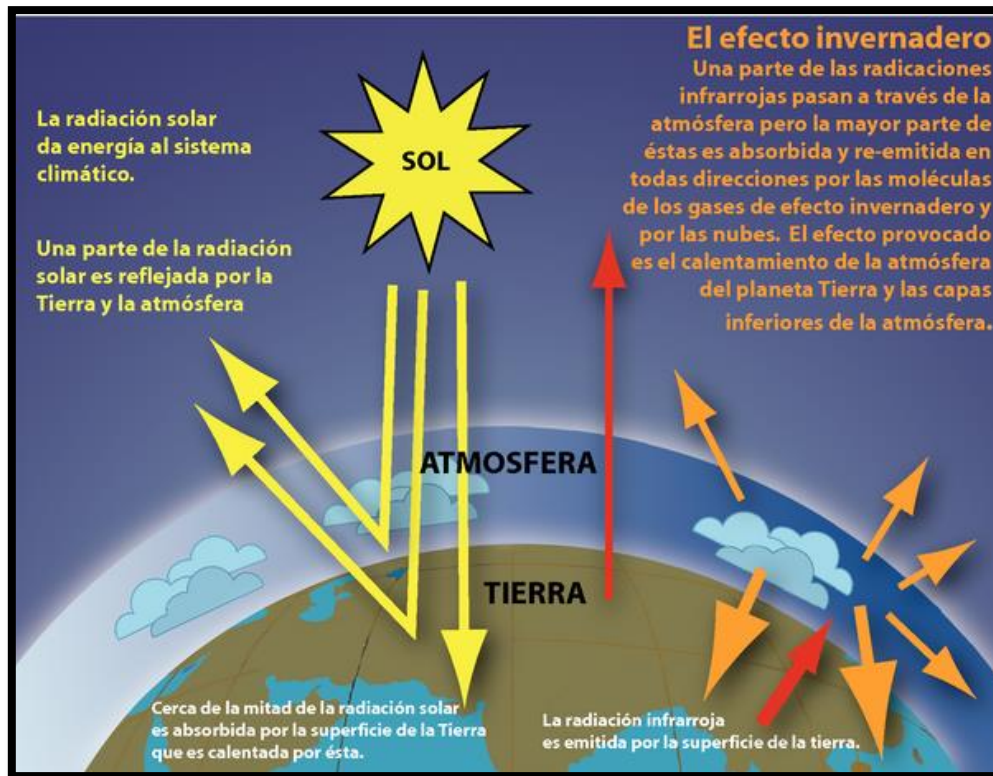


Figura 2.3: Efecto Invernadero (Fuente: www.ipcc.ch)

En el año 2013 se definieron nuevos escenarios llamados RCPs (IPCC, 2013). Estos escenarios se centran en las emisiones antropogénicas y no incluyen cambios en impulsores naturales como el forzamiento solar o volcánico o las emisiones naturales de CH_4 o N_2O :

- **RCP2.6:** Mitigación estricta del cambio climático, representa un escenario donde se planea mantener el calentamiento global a menos de 2°C sobre los valores del periodo preindustrial. Supone un cambio drástico en las políticas medioambientales, emisiones de metano (CH_4) reducidas en 40%. En este escenario se encuentran los menores valores de emisión de contaminantes.
- **RCP4.5:** Escenario desarrollado por el Pacific Northwest National Laboratory's Joint Global Change Research Institute (JGCRI) en EE. UU, es un escenario de estabilización en la cual la forzante radiativa se estabiliza después del año 2100.
- **RCP6:** es un escenario de estabilización en la cual la forzante radiativa se estabiliza después del año 2100 mediante la utilización de nuevas tecnologías y la reducción de gases de efecto invernadero.
- **RCP8.5:** Escenario caracterizado por el aumento de la emisión de gases de efecto invernadero, análogo a escenarios donde se dejan aumentar dichos gases, es decir, escenarios desfavorables.

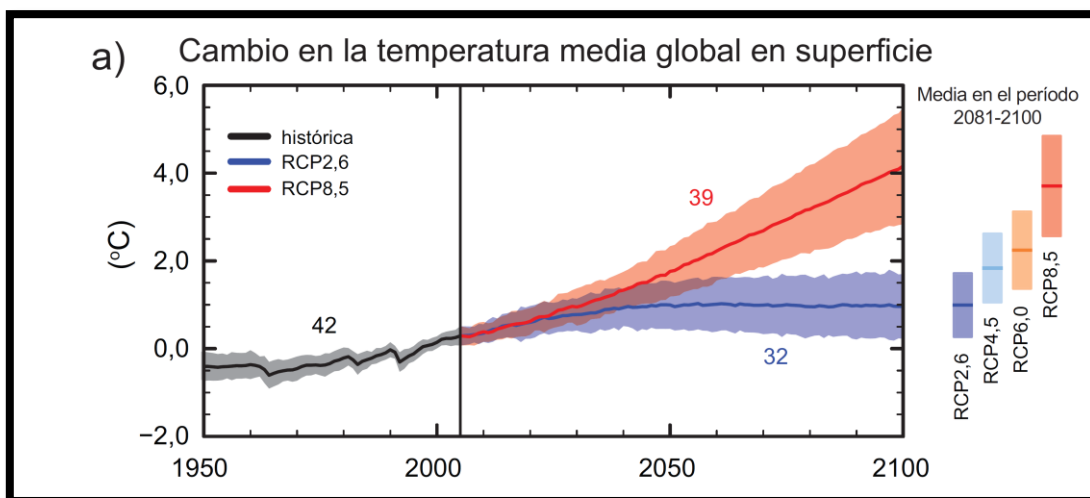


Figura 2.4: Cambio en la temperatura media global. (Fuente: IPCC Quinto Informe)

En la Figura 2.4 se muestran las series temporales del cambio anual de la temperatura media global en superficie con respecto al horizonte 1986-2005 simuladas en el marco del quinto informe del IPCC (IPCC,2013). Con color rojo se muestra la serie de tiempo del escenario **RCP8,5** y con azul la serie del escenario **RCP2,6**. Así mismo, las áreas sombreadas corresponden a las incertidumbres asociadas. En color negro por otra parte, se muestra la evolución histórica de los escenarios utilizando forzantes históricas reconstruidas. Finalmente, en el extremo derecho de la figura se muestran las medias de las series temporales en el periodo 2081-2100.

2.3.2.- Modelos de circulación global

Los modelos de circulación global son modelos climáticos que se caracterizan por representar procesos físicos en la atmósfera, océanos, criosfera y en la superficie terrestre. Además, son una de las herramientas más avanzadas para simular la respuesta del cambio climático ante el aumento de los gases del efecto invernadero. Estos modelos se representan grillas (IPCC, 2009). Estas grillas, por lo general poseen una resolución horizontal de 250 y 600 km, verticalmente se tienen por lo general 10 a 20 capas. Cabe mencionar que existe una extensa lista (http://www.ipcc-data.org/sim/gcm_global/index.html) de GCM desarrollados por científicos.

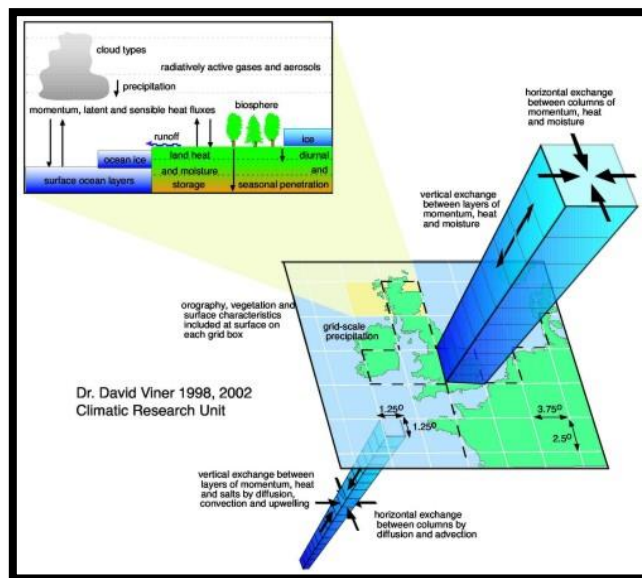


Figura 2.5: Grilla GCM (http://www.ipcc-data.org/guidelines/pages/gcm_guide.html)

Con el fin de obtener la información proveniente de los GCM en la resolución buscada, se debe realizar el proceso de escalado, este proceso consta de 2 partes:

- Escalado espacial: se realiza bajo el supuesto de que aquellos puntos de información que están más cercanos al punto de interés tienen mayor preponderancia que aquellos que están más lejos. Para este procedimiento se puede realizar un escalamiento proporcional al inverso cuadrado de la distancia.
- Escalado temporal: Con el fin de representar de la manera más exacta posible las proyecciones de variables, se necesita realizar un ajuste de funciones de transferencia entre los datos proyectados en el periodo de línea base para el punto de interés, con respecto a los datos meteorológicos disponibles en el área de estudio. Este escalamiento puede ser estadístico (Wood et al., 2002) o dinámico. Distintos tipos de escalamiento, ventajas y desventajas se observan en Figura 2.6.

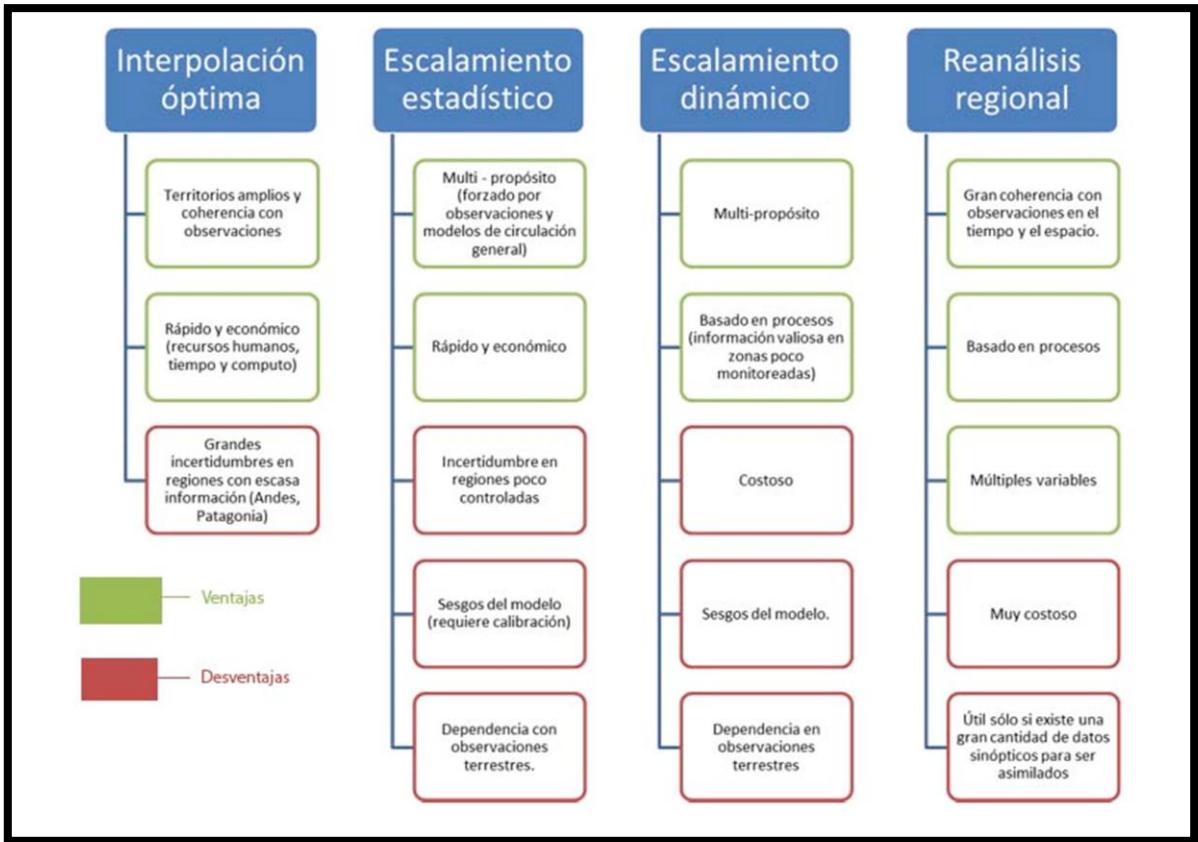


Figura 2.6: Comparación entre distintas técnicas de escalamiento (DGA, 2017)

2.4.- Redes Neuronales Artificiales

En el marco del “*Machine Learning*” (Brynjolfsson, 2017) las redes neuronales artificiales se utilizan en la simulación de actividades cognitivas complejas mediante métodos computacionales inspirados en el funcionamiento del sistema nervioso central humano.

El cerebro humano está compuesto por billones de neuronas (Figura 2.7) interconectadas entre si formando redes que tienen funciones especiales. Estas neuronas recogen señales de otras neuronas a través de estructuras llamadas dendritas. Las neuronas emiten impulsos a lo largo de una fibra denominada axón la cual tiene millones de ramificaciones. La conexión neurona-neurona se produce gracias a la sinapsis (conexión ramificación axonal-dendrita) donde además se transforma el impulso eléctrico en una señal neuroquímica mediante la liberación de los neurotransmisores (Purves, D. et al, 2001).

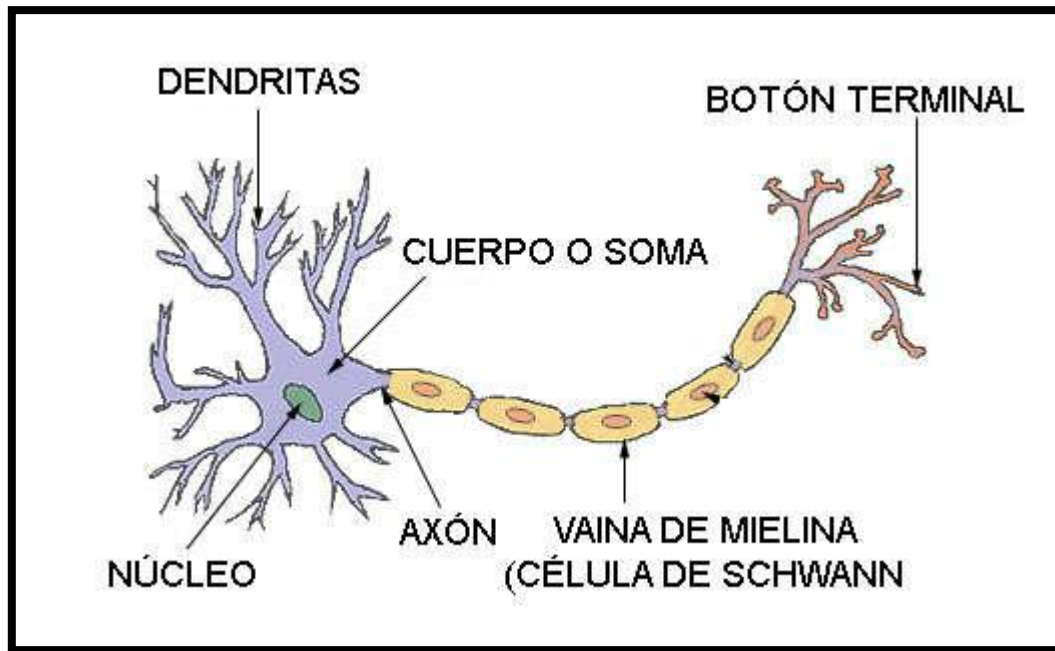


Figura 2.7: Neurona y sus partes (Fuente: <http://www.innovacd.net/zona-cerebro/la-neurona>)

Los neurotransmisores tienen 2 efectos: **excitatorios o inhibitorios** los cuales son variables, de manera que se habla que la efectividad de la sinapsis. Las señales recibidas por una neurona se combinan y en función de la **estimulación total recibida**, la neurona alcanza su **nivel de activación**, lo cual se traduce en la **frecuencia** con la cual emite las señales.

El aprendizaje de la red neuronal se produce mediante la variación de la efectividad de la sinapsis, de modo que también varía el grado de influencia de una neurona sobre otra, así, es posible reconocer la arquitectura de la red.

2.4.1.- Neurona

Las redes neuronales artificiales (RNA) se utilizan en la simulación de actividades cognitivas complejas mediante métodos computacionales. El elemento básico de las RNA es la neurona. En la Figura 2.8 se puede observar la estructura de una neurona donde las entradas X_i representan la información entregada por las otras neuronas, las cuales se ponderan por su peso (influencia W_i), luego la información ponderada entra a la función de transferencia resultando en las variables de salida.

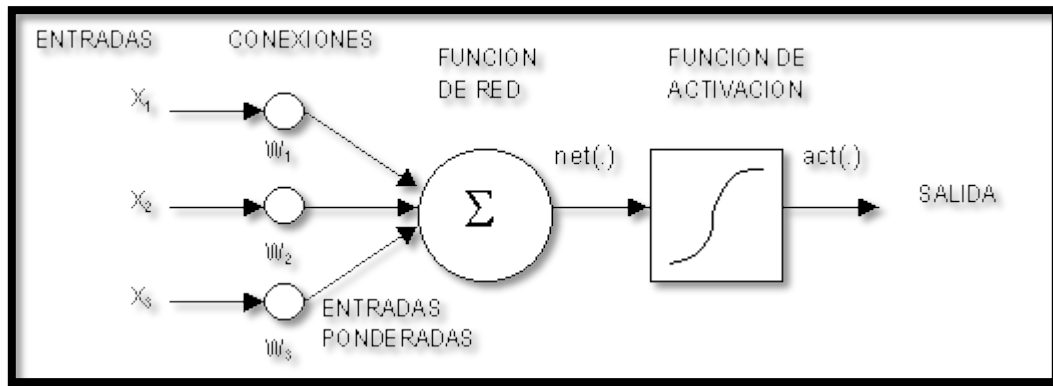


Figura 2.8: Esquema RNA (Fuente: <http://www.redes-neuronales.com.es>)

La función de red está definida como:

$$F_{red} = \sum w_i p_i + b \quad (1)$$

Donde \$w_i\$ son los pesos, \$p_i\$ corresponden a las entradas y \$b\$ es el escalar “bias” que es similar a una entrada con un peso igual a 1 (Hudson, 2015).

Existen variadas funciones de transferencia (Figura 9.43), donde las principales son la lineal, escalón y sigmoide (Hudson, 2015):

1. Lineal: las funciones de transferencia de este tipo son utilizadas en la capa final.
2. Log-Sigmoidea: comúnmente utilizada en la capa oculta, en parte por ser una función diferenciable.

En la Figura 2.9 se puede observar una RNA donde se distinguen 3 capas:

1. Capas de entrada: Reciben la información del exterior.
2. Capas Ocultas: Procesan, memorizan y transmiten la información.
3. Capas de salida: Procesan la información desde las capas ocultas y entregan la salida del modelo.

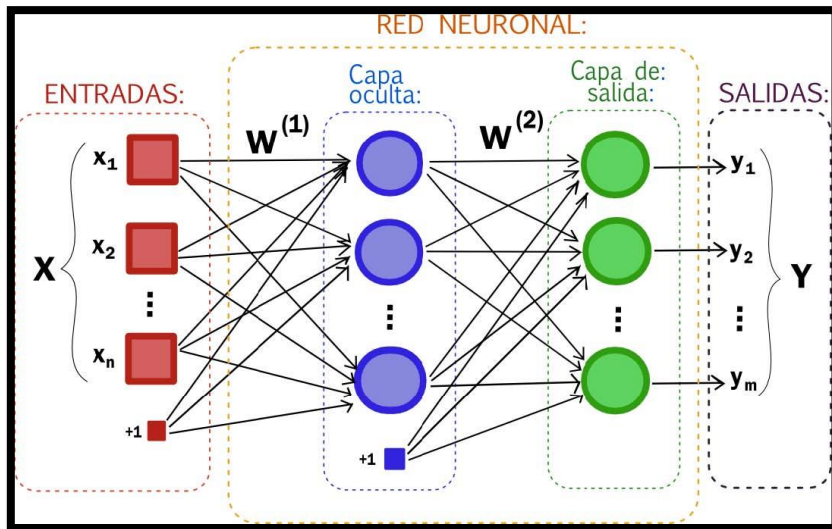


Figura 2.9: RNA (Fuente: <https://artfromcode.wordpress.com/2017/04/18/red-neuronal-en-python-con-numpy-parte-1/>)

2.4.2.- Arquitectura de las redes neuronales artificiales

La arquitectura de la red neuronal consiste en la disposición de las neuronas, es decir, la cantidad de capas y la dirección en que viaja la información la cual puede ser unidireccional (red estática) o bidireccional (red dinámica). Usualmente se habla de los siguientes tipos (Figura 2.10):

- Redes prealimentadas o feedforward: son redes con conexiones unidireccionales y totalmente conectadas, donde la información fluye desde la capa de entrada a la de salida.
- Redes competitivas: son redes dinámicas donde las neuronas de la misma capa poseen conexiones bidireccionales.
- Redes recurrentes o feedback: redes donde las capas pueden crear conexiones con las neuronas de las capas previas.

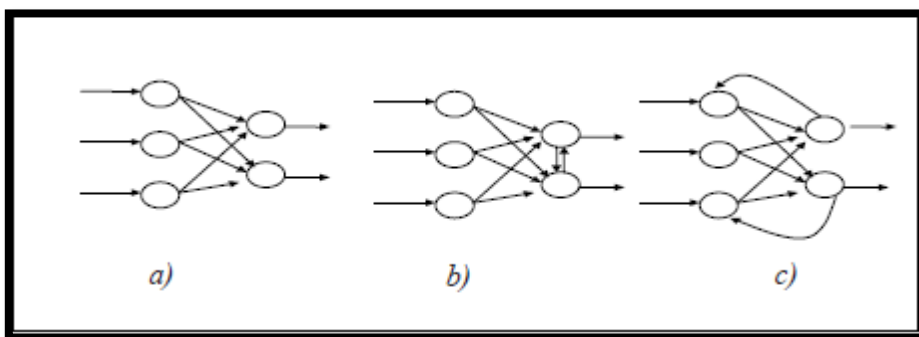


Figura 2.10: Arquitectura RNA a) prealimentadas, b) competitivas, c) recurrentes (García, et al ,2002)

2.4.3.- Ventajas de las RNA

Las RNA tienen múltiples ventajas:

- Representación de las relaciones entrada y salida: Mediante algoritmos de entrenamiento es posible optimizar los pesos de la red dada una secuencia de entradas y salidas deseadas de manera de obtener los resultados deseados (Gonzalez, 2003).
- Capacidad de generalización: Generalmente se utiliza un set de datos distintos en los procesos posteriores al entrenamiento, de manera que el correcto resultado de la red en los procesos de validación y prueba dependerá de la capacidad de generalización de la red (Gonzalez, 2003).
- No linealidad: Los procesos dentro de la neurona son generalmente no lineales. La conexión de los distintos procesos transmite esta no linealidad a lo largo de la red (Muñoz, 1996).
- Adaptabilidad: Las RNA son estructuras adaptativas debido a sus diferentes funciones de transferencia y actualización de pesos (Muñoz, 1996).
- Procesamiento paralelo: La estructura y modo de operación de las RNA las hace adecuadas para el procesamiento paralelo real (Gomez, 1999).
- Memoria Distribuida: el conocimiento acumulado por la red se encuentra distribuido en la totalidad de sus conexiones, lo que tiene como consecuencia la tolerancia a los fallos. Así, la RNA es capaz de funcionar a pesar de presentar algunos fallos en neuronas y conexiones ya que la información se encuentra distribuida por toda la red. Las RNA son capaces de abstraer las características esenciales de las entradas, de manera que son capaces de procesar correctamente datos incompletos o distorsionados (Gómez, 1999).

2.4.4.- Aprendizaje o Entrenamiento de RNA

El aprendizaje o entrenamiento consiste en la presentación de patrones y la posterior actualización de los pesos de las conexiones siguiendo algún método de aprendizaje la cual tiene como objetivo minimizar el error. Los métodos de entrenamiento se presentan en el Apéndice A

La calidad del aprendizaje dependerá exclusivamente de la arquitectura de la red y el algoritmo de entrenamiento seleccionado. Algunos tipos de entrenamientos son:

1. Aprendizaje supervisado: La red es alimentada con información acerca de cómo se está resolviendo el problema. Requiere un conjunto de datos entrada/salida deseados.
2. Aprendizaje no supervisado: No se requiere ingresar patrones de salida deseados, de manera que, al presentar un patrón aprendido o parecido, se producirá la misma salida.
3. Aprendizaje Reforzado: Se utiliza cuando no existe un conjunto de datos de entrenamiento ni de datos de salida deseados. La única información disponible es un escalar que indica el desempeño de la red neuronal con un retardo, ya que se obtiene de acuerdo con las salidas ya entregadas. La red se ajusta forzándola con una predicción de esta evaluación, la cual se utiliza además para ajustar la misma predicción.

2.4.4.1.- Aprendizaje Supervisado

Generalmente el entrenamiento requiere una gran cantidad de datos entrada/salida del problema a modelar. Usualmente, los datos son separados siguiendo esta distribución:

- Entrenamiento (50%): Ajuste de pesos para lograr la salida requerida.
- Validación (25%): Monitoreo el proceso de entrenamiento.
- Prueba (25%): Observar el desempeño de la red.

El aprendizaje supervisado necesita de datos de entrada y salida, esta regla de aprendizaje busca minimizar el error cuadrático medio (ECM) entre la salida deseada y la observada. El algoritmo es el siguiente (Minsky and Papert, 1972):

1. Asignar en forma aleatoria los pesos de todas las conexiones.
2. Seleccionar un par de entrenamiento (patrón de entrada y salida deseada)
3. Presentar el patrón de entrada y calcular la salida total de la red.
4. Se calcula el error entre la salida obtenida y la deseada, el error se calcula como la suma de los cuadrados de las diferencias entre las salidas mencionadas.
5. Ajustar los pesos
6. Volver al paso 2

Los pesos pueden ser ajustados de la siguiente manera (Figura 2.11):

- Actualizar cada vez que se le presente un ejemplo a la RNA
- Actualizar una vez que se lo mostraron todos ejemplos a la RNA
- Actualizar los pesos cada vez que se le presenten ciclos de datos.

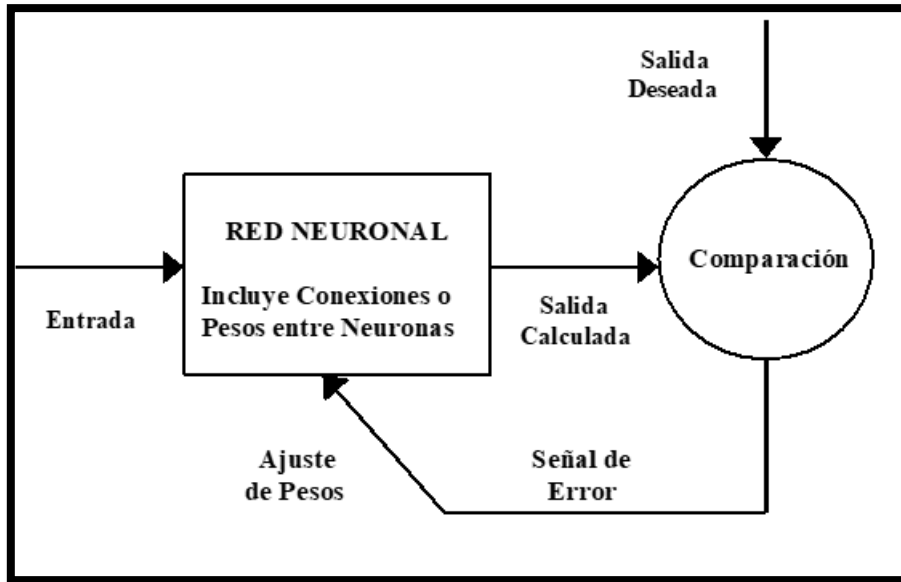


Figura 2.11: Actualización de pesos RNA usando aprendizaje supervisado

2.4.5.- Redes neuronales secuenciales

En un intento por mejorar la calidad de los pronósticos de caudales para tiempos de adelanto mayores es que algunos autores (Prakash, et. al 2014) aplicaron el concepto de redes neuronales secuenciales que consiste básicamente en la construcción de varias RNA unidas secuencialmente. Dado que la eficiencia de las RNA resulta mejor cuando los tiempos de pronóstico son pequeños, este método constituye un gran avance en cuanto a mejorar la eficiencia para tiempos de pronóstico mayores. La estructura de la RNA secuencial se observa en la Figura 2.12:

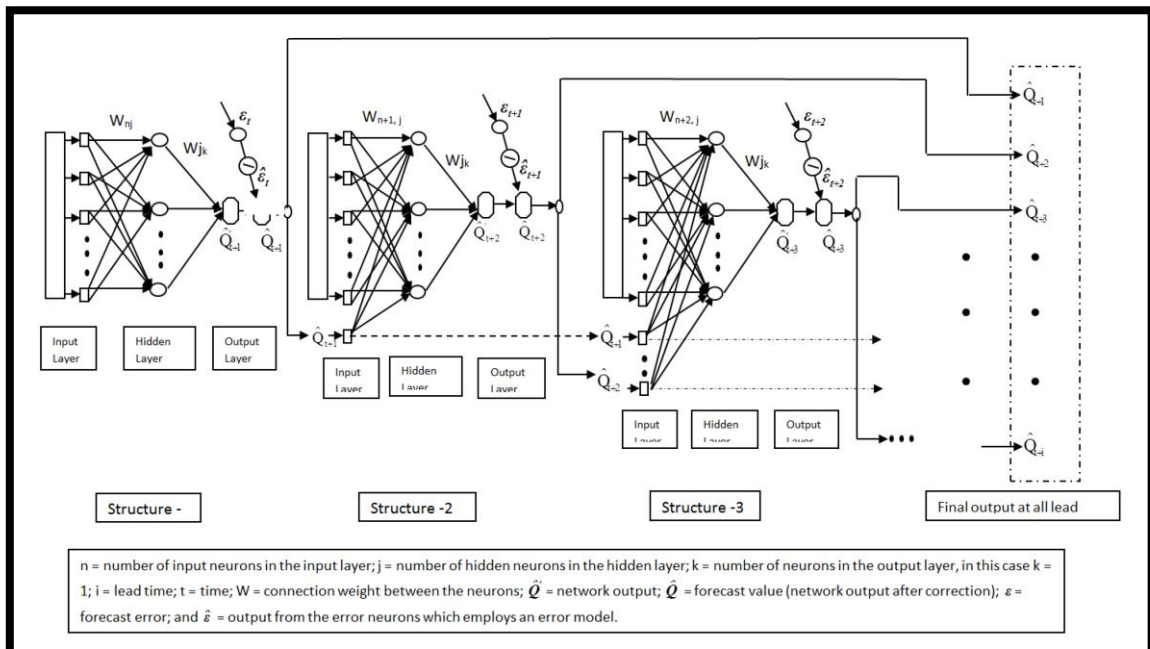


Figura 2.12: RNA secuencial.

El procedimiento consiste en construir varias RNA, al construir la primera se corrige la salida con función que reduce la varianza. Esta función generalmente sigue una distribución normal, función cuadrática o cubica.

Corregida la salida de la primera red se construye la segunda agregando en las entradas los resultados de la red antecesora y el proceso avanza sucesivamente hasta lograr el tiempo de pronóstico deseado.

Como se dijo anteriormente, la ventaja sobre las RNA convencionales radica en que la eficiencia de redes con tiempo de pronóstico pequeño es mejor que para tiempos mayores, además de que en cada iteración se reduce la varianza de los resultados de la red aumentando la calidad del pronóstico.

2.5.- Modelo Variable Infiltration Capacity (VIC)

Variable Infiltration Capacity **VIC** (Liang et al., 1994) es un modelo hidrológico de macro escala el cual resuelve balances hídricos y de energía, Originalmente desarrollado por Xu Liang en la Universidad de Washington, VIC es un modelo de investigación y en sus variadas formas ha sido aplicado a variadas cuencas incluyendo las del Río Columbia, Río Ohio, Río Arkansas-Red y los Ríos del alto Mississippi, así como también en otros lugares del mundo. El desarrollo y mantención de la versión oficial actual de VIC es manejada por el Departamento de Ingeniería Civil y Ambiental de la Universidad de Washington, bajo la dirección de Dennis P. Lettenmaier. En VIC la unidad son celdas, y puede ser asimilada como el equivalente a la unidad de respuesta hídrica, ya que VIC exige la independencia de flujos entre cada celda a ser ingresada, realizando un balance energético en cada celda independientemente.

El paso de tiempo utilizado por el modelo depende de si el balance realizado es de masa o energético, siendo el valor del paso de tiempo de un día en el primer caso, y menor a un día en el segundo.

La Figura 2.13 muestra la estructura de cada celda en VIC. De esta manera, la cuenca modelara debe estar grillada con un tamaño de celda definido. Así, cada celda simula procesos hídricos y energéticos, por lo que cada una de estas debe contar con todos los parámetros forzantes necesarios para la correcta simulación de los procesos.

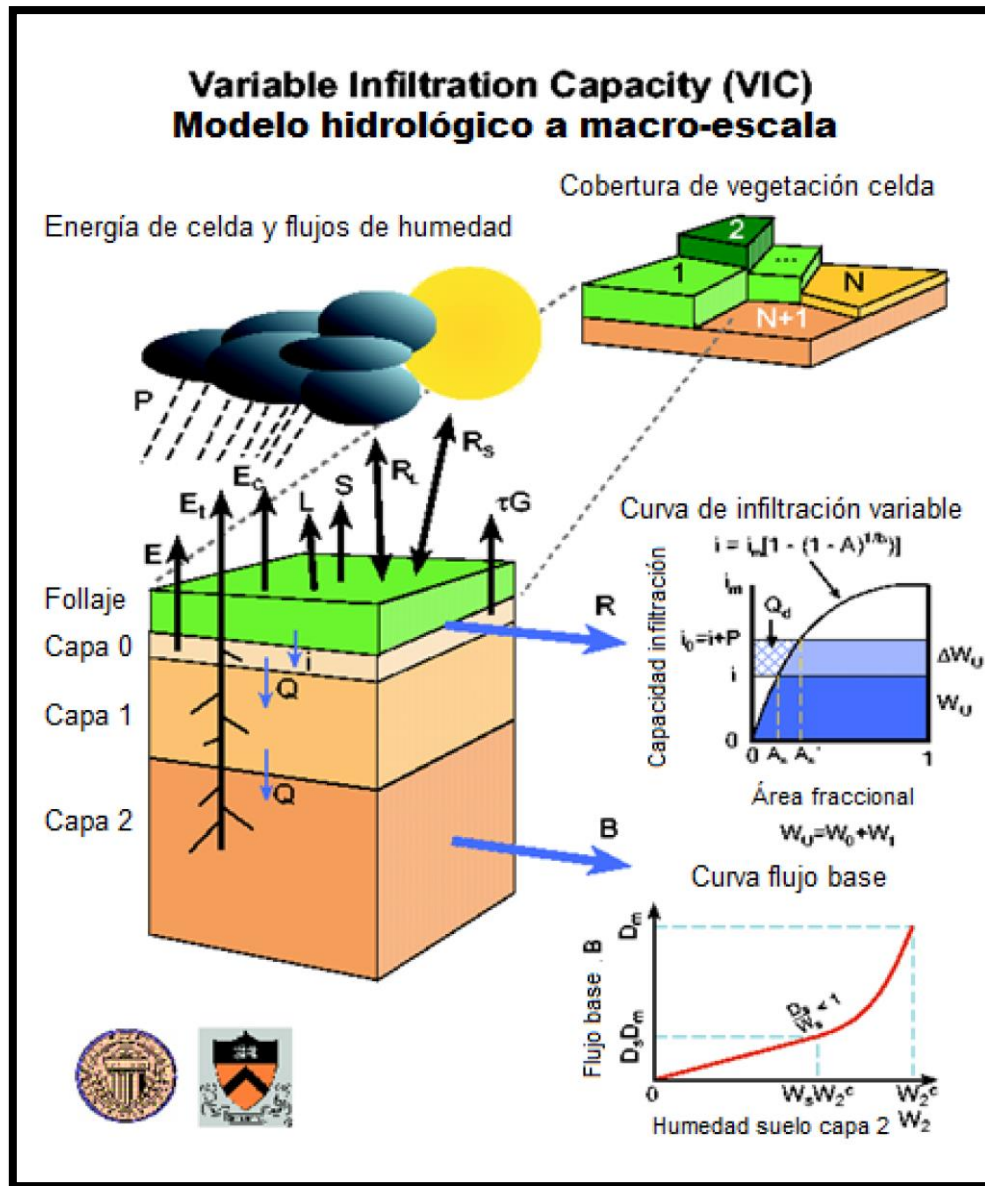


Figura 2.13: Representación de celda modelo VIC (Gao et al, 2009)

Las celdas no interactúan entre sí en la fase de balance energético. Esto requiere una serie de supuestos, siendo el más importante que los flujos verticales son mucho más grandes que los horizontales. Así también se considera que el flujo subterráneo es pequeño en relación con el de superficie y tanto el subsuperficial como el proveniente de lagos y humedales no tienen aportes significativos a los canales, y las inundaciones sobre las planicies no son importantes.

Este tipo de supuestos son cumplidos satisfactoriamente toda vez que el tamaño de la celda es lo suficientemente grande. Este tamaño es variable según la cuenca a estudiar y se recomienda que se encuentre en el rango de 1 km a 2° de circunferencia (~200 km).

Los parámetros que se definen en cada celda son variables dependiendo de la zona de estudio. Sin embargo, existen parámetros fijos que siempre deben mostrarse, tales como la vegetación y los parámetros de suelo. Además de esto existen variables meteorológicas que deben presentarse también, como la precipitación y las temperaturas mínimas y máximas en cada celda, para un paso de tiempo previamente definido, a esto se le puede agregar variables como la velocidad del viento, humedad, albedo, radiación entre otras.

La representación de la vegetación es mediante el uso de una librería de parámetros, el cual se basa en el índice de área de follaje, o Leaf Area Index (LAI), esto entrega los valores de follaje relevantes.

La evapotranspiración en VIC se define bajo tres componentes para cada tipo de vegetación, estos son la evaporación de la cubierta húmeda, la transpiración de la cubierta seca y la evaporación del suelo de cubierta desnuda. La evaporación de la vegetación húmeda y la transpiración de la vegetación seca son estimadas la aproximación física de Penman-Monteith (Figura 2.14).

$$E_p = \frac{s(R_n - G) + \rho c_p d_a / r_a}{s + \gamma(1 + r_s / r_a)}$$

Figura 2.14: Aproximación de Penman Monteith (FAO, 2006)

En cuanto a los suelos, su parametrización se basa en las texturas de suelo, información que se extrae desde La Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO).

La composición del suelo es utilizada para estimar porosidad, saturación, capacidad de campo, marchitamiento, capacidad residual y otros parámetros para casos no saturados. Los tipos de suelos existentes según su textura se muestran en la Figura 2.15, dependiendo estos de su porcentaje de arcilla, limo y arena.

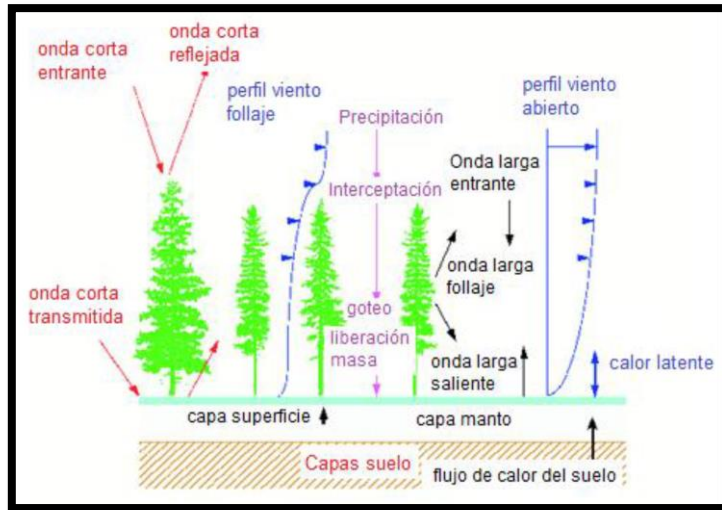


Figura 2.17: Simulación de Nieve en VIC

2.6.- Análisis de Componentes Principales

El análisis de componentes principales es un método mediante el cual se busca reducir la dimensionalidad de un conjunto n de variables correlacionadas entre sí. El resultado es una nueva matriz de variables m ($m < n$) linealmente independientes (componentes principales), las cuales son combinaciones lineales de las variables originales.

La primera componente principal explica la mayor varianza de los datos, es decir, minimiza la suma de los cuadrados de las distancias de los valores observados. De esta manera, a medida que se obtienen nuevas componentes la varianza disminuye hasta poder explicarla completamente (100% de la varianza).

El procedimiento de cálculo es el siguiente:

1. Se centran y normalizan las variables originales (x_i):

$$w_i = \frac{x_i - \mu_i}{\sigma_i} \quad (2)$$

Donde:

x_i : Variable original.

μ_i : Media de la variable original.

σ_i : Desviación estándar de la variable original.

Así, se obtiene una matriz con N_θ filas (observaciones) y N_V columnas por cada una de las variables estandarizadas.

2. Se calcula la matriz de correlaciones:

$$R = (r_{ij}) = \begin{bmatrix} 1 & \cdots & r_{1,NV} \\ \vdots & \ddots & \vdots \\ r_{N\phi,1} & \cdots & 1 \end{bmatrix} \quad (3)$$

Donde r_{ij} es la correlación entre las variables w_i y w_j .

3. Se obtienen un nuevo sistema de coordenadas rotando el sistema de ejes w_j de modo que cada componente CP_i sean linealmente independientes entre sí. Lo anterior, se logra calculando los valores propios de la matriz R:

$$\text{Det}[R - \lambda I] = 0 \quad (4)$$

Donde:

R: Matriz de correlación

I: Identidad

λ : Valores propios de la matriz R.

Luego el porcentaje de variación explicada P_i por cada componente se calcula como sigue:

$$P_i = (\lambda_i)/(NV) \quad (5)$$

Cada λ_i permite obtener un conjunto de NV coeficientes c_{ij} (eigen values) definidos por:

$$(R - \lambda_i I)C_i = 0 \quad (6)$$

$$\sum_j C_{i,j}^2 = 1 \quad (7)$$

Donde

i: Se refiere a cada CP_i

j: número de elementos de w

C_i : Es el vector propio unitario asociado a λ_i

4. Obtención de las componentes principales:

$$CP_i = \sum_{j=1}^{NV} c_{i,j} w_j \quad (8)$$

5. Ya determinadas las componentes principales se procede a elegir n de ellas, las cuales en conjunto expliquen determinada cantidad de la varianza total. Generalmente, esto se realiza

observando un gráfico de sedimentación (scree plot). En la Figura 2.18 se puede observar un ejemplo:

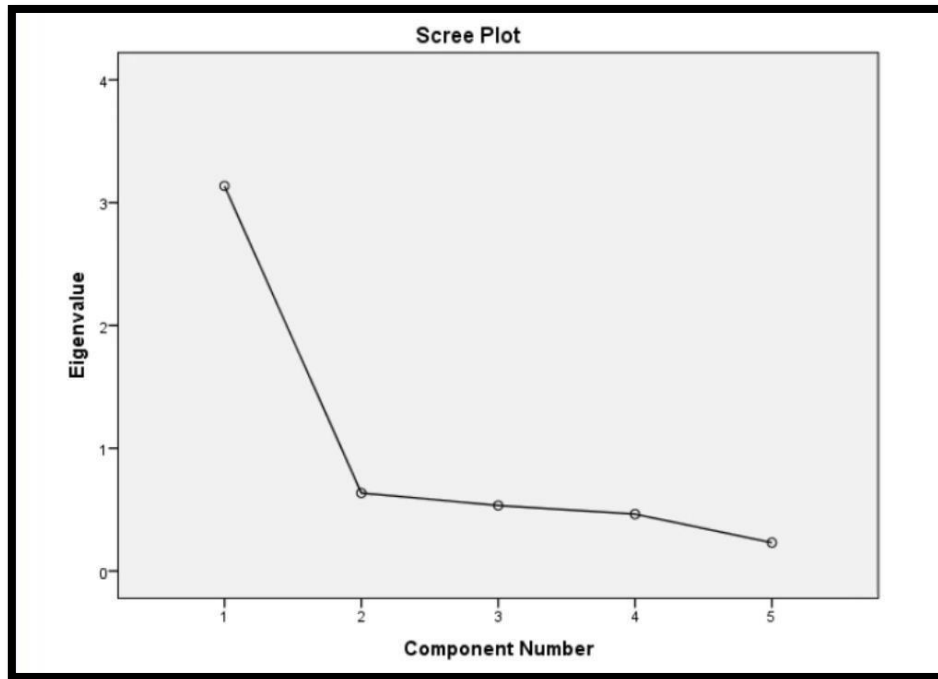


Figura 2.18: Ejemplo de grafico de sedimentación.

2.7.- Métricas de Evaluación

En este trabajo se utilizaron 4 índices para medir el desempeño de los modelos de pronóstico.

2.7.1.- Error porcentual

El error porcentual se define de la siguiente manera:

$$E = \frac{Q_o - Q_s}{Q_o} \quad (9)$$

Donde:

Q_s : Es el caudal simulado.

Q_o : Es el caudal observado

μ_o : Es la media de la serie de caudales observados.

Mientras el error sea lo más cercano 0 mejor es el desempeño del modelo.

2.7.2.- Normalized Mean Squared Error (NMSE)

El indicador NMSE es una variación del error cuadrático medio MSE. Se calcula de la siguiente manera:

$$NMSE = \frac{1}{N} \sum_t \frac{(Q_s(t) - Q_o(t))^2}{\mu_o * \mu_s} \quad (10)$$

Donde:

Q_s : Es el caudal simulado.

Q_o : Es el caudal observado

μ_o : Es la media de la serie de caudales observados.

2.7.3.- Nash-Sutcliffe Model Efficiency (NSE)

El coeficiente Nash-Sutcliffe (Nash and Sutcliffe, 1970) conocido comúnmente como NSE es un indicador del poder predictivo de un modelo hidrológico. El NSE se determina de la siguiente forma:

$$NSE = 1 - \frac{\sum_{t=1}^n (Q_s(t) - Q_o(t))^2}{\sum_{t=1}^n (Q_o(t) - \mu_o)^2} \quad (11)$$

Donde:

Q_s : Es el caudal simulado.

Q_o : Es el caudal observado

μ_o : Es la media de la serie de caudales observados.

Mientras el NSE sea cercano a 1 mejor es el desempeño del modelo.

2.7.4.- Kling-Gupta Efficiency

El indicador Kling-Gupta (Kling and Gupta, 2009) conocido comúnmente como KGE es un indicador de desempeño de modelos hidrológicos. Este indicador es producto del análisis de la descomposición de otros índices como NSE y MSE. Trata de buscar el punto óptimo de la frontera de Pareto compuesta por α (variabilidad relativa entre los valores observados y simulados), β (comparación entre la media de la serie observada y simulada) y r (correlación lineal entre la serie observada y simulada). El cálculo del KGE se observa en la Eq(12):

$$KGE = 1 - \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\alpha - 1)^2} \quad (12)$$
$$\alpha = \frac{\sigma_s}{\sigma_o}$$
$$\beta = \frac{\mu_s}{\mu_o}$$

Donde

σ_o : Desviación estándar de la serie observada.

σ_s : Desviación estándar de la serie simulada.

μ_s : Media de la serie simulada.

μ_o : Media de la serie observada.

2.7.5.- Criterio de información de Akaike

El criterio de información de Akaike (Akaike, 1974) (AIC), es una medida de la calidad relativa de un modelo estadístico, para un conjunto de datos. El AIC proporciona un criterio de selección de modelos.

El AIC maneja un trade-off entre la bondad de ajuste del modelo y su complejidad. Ofrece una estimación relativa de la información perdida cuando se utiliza un modelo determinado para representar el proceso que genera los datos. Este indicador no proporciona información acerca de la calidad de un modelo.

El AIC se calcula de la siguiente manera:

$$AIC = 2k - 2\ln(L) \quad (13)$$

Donde:

k: Número de parámetros del modelo estadístico,

L: Máximo valor de la función de verosimilitud para el modelo estimado.

Ahora, cuando el tamaño de la muestra es finito se utiliza la siguiente relación:

$$AIC_c = AIC + \frac{2k^2 + 2k}{n - k - 1} \quad (14)$$

2.8.- Validación cruzada

La validación cruzada (Larson, 1931) es un método estadístico para evaluar y comparar algoritmos de aprendizaje mediante la división de los datos en 2 segmentos: la primera para entrenar y la segunda para validar. En este trabajo se ocupó la validación cruzada “Leave one out” para la selección de la RNA óptima, el procedimiento es el siguiente (Stone, 1974):

1. Se separa 1 observación de las n que se tienen.
2. Se entrena con las n-1 observaciones restantes
3. Con los pesos de las RNA obtenidas se simula el valor con los predictores asociados.
4. Se contrasta el valor simulado con el valor observado y se registra el error.
5. Se repite n veces hasta haber simulado todos los datos.

3.-Metodología

La metodología utilizada en este trabajo de título se resume a continuación:

1. Se analiza el modelo VIC implementado a escala horaria (DGA, 2017) en la cuenca definida por la estación fluviométrica Maipo en el Manzano para obtener las series temporales diarias de las variables hidrometeorológicas disponibles y se determina cuáles se utilizan para definir aquellas a incorporar en la capa de entrada de las redes neuronales. Para dichos efectos se considera usar el método de componentes principales o juicio de expertos a base de la información simulada por el modelo hidrológico VIC en el período histórico.
2. Se considerarán 2 modelos referentes a la disponibilidad de información (M1 y M2), para ver el impacto de esta en los pronósticos de caudales.
3. Se considera analizar, al menos, tres conjuntos de variables en la capa de entrada. Para ello se separa el set de datos del período histórico 1985 a 2015 en tres conjuntos: entrenamiento (50%), validación (25%) y prueba (25%). Se analizan escalas de modelación semanales, mensuales y semestrales lo que permite incluir mayor número de variables de manera de asegurar que los ejemplos disponibles en cada caso verifiquen la restricción de ser al menos 10 veces el número de parámetros de la red.
4. Se entrena cada red con distintas combinaciones de variables del periodo 1985 a 2015 y número de neuronas en la (s) capa (s) oculta (s) de manera de seleccionar la arquitectura óptima revisando su desempeño en los ejemplos de validación.
5. Se aplica cada arquitectura de red seleccionada a los ejemplos de prueba y se analizarán los resultados obtenidos para definir el modelo a utilizar en la etapa siguiente. Se considera realizar 4 entrenamientos adicionales con el periodo histórico de los modelos MIROC-ESM, CCSM4, CSIRO e IPSL manteniendo la estructura de predictores, con el fin de determinar el vector de pesos y bias que entregan los mejores resultados en conjunto para los 4 modelos utilizados.
6. Se aplica la red óptima, obtenida en la etapa anterior, usando la información futura (2020-2050) derivada del uso del modelo VIC bajo los escenarios definidos en el proyecto de actualización del balance hídrico nacional (DGA, 2018) usando los modelos mencionados anteriormente.
7. Se contrastan los valores de los volúmenes totales de deshielo y caudales medios de los períodos de deshielo (Octubre Marzo) obtenidos entre los años 2020 a 2050 con los simulados por el modelo VIC en dichos años.
8. Según los resultados obtenidos se concluye respecto de la validez de utilizar un modelo estocástico implementado a base de información pasada en un escenario futuro no estacionario de cambio climático en la cuenca de estudio.

En la Figura 3.1 se puede observar el diagrama de flujo con las principales actividades:

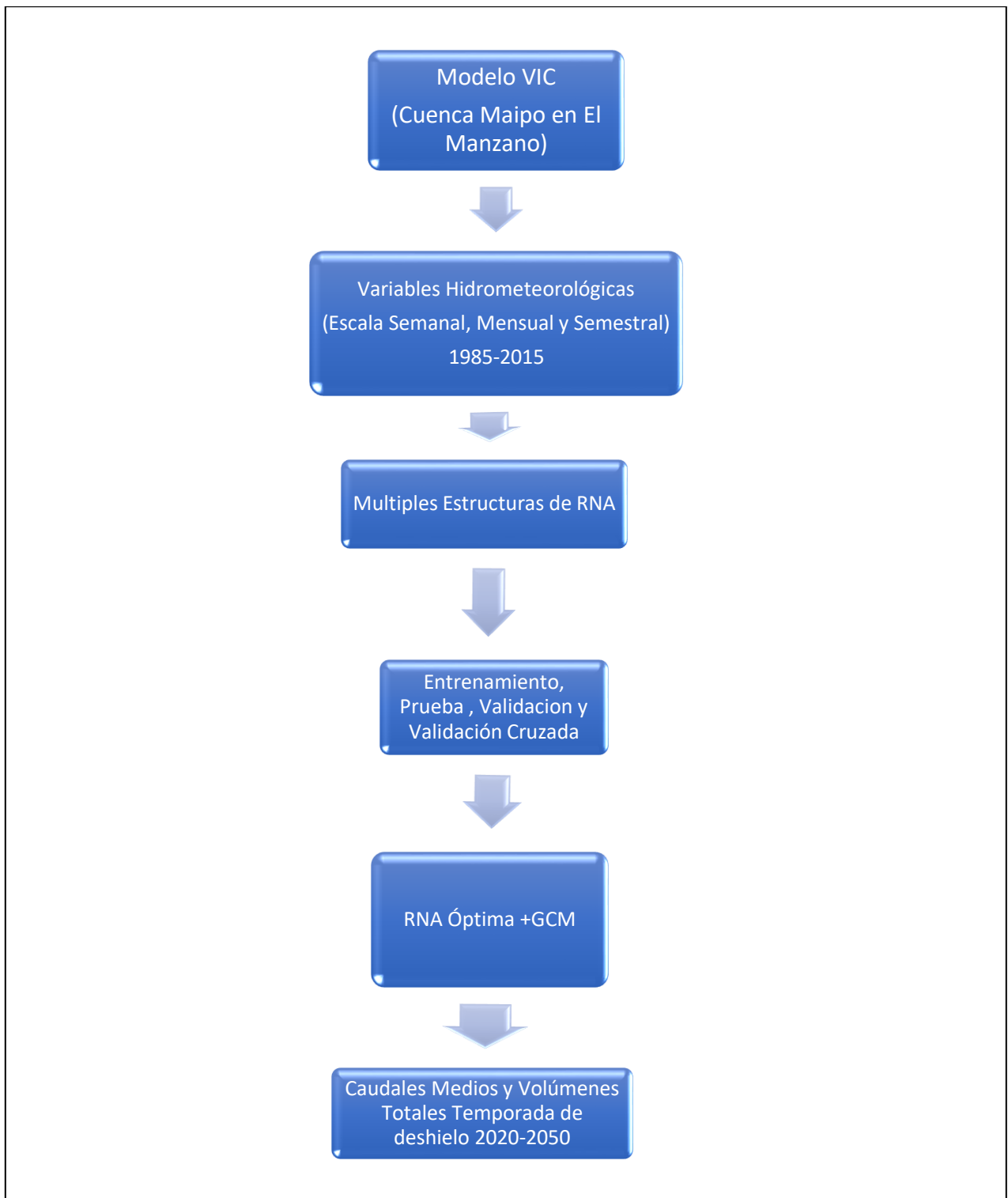


Figura 3.1: Diagrama de flujo de la Metodología

4.-Zona de Estudio

El río Maipo nace en las laderas del volcán Maipo a 5623 msnm con las características de una quebrada cordillerana. La gran altura que presenta la cordillera de los Andes en esta zona permite la presencia de zonas glaciales y nieve estacional que controla las características hidrológicas del río. Ya en los 1200 msnm el río toma un caudal importante debido a los aportes de los ríos Volcán y Yeso además del aporte del río Colorado el cual toma un valor promedio anual de 26 m³/s. Finalmente, al llegar a la estación Maipo en el Manzano el río tiene caudales peak promedio de 200 m³/s en la temporada de deshielos y un caudal medio anual del orden de 100 m³/s registrados según la DGA (1970-2016). Los caudales medios anuales desde 1970 a la fecha se observan en la Figura 4.2.

En la parte alta de la cuenca del Maipo (Figura 4.1) se registra una temperatura media anual de 10° con grandes contrastes térmicos pudiendo registrar temperaturas máximas diarias de 27°. Según autores (Videla, 2013), se considera la existencia de un aporte de agua de fusión glacial y que la línea de nieves es variable durante el invierno, rondando los 2000 msnm. Se establece que la Línea de Nieve Estacional (LNE) de invierno suele ubicarse unos 100 metros más baja de la altura media de la isoterma 0 (°C), es decir a unos 1650 msnm.

Según la DGA (2000) en la cuenca esta particionada en 8 subcuencas (Tabla 4.1):

Tabla 4.1: Delimitación de la cuenca alta del río Maipo (DGA)

COD Subcuenca	NOMBRE	AREA [km2]
05700	Río Maipo bajo junta Río Negro	768
05701	Río Maipo entre Río Negro y Río Volcán	737
05702	Río Volcán	523
05703	Río Yeso	627
05704	Río Maipo entre Río Volcán y Río Colorado	533
05705	Río Colorado antes junta Río Olivares	789
05706	Río Olivares	543
05707	Río Colorado entre Río Olivares y Río Maipo	338
	Total	4858

La cuenca muestra una morfología glacial y peri-glacial redondeada en las partes medias y bajas por los procesos fluviales. Se caracteriza por grandes valles encerrados con pendiente promedio de 30°.

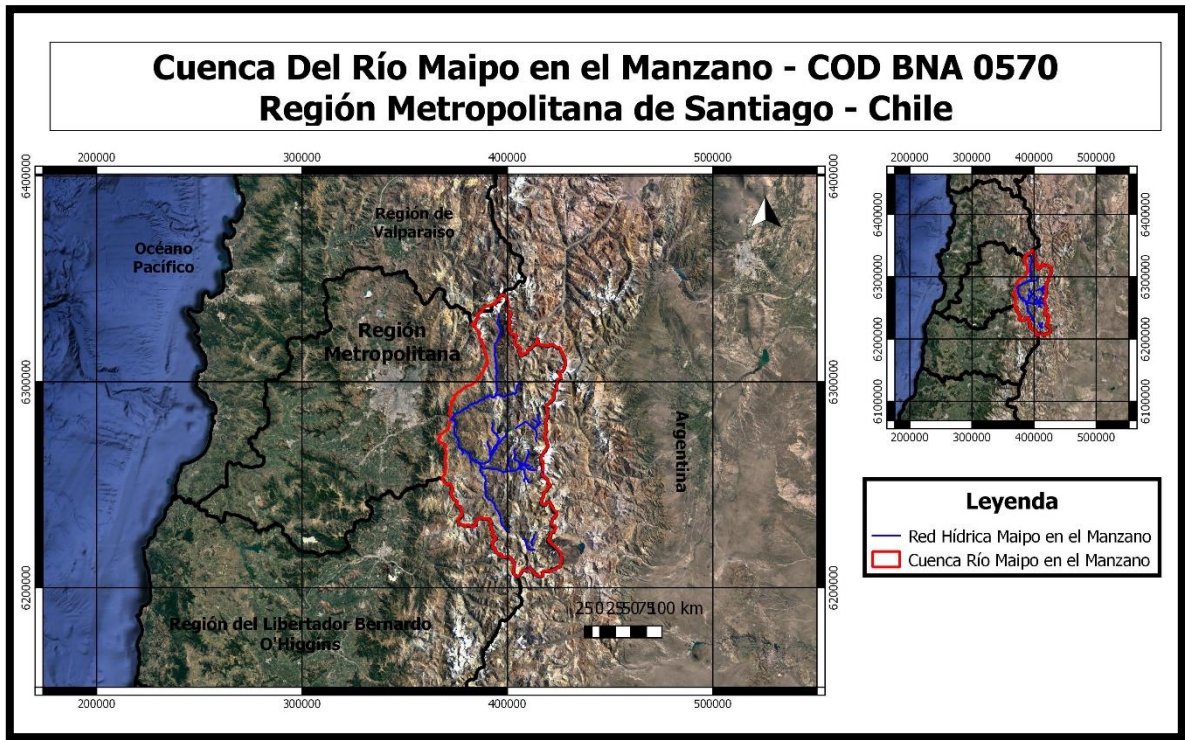


Figura 4.1: Cuenca Alta del Río Maipo (Fuente: Elaboración propia)

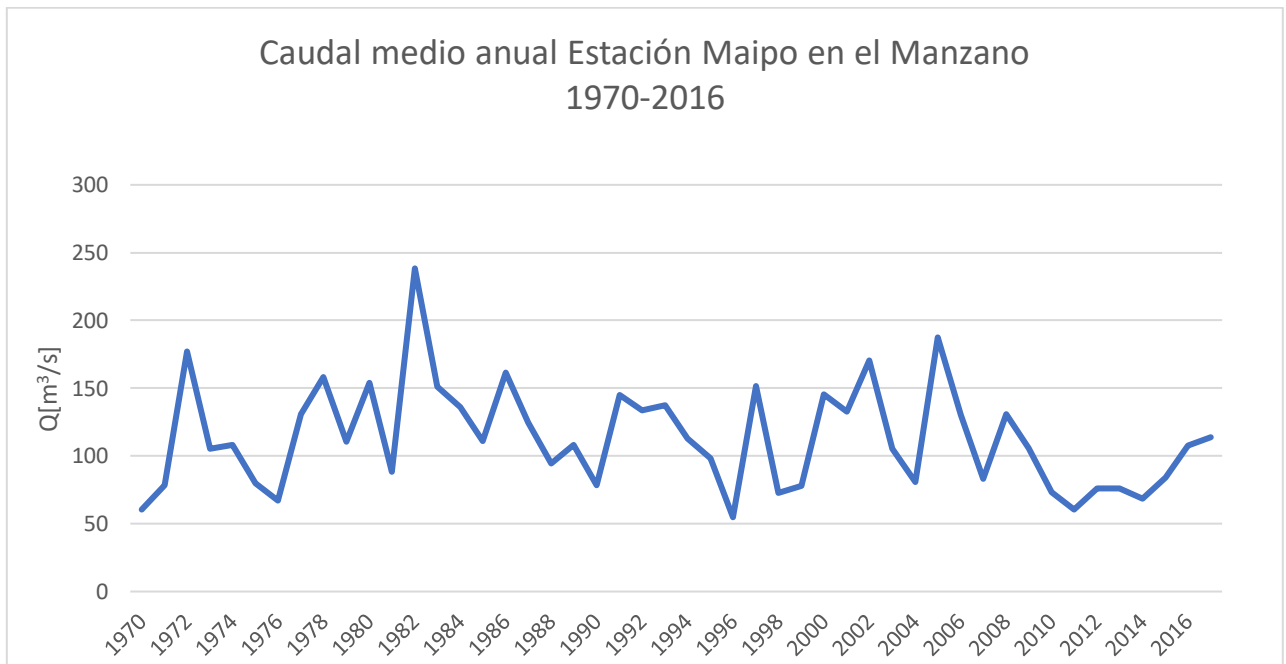


Figura 4.2: Caudales Medios Anuales Cuenca Maipo en el Manzano Periodo 1970 – 2016 (DGA)

El perfil altitudinal varía entre los 1000 y 6500 msnm como se puede se puede observar en la Figura 4.3:

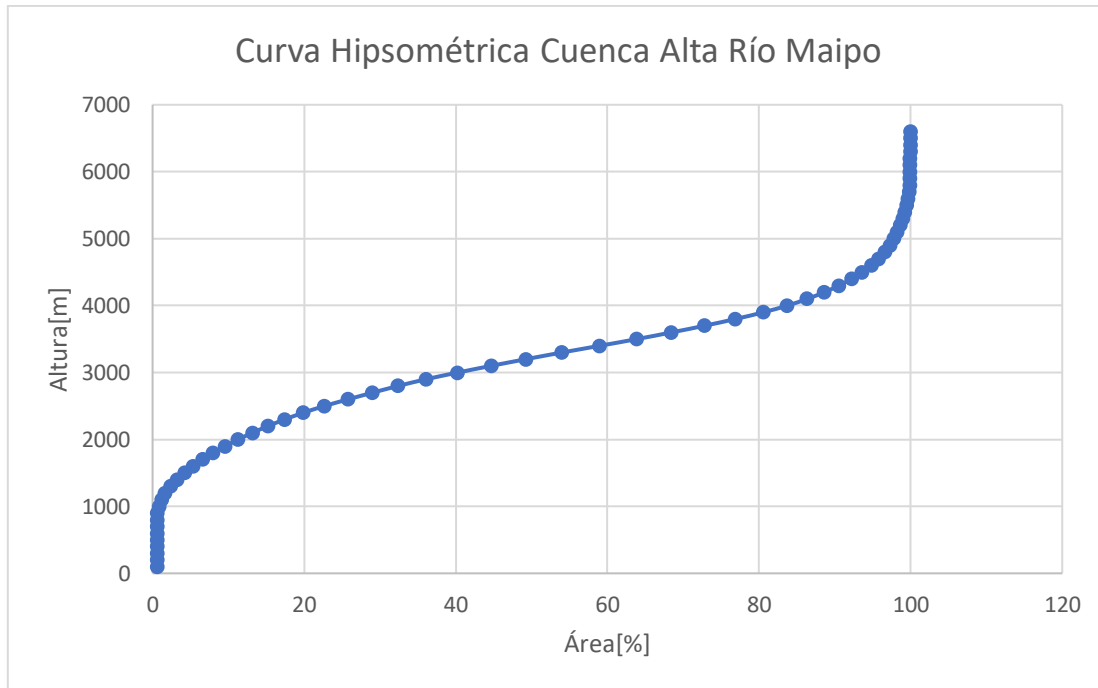


Figura 4.3: Curva Hipsométrica Cuenca Alta Del Maipo (ASTER GDEM NASA: 20170814125817_550532624)

Según el registro de las estaciones Fluviométricas (Tabla 4.2) actualmente operando en la cuenca alta del río Maipo se pueden distinguir dos regímenes hidrológicos: Nival y Nivo-Pluvial. El primero se desarrolla en la zona de influencia de la cordillera de Los Andes, donde los mayores caudales se presentan en los meses de deshielo (oct.-Mar). Además, la estacionalidad está fuertemente marcada como se observa en la Figura 4.4 donde se muestran los coeficientes de Pardé:

Tabla 4.2: Estaciones Fluviométricas y régimen hidrológico

N°	Régimen	Estación	Latitud [°]	Longitud [°]	Altitud [msnm]	Vigencia
1	Nival	Río Maipo En Las Hualtatas	33.98	70.15	1820	1979-2013
2		Río Olivares Antes Junta Río Colorado	33.49	70.14	1532	1991-
3		Río Maipo En Las Melosas	33.85	70.2	1527	1966-
4		Río Colorado Antes Junta Río Olivares	33.59	70.13	1500	1977-
5		Río Volcán En Queltehues	33.81	70.21	1365	1966-
6		Río Colorado Antes Junta Río Maipo	33.49	70.37	920	1970-
7	Nivo-Pluvial	Río Maipo En San Alfonso	33.73	70.3	1092	1985-
8		Río Maipo En El Manzano	33.76	70.16	864	1965-

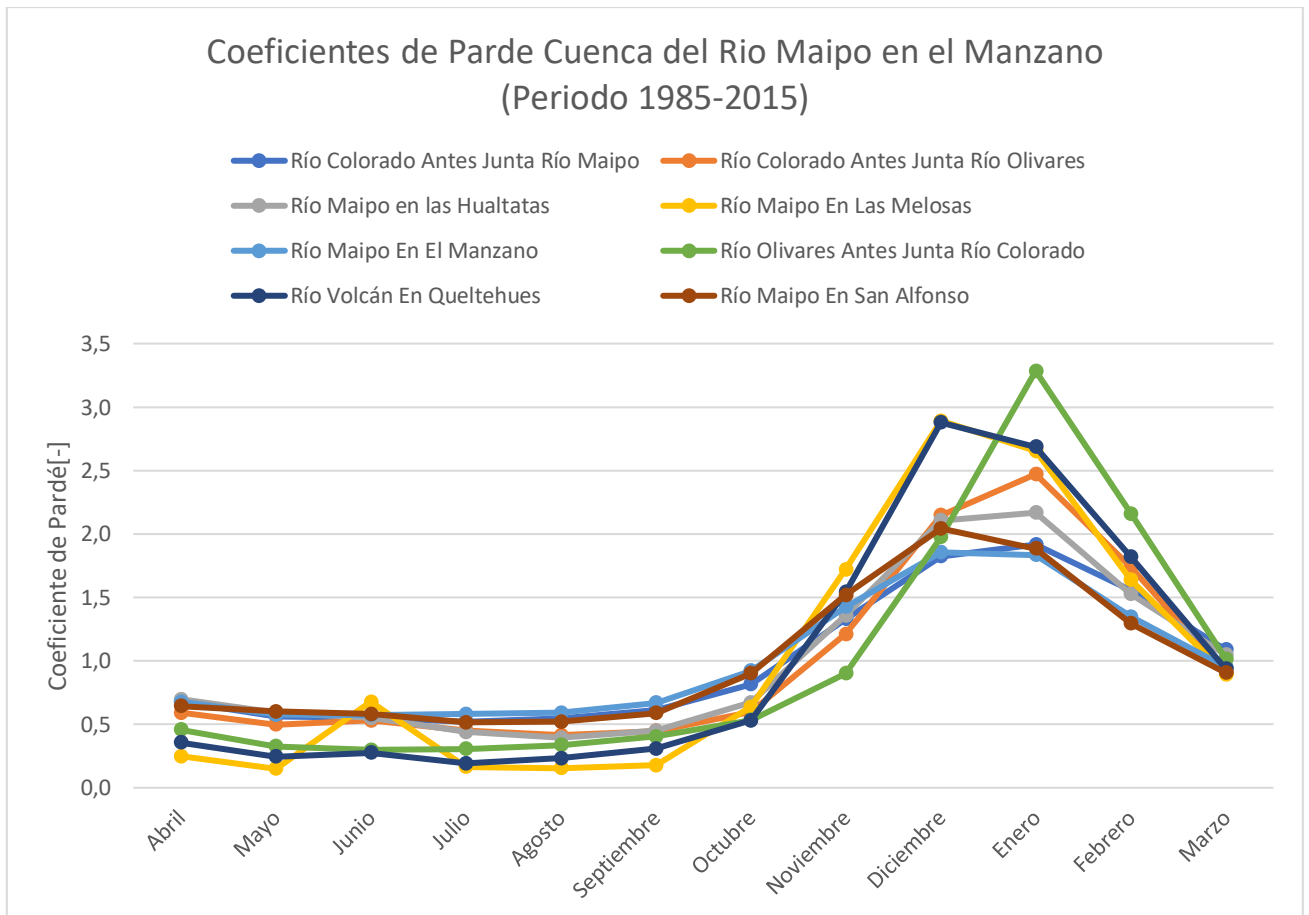


Figura 4.4: Coeficientes de Parde de las subcuencas del Rio Maipo en el Manzano 1980-2015

El caudal efluente de la zona en estudio se mide en la estación Maipo en El Manzano, generando la curva de variación estacional (CVE) que se observa en la Figura 4.5.

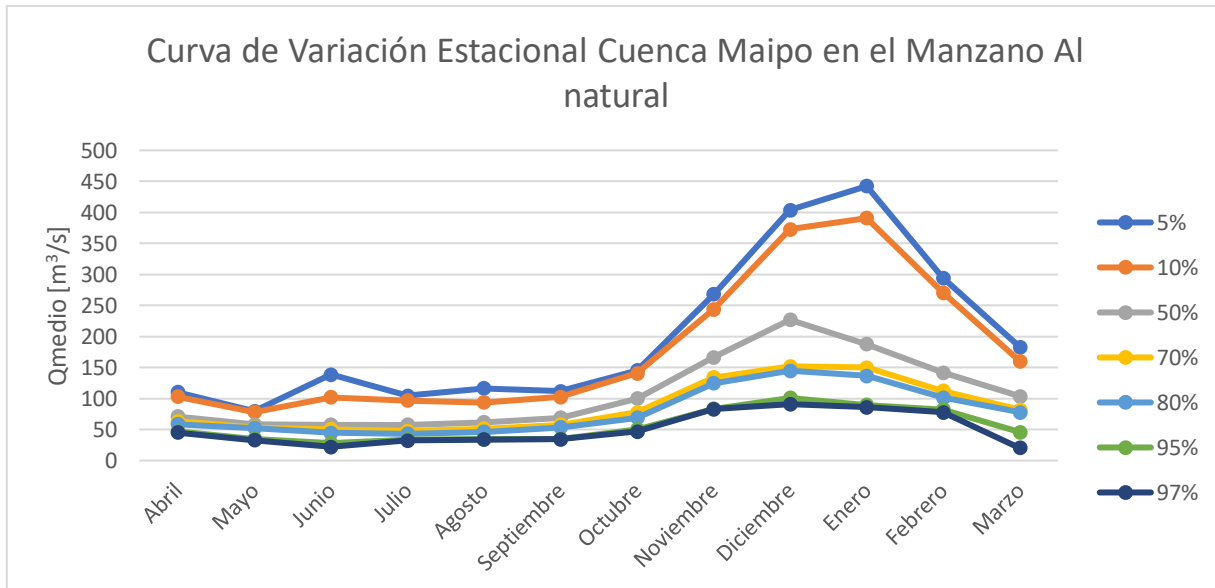


Figura 4.5: Curvas de variación estacional cuenca Maipo en el Manzano período 1980 a 2016

En cuanto a derechos de aguas el río Maipo se divide en tres secciones (Figura 4.6), donde la primera sección está ubicada en la cuenca alta (zona de estudio). En esta sección se administran cerca de 8000 acciones (160.000 l/s) correspondientes a derechos consuntivos permanentes y aproximadamente 12000 (180.000 l/s) no permanentes.

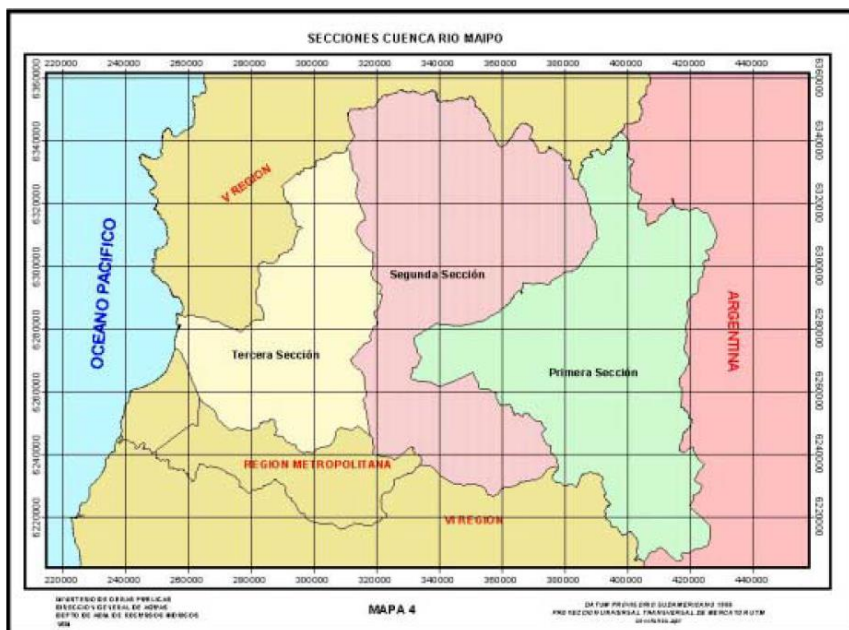


Figura 4.6: Secciones de Río Maipo

En esta zona se encuentran las centrales hidroeléctricas que se detallan en la Tabla 4.3 (DGA, 2004).

Tabla 4.3: Centrales Hidroeléctricas en la Cuenca alta del Río Maipo (DGA, 2004)

Nombre Central Hidroeléctrica	Propietario	Puesta en Servicio	Utilización de Recursos	Caudal de Diseño [m ³ /s]
Queltehues	CHILEGENER S.A.	1928	Río Maipo	28,1
El Volcán	CHILEGENER S.A.	1942	Río Volcán	9
Alfalfal	CHILEGENER S.A.	1991	Río Colorado y Olivares	30
Maitenes	CHILEGENER S.A.	1923	Río Colorado	11
Puntilla	Cía. Manufacturera de Papeles Y Cartones	S/I	Río Maipo	18
La Florida	Sociedad Canal De Maipú	1909	Río Maipo	22
Los Bajos y Caemsa	Carbomet Energía S.A	S/I	Río Maipo	14,3
Los Morros	Cía. Eléctrica Los Morros	S/I	Río Maipo	17

Las principales fuentes de captación de agua potable se ubican en:

- Laguna Negra
- Laguna Lo Encañado
- Estero El Manzano
- Estero El Canelo
- Río Maipo
- Embalse El Yeso (Almacenamiento disponible 250 Mm³)

5.-Resultados

5.1.- Análisis de las variables del modelo VIC.

Del modelo VIC (DGA, 2017) e información meteorológica se obtuvieron 5 variables a nivel diario para el período 1980-2015:

- Escorrentía o Caudal (Q) [m^3/s] (Modelado VIC)
- Precipitación (Pp) [mm]
- Temperatura media (T) [mm]
- Evapotranspiración [mm] (Modelado VIC)
- Equivalente en agua de la nieve (SWE) [mm] (Modelado VIC)

Posteriormente se realizó la siguiente agregación de las variables para cada paso temporal de pronóstico (semanal, mensual, período estacional):

- Promedio de la Escorrentía o Caudal [m^3/s]
- Suma de la Precipitación [mm]
- Promedio de la Temperatura [$^{\circ}\text{C}$]
- Suma de la Evapotranspiración [mm].
- Suma del Equivalente de Agua en Nieve [mm].

En esta sección solo se presentan los resultados a nivel mensual. Los demás resultados, incluyendo correlaciones y análisis de las componentes principales se presentan en el Anexo A.

En las figuras 5.1 a 5.5 se presentan las series históricas de las variables disponibles a nivel temporal mensual. Preliminarmente, se puede observar una tendencia a la disminución tanto para los caudales (Figura 5.6), precipitación y equivalente de agua en nieve (SWE).

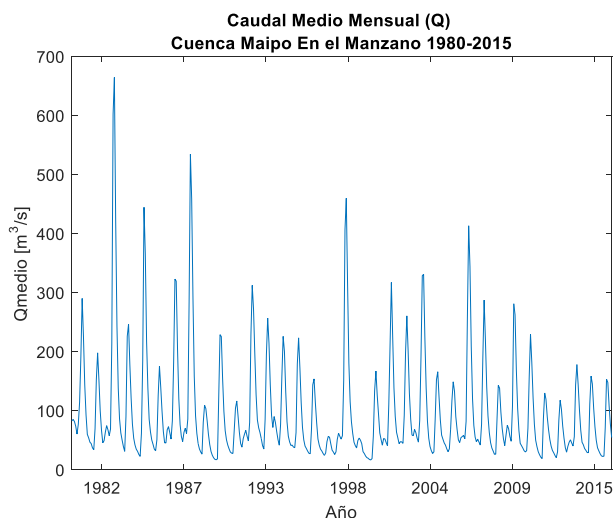


Figura 5.1: Caudal medio mensual

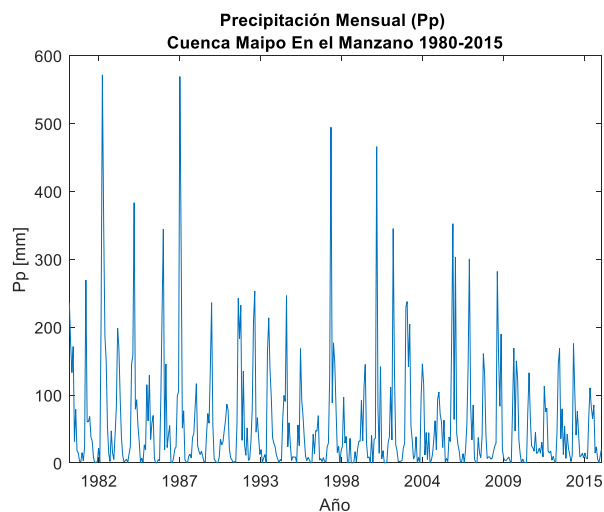


Figura 5.2: Precipitación Mensual

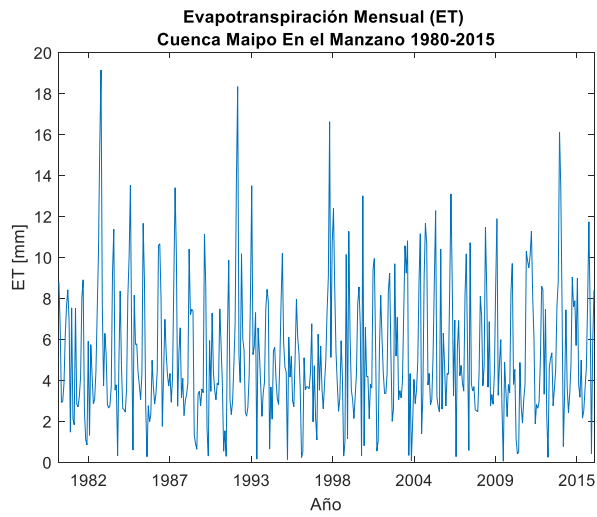


Figura 5.3: Evaporación Mensual

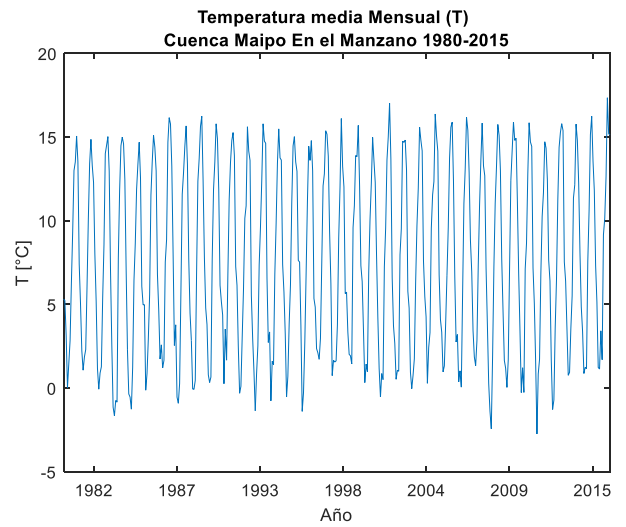


Figura 5.4: Temperatura Media Mensual

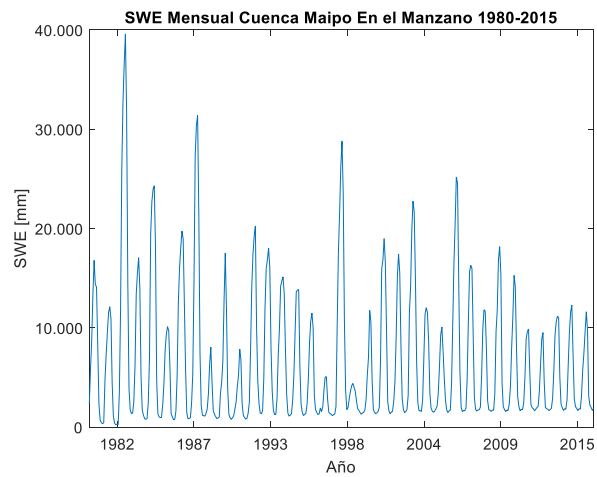


Figura 5.5: SWE medio Mensual

Esta tendencia a la baja en los caudales tiene sentido debido al comportamiento de la cuenca. Al ser una cuenca nival, donde la principal fuente de caudales es la cantidad de nieve acumulada hasta el comienzo de la temporada de deshielos, producto de las precipitaciones entre los meses de abril y septiembre (véase Figura 5.6), se esperarían disminuciones de caudales producto de la baja de precipitaciones en invierno.

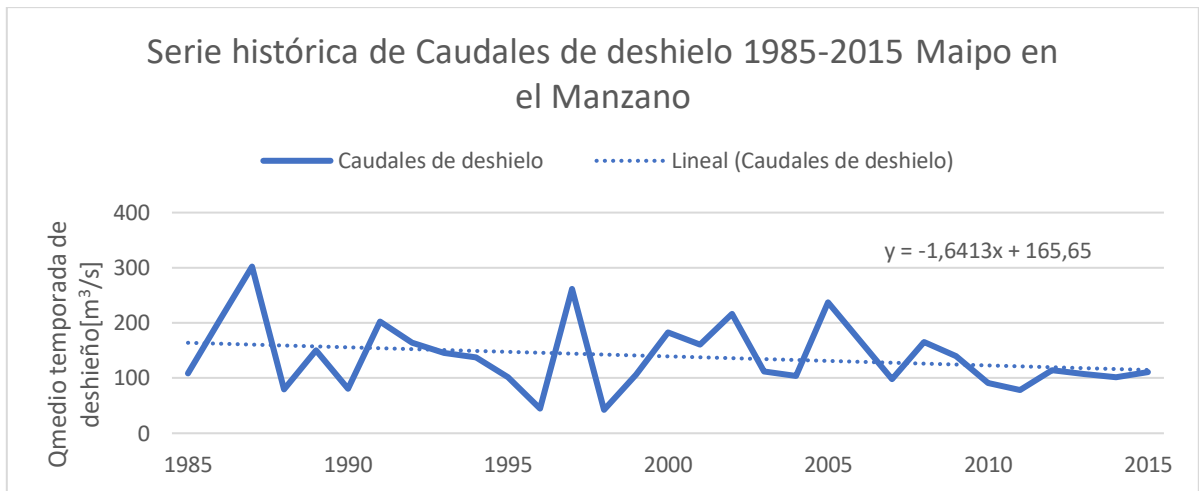


Figura 5.6: Serie histórica de caudales de deshielo 1985-2015.

Los estadísticos del período de deshielo de las series temporales obtenidas a escala mensual se muestran en la tabla 5.1. El análisis de las componentes principales (tablas 5.2, 9.2 y 9.4) muestra que para las series semanales y mensuales se explica aproximadamente un 97% de la varianza (figuras 5.7, 9.6 y 9.17) con solo las 4 primeras componentes. En dichas componentes, las variables con mayor importancia son Q, P, T y ET. En cambio, el análisis semestral muestra que las variables que explican mejor la varianza son SWE, T y ET.

Tabla 5.1: Estadísticos Variables Mensuales

	Estadísticos Variables Vic Escala Mensual				
	Q [m³/s]	SWE [mm]	Pp [mm]	T [°C]	ET [mm]
Promedio	150,4	4942,7	13,3	12,3	6,4
Std	104,8	6506,6	15,9	3,0	4,0
Max	664,2	39600,5	85,4	17,4	19,1
Min	27,2	236,9	0,0	3,8	0,1

Tabla 5.2: Eigen loads PCA mensual

	PC1	PC2	PC3	PC4	PC5
Q	0,46	0,39	0,10	0,69	-0,39
SWE	-0,41	0,52	-0,55	0,30	0,41
Pp	-0,46	0,16	0,80	0,24	0,24
T	0,60	-0,10	0,10	0,08	0,79
ET	0,22	0,74	0,17	-0,61	-0,04

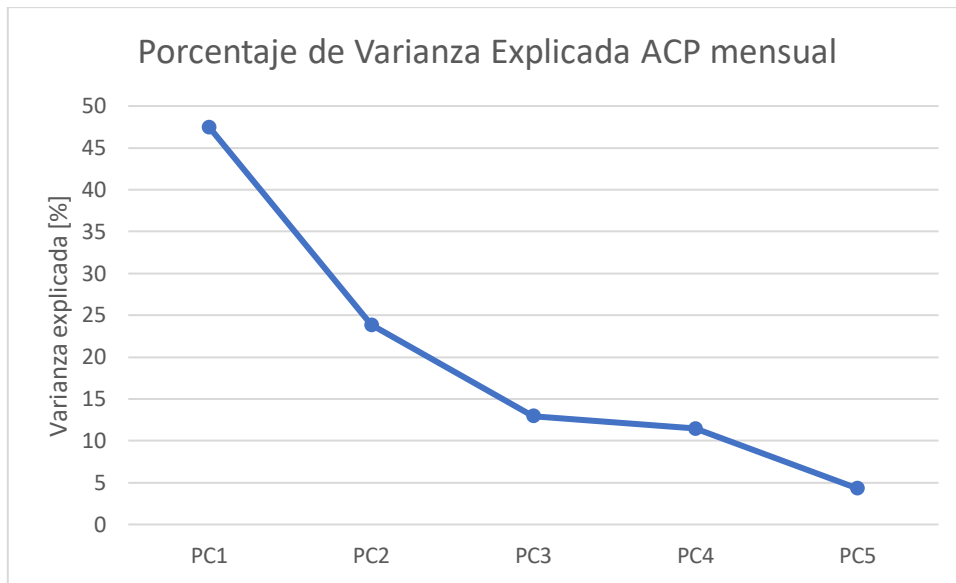


Figura 5.7: Porcentaje de Varianza explicada por cada componente principal nivel mensual

Luego, se analiza la autocorrelación parcial y correlación cruzada con el fin de determinar la cantidad de lags a la RNA (Figuras 5.8 a 5.12). Estas figuras son importantes para definir la dependencia todas las variables sobre el caudal. Sin embargo, no entregan una idea definitiva de los lags que podrían ser los mejores predictores. La Figura 5.8 indica altas correlaciones del caudal con dos observaciones anteriores ($Q(t-1)$ y $Q(t-2)$). Por otra parte, las demás variables indican una fuerte dependencia Inter-estacional, es decir, la correlación aumenta a medida que el tiempo retardo aumenta, hasta llegar a un peak que se encuentra en $t-3$ y $t-4$ específicamente, indicando un patrón repetitivo.

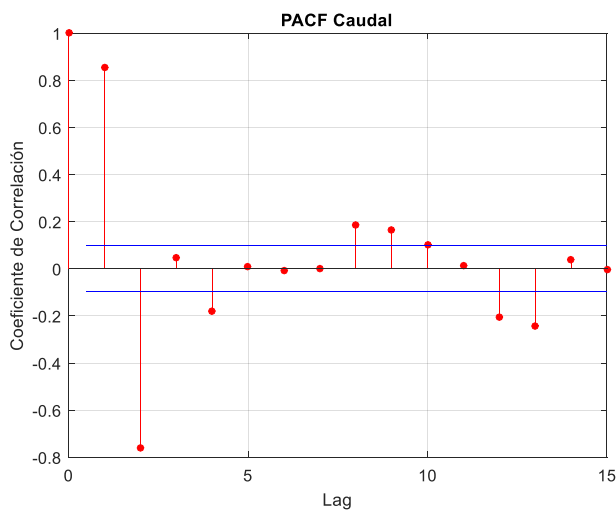


Figura 5.8: PACF Q medio mensual

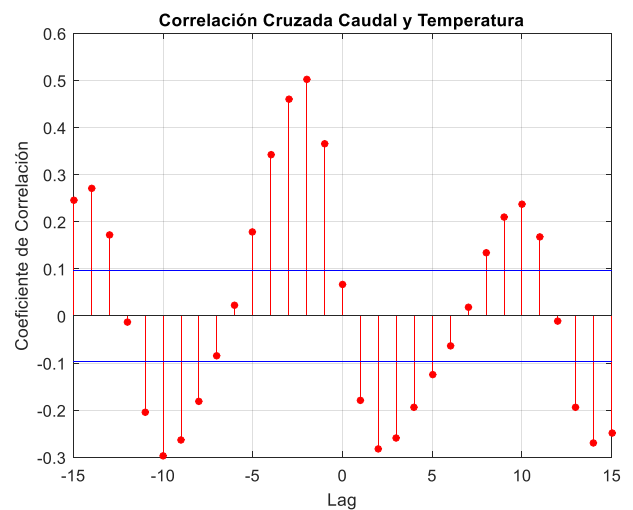


Figura 5.9: Correlación cruzada $Q-T$

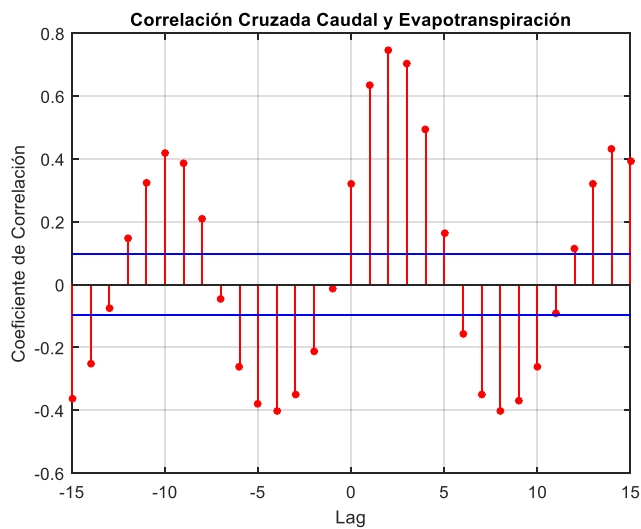


Figura 5.10: Correlación Q-E

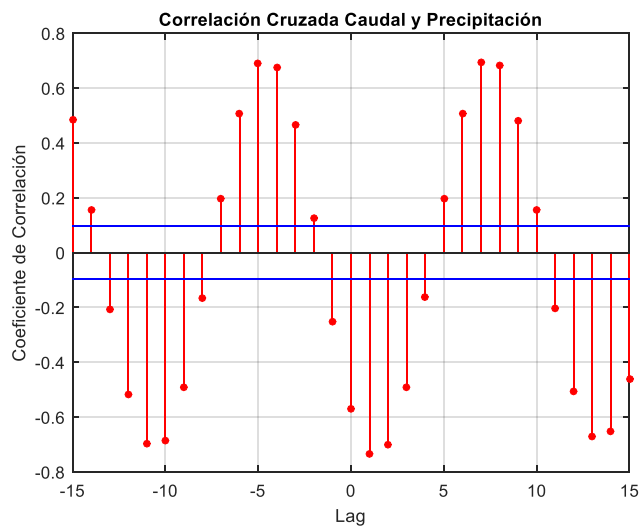


Figura 5.11: Correlación Q-P

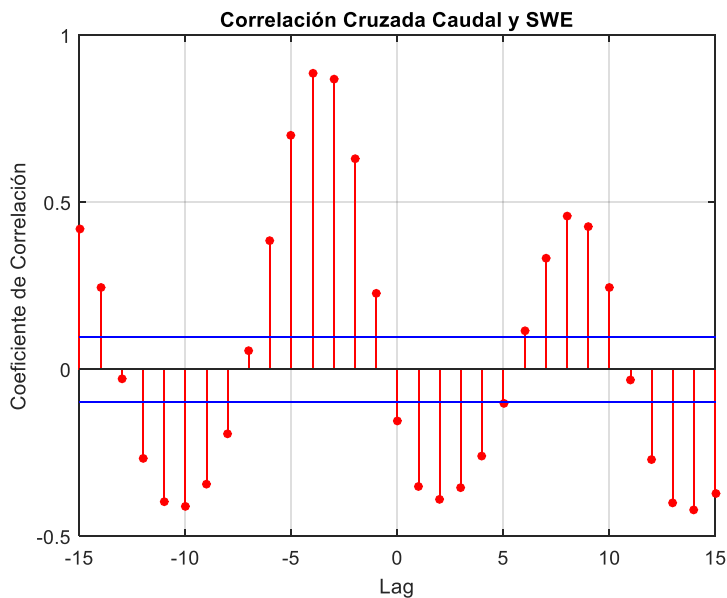


Figura 5.12: Correlación Q-SWE

Con relación a los otros pasos temporales (semanal y semestral), el comportamiento del sistema es similar, es decir, la correlación aumenta a medida que el retardo aumenta hasta llegar a un punto. No obstante, temporalmente la dinámica es distinta (2 semanas de retardo para el caso semanal en contraste con 2 meses para el caso mensual si se analizan los caudales).

Finalmente, en el caso semestral, el análisis obliga a utilizar como predictor la temporada anterior (invierno) por lo que el análisis resulta más restrictivo.

Las funciones de correlación cruzada entregan una idea parcial de los lags que tienen influencia sobre el caudal. Sin embargo, no representan la influencia que se está buscando debido a que se calculan con toda la serie temporal y no representa el análisis de la dependencia Inter-estacional específico que se busca. Dicho lo anterior, se calcula la correlación para cada agregación temporal de las variables para la temporada a pronosticar.

Para entender mejor la relación estacional de las variables se muestra en mapa de colores parte de la matriz de correlaciones (Anexo G). Con color verde se muestran las correlaciones cercanas a 1 y con rojo las correlaciones cercanas a -1. Lo que se está buscando es un patrón diagonal idealmente de color verde que refleje las altas correlaciones de determinados tiempos de retardo (lags).

5.2.- Determinación de los Predictores

Con las matrices y gráficos de correlación analizadas, se seleccionan los mejores predictores en cuanto a correlación para cada variable:

- 5 lags a nivel semanal.
- 5 lags nivel mensual.
- 1 lag nivel semestral.

Para que el análisis fuese lo más robusto posible, se generaron conjuntos aleatorios de predictores, es decir, para cada variable (Q, SWE, P, T, E) se eligieron subconjuntos de todos los predictores (Lags) obtenidos en el subcapítulo anterior. Para tener un amplio espectro de posibilidades, se generaron conjuntos para cada uno de los pasos temporales (60 semanal, 300 mensual y 60 semestral). Cada uno de ellos pasó al proceso de calibración de la red donde se calculan los indicadores NSE, KGE y MSE. Al mismo tiempo, se realizó una validación cruzada (“Leave one out”) de cada conjunto y finalmente se calculó el índice AIC que mide la parsimonia de un modelo.

Todas las redes se entrenaron con el método Levenberg-Marquardt que es el método que en general entrega buenos resultados (Khan et al., 2013), además de considerar un número específico de neuronas en la capa de oculta (Tabla 5.3) y función de transferencia tangente-sigmoidal en la capa oculta debido a la experiencia con múltiples iteraciones de prueba anteriores.

El ranking de los mejores conjuntos de variables y predictores se encuentra en el Anexo B.

Tabla 5.3: Número de neuronas en cada paso temporal

Paso Temporal	Número de Neuronas	
	Capa oculta	Capa de Salida
Semanal	10	1
Mensual	1	1
Semestral	1	1

Los conjuntos de predictores que se presentan en el anexo B permiten filtrar aquellas arquitecturas que son parsimoniosas de aquellas que no lo son. Sin embargo, el criterio de parsimonia no implica que las redes tengan un excelente desempeño en el pronóstico. Dicho esto, se buscó aquella arquitectura dentro de esta preselección que entregara los mejores resultados en el pronóstico.

La red óptima se considera como aquella que entrega el mayor número de temporadas con error bajo el 20%, que es el criterio que ocupa la DGA para calificar un buen pronóstico.

5.2.1.- Datos de adicionales para el Entrenamiento

Paralelamente al set de datos de calibración original (HIST), se tienen otros cuatro adicionales provenientes de los GCM (Global Circulation Model), estos sets de datos de entrada son distintos de HIST (Figura 5.13) debido a que el escalamiento temporal tuvo como objetivo obtener curvas de variación estacional (Figura 5.14) y curvas de duración (Figura 5.15) similares para el periodo histórico. De esta manera, las configuraciones que se utilizaron se describen en la Tabla 5.4 :

Tabla 5.4: Set de datos para el pronóstico 2020-2050

Set	Parámetros RN con forzantes LB	Descripción
1	HIST	Serie de datos modelados por VIC y calibrado con la información meteorológica histórica disponible en la cuenca.
2	CCSM4	Serie de datos obtenidos a partir del escalamiento del GCM CCSM4.
3	CSIRO	Serie de datos obtenidos a partir del escalamiento del GCM CSIRO.
4	IPSL	Serie de datos obtenidos a partir del escalamiento del GCM IPSL.
5	MIROC	Serie de datos obtenidos a partir del escalamiento del GCM MIROC.

Nota: LB se refiere a Línea base 1985-2015

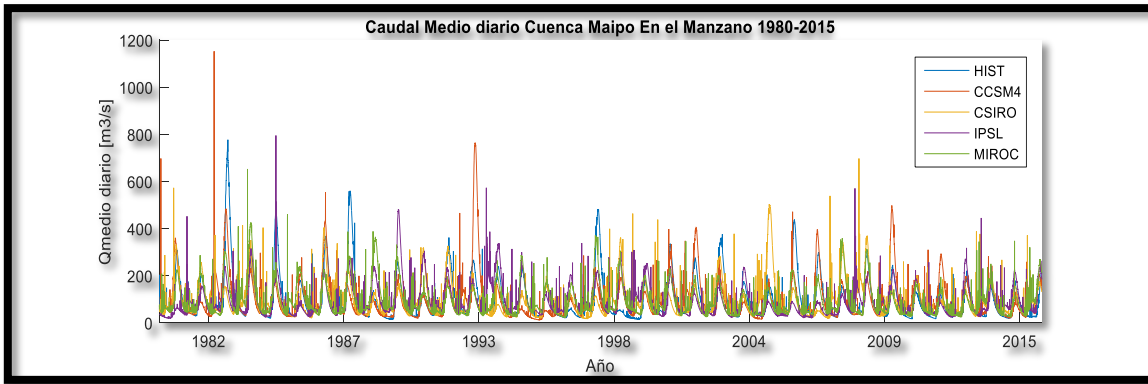


Figura 5.13: Caudales medios semanales en Maipo en El Manzano, simulados por forzantes históricas y por forzantes obtenidas a base de modelos GCM. Periodo histórico 1985-2015

Los estadísticos de cada set de datos se presentan en la Tabla 5.5:

Tabla 5.5: Estadísticos de las Series Históricas de Caudales medio diarios Cuenca Maipo en el Manzano periodo 1980-2015

	HIST	CCSM4	CSIRO	IPSL	MIROC
Promedio	97,4	92,4	90,9	89,2	92,3
std	93,6	87,0	71,1	66,1	68,1
max	777,3	1152,9	697,8	795,4	653,2
min	14,9	12,5	17,5	18,5	22,2

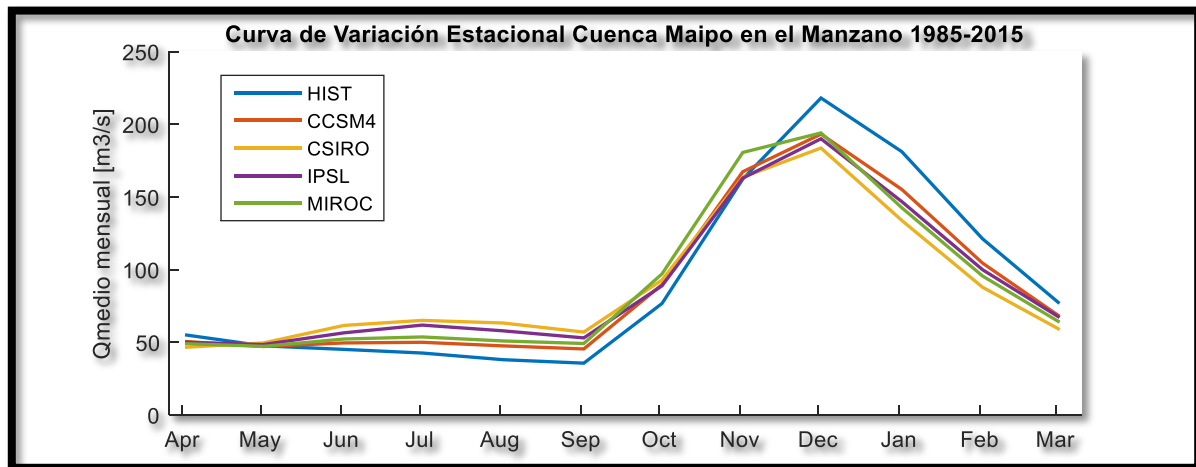


Figura 5.14: CVE Periodo 1985-2015 Cuenca Maipo en el Manzano

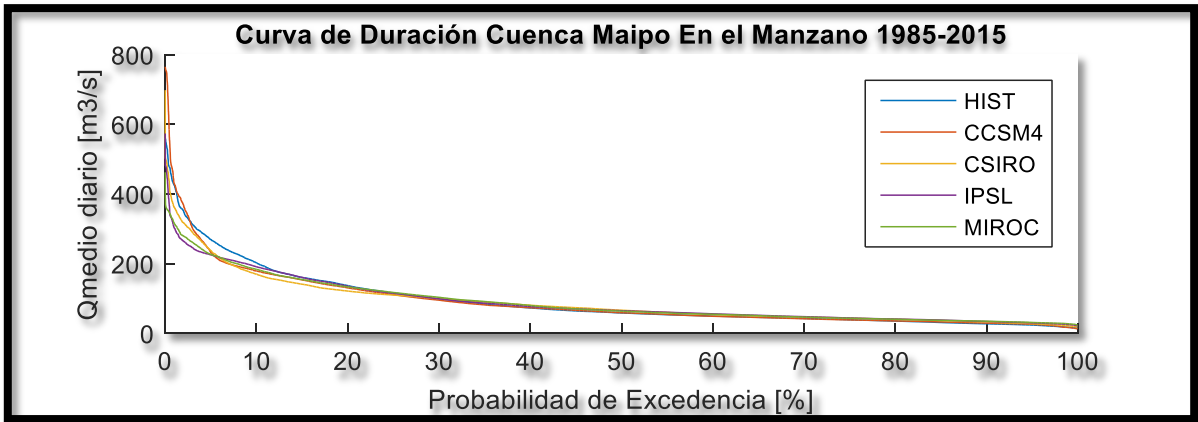


Figura 5.15: Curva de Duración periodo de deshielo 1985-2015 Cuenca Maipo en el Manzano

A modo de resumen, el procedimiento de trabajo con los sets de datos de calibración es el siguiente:

1. Calibración de las RNA con la serie HIST: Se obtendrán 60 conjuntos de predictores a nivel semanal, 300 a nivel mensual y 60 a nivel semestral.
2. Se elegirán para cada paso temporal una red óptima.
3. A partir de la estructura neuronal obtenida en el punto 2 (conjunto de predictores) se entrenará adicionalmente con cada GCM en su periodo 1985-2015. Se realiza este proceso para obtener 5 estructuras de pesos de la red neuronal.
4. Se toma la red óptima entre las 5 obtenidas en el punto 3 como aquella que entregue el máximo número de temporadas con error por bajo el 20% una vez hecho el pronóstico.

5.3.- Pronóstico del periodo de deshielo 2020-2050

5.3.1.- Bases del Pronóstico y definición de Modelos de pronóstico

En esta sección se proyectan los caudales de deshielo para el horizonte 2020-2050 a nivel semanal, mensual y semestral. Del mismo modo se muestran los resultados del pronóstico al utilizar los sets de datos de entrada mencionados en el subcapítulo 5.2.1.-.

Se definen los siguientes modelos de pronóstico en cuanto a la disponibilidad de la información:

1. Modelo con la información imperfecta (Modelo 1): en este modelo se utilizan valores conocidos de las forzantes y otras variables en caso de necesitarlas durante el pronóstico (valor de variables desde octubre). El tiempo de adelanto es una temporada completa.
2. Modelo con información perfecta (Modelo 2): en este modelo el tiempo de adelanto está determinado por el menor de los lags de todas las variables de la estructura de la RNA, es decir, para pronosticar la temporada completa se debe hacer por lo menos una iteración. Este modelo solo aplica para el pronóstico semanal y mensual.

Las bases del pronóstico en el modelo 1 son las siguientes:

- Se entrenará la RNA con información meteorológica disponible en el periodo 1985-2015 conforme a las indicaciones dispuestas en el subcapítulo 5.2.1.-.
- El periodo de pronóstico corresponde a la temporada de deshielo (octubre a marzo) de los años 2020 a 2050.
- Se pronostican caudales medios y volúmenes para las temporadas señaladas.
- Se consideran como variables independientes todas las variables (Q, P, T, SWE, E) hasta el 30 de septiembre del año hidrológico de pronóstico.
- Se han supuesto valores conocidos de Pp, SWE en función de la curva de duración de la precipitación y SWE acumulados en la temporada abril-septiembre. Para esto se discretizó la curva de duración como se observa en el ejemplo del caso mensual (Anexo C). Lo anterior se ha realizado para cada paso temporal de pronóstico.
- En caso de necesitar algún caudal desde el mes de octubre en adelante se utilizan los valores pronosticados.
- Se ocupan Temperaturas promedio históricas, dado que se puede considerar como una variable estacionaria.

Los volúmenes pronosticados corresponden a valores del régimen natural del Río.

Del mismo modo las bases del pronóstico en el modelo 2 son las siguientes:

- Se entrenará la RNA con información meteorológica disponible en el periodo 1985-2015 conforme a las indicaciones dispuestas en el subcapítulo 5.2.1.
- El periodo de pronóstico corresponde a la temporada de deshielo (octubre a marzo) de los años 2020 a 2050.
- Se pronostican caudales medios y volúmenes para las temporadas señaladas con un adelanto determinado por el menor de los lags de todas las variables.
- Se consideran como variables independientes todas las variables (Q, P, T, SWE, E) hasta el día anterior del periodo a pronosticar.
- Los volúmenes pronosticados corresponden a valores del régimen natural del Río.

La cuantificación de los errores se hará de la siguiente manera:

- Se calcularán porcentuales los errores semanales, mensuales y semestrales según lo descrito en el subcapítulo 2.7.
- Adicionalmente se cuantificará el error porcentual total de la temporada.

Los vectores de pesos y bias de las redes neuronales se encuentran en el Anexo D.

5.3.2.- Pronóstico Semanal

De las sucesivas evaluaciones de los subconjuntos de predictores se seleccionó como una modificación del subconjunto de variables 37 tanto para el modelo 1 y 2. En la Tabla 5.6 se señalan las variables consideradas para el modelo 1 mientras en la tabla 5.7 se indican las del modelo 2.

Tabla 5.6: Variables RNA semanal modelo 1

Lag	Variable
t-2	Q
t-12	SWE
t-13	
t-14	
t-15	
t-16	
t-22	Pp
t-23	
t-24	
t-25	
t-26	
t-1	T
t-3	
t-24	
t-25	

Tabla 5.7: Variables RNA semanal modelo 2

Lag	Variable
t-2	Q
t-12	SWE
t-13	
t-14	
t-15	
t-16	
t-22	Pp
t-23	
t-24	
t-25	
t-26	
t-3	T
t-24	
t-25	

Dado que el menor de los lags corresponde a $Q(t-2)$ el pronóstico debe realizarse cada 2 semanas. Se observa una alta dependencia Inter-estacional, donde la precipitación y equivalente en nieve caídos en invierno y comienzo de temporada están presentes como predictores.

Bajo esta configuración el resumen del porcentaje de temporadas con error bajo el 20% (DGA, 2018) para los 2 modelos se observa en las tablas 5.8 y 5.9. En la Tabla 5.8 se muestra la importancia de entrenar adicionalmente con los periodos históricos provenientes de los GCM, en caso de haber tomado los pesos provenientes de la calibración de la serie HIST los resultados no serían buenos en comparación con las demás series (ej: CCSM4). Lo anterior es producto de los supuestos tomados para el modelo 1.

Por otra parte, los resultados en el modelo 2 son satisfactorios para todos los sets de datos en evaluación.

Tabla 5.8: Resumen de temporada con error bajo el 20% en el pronóstico semanal modelo 1

Porcentaje de temporadas con error total de la temporada bajo el 20%					
GCM	SET DE ENTRENAMIENTO				
	CCSM4	CSIRO	IPSL	MIROC	HIST
CCSM4	90%	68%	81%	65%	71%
CSIRO	94%	77%	68%	61%	68%
IPSL	81%	55%	77%	52%	74%
MIROC	77%	68%	74%	81%	68%

Tabla 5.9: Resumen de temporada con error bajo el 20% en el pronóstico semanal modelo 2

Porcentaje de temporadas con error total de la temporada bajo el 20%					
GCM	SET DE ENTRENAMIENTO				
	CCSM4	CSIRO	IPSL	MIROC	HIST
CCSM4	100%	100%	100%	100%	100%
CSIRO	100%	100%	100%	94%	100%
IPSL	100%	100%	100%	100%	100%
MIROC	100%	100%	100%	94%	100%

De esta manera, para el modelo 1 se elige el set de entrenamiento CCSM4 y para el modelo 2 se elige el set HIST, cuyos resultados de calibración se presentan en las figuras 5.16 a 5.19. Los indicadores de desempeño del entrenamiento y la parsimonia (AIC) se pueden ver en la Tabla 5.10.

La regresión de los datos para cada proceso del entrenamiento, resultan en valores de R cercanos 1 y los histogramas de errores tienen distribución aparentemente normal centrada en 0. Los valores de NSE, KGE y NMSE tanto en entrenamiento como validación cruzada son satisfactorios, indicando buenos resultados.

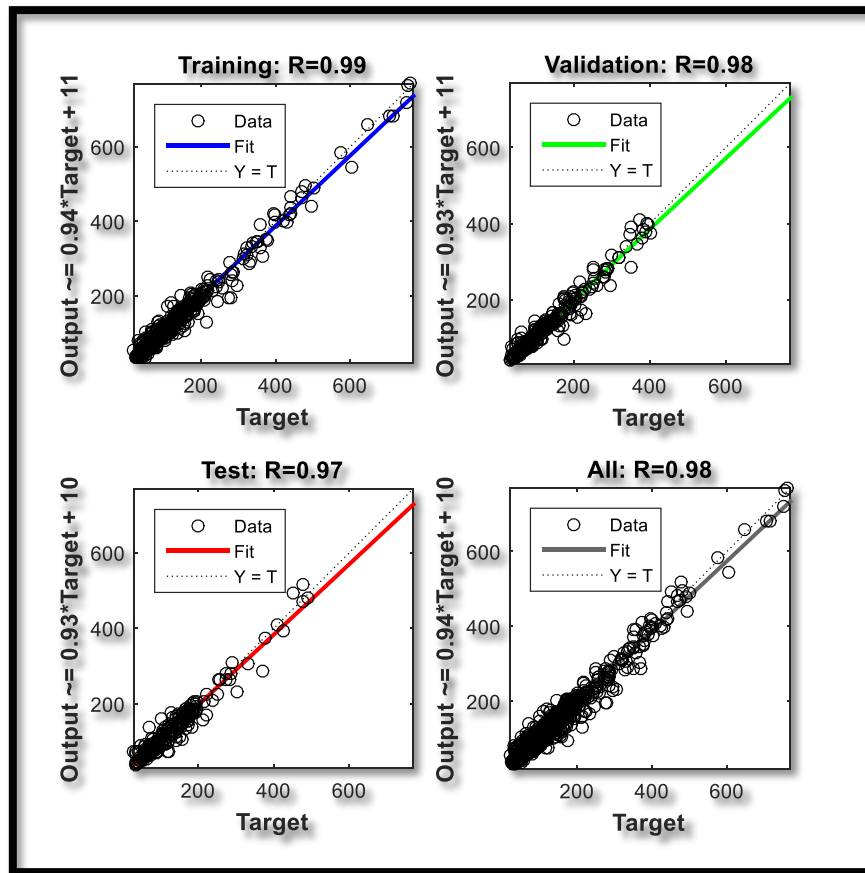


Figura 5.16: Regresión del entrenamiento, prueba y validación RNA semanal modelo 1

En general, los resultados de ambos modelos son muy parecidos tanto en la regresión como en las métricas de evaluación. Algunas diferencias se presentan en los histogramas de errores si se observan la magnitud de estos.

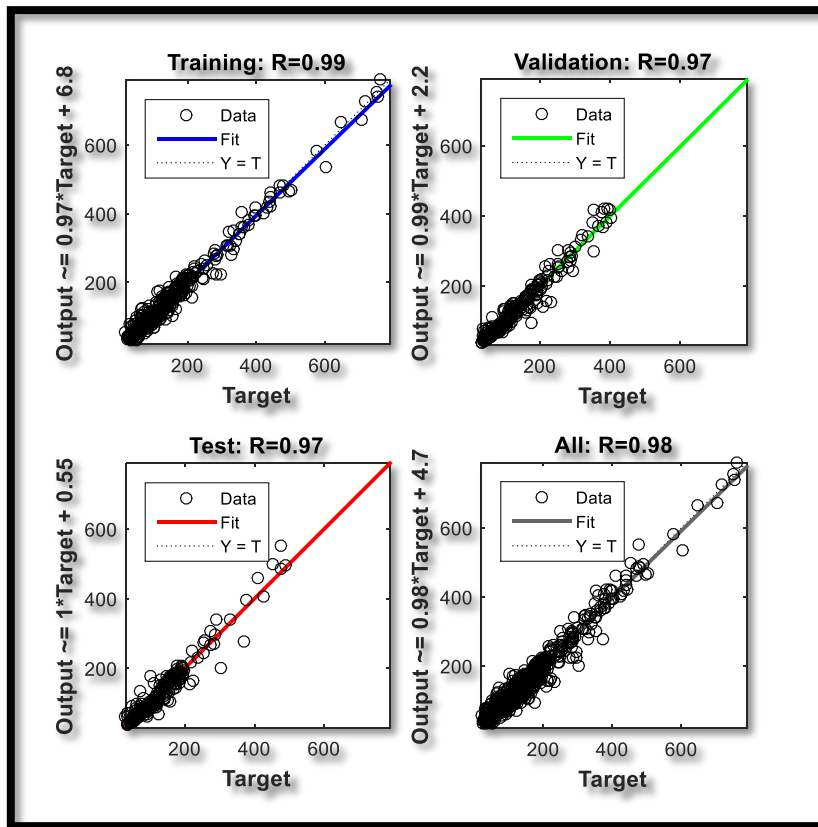


Figura 5.17: Regresión del entrenamiento, prueba y validación RNA semanal modelo 2

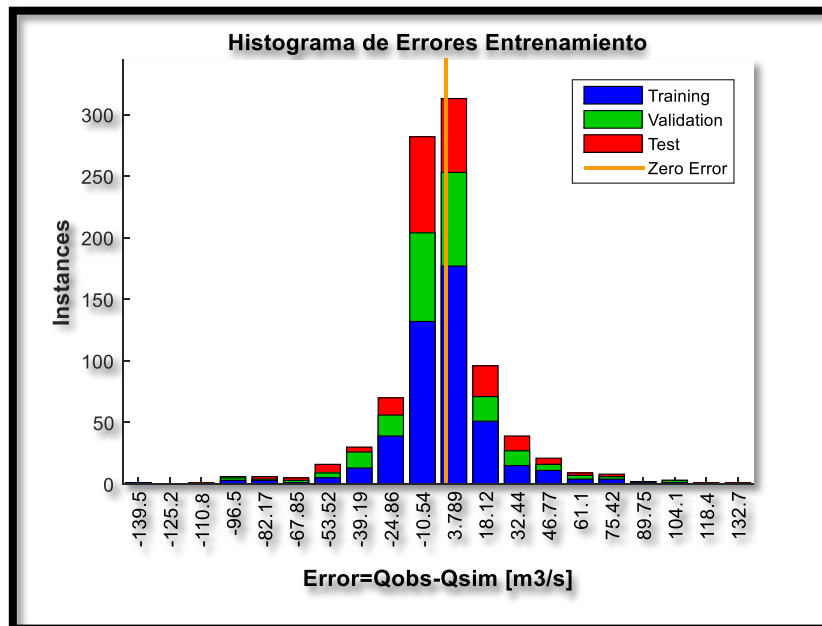


Figura 5.18: Histograma de Errores RNA semanal modelo 1

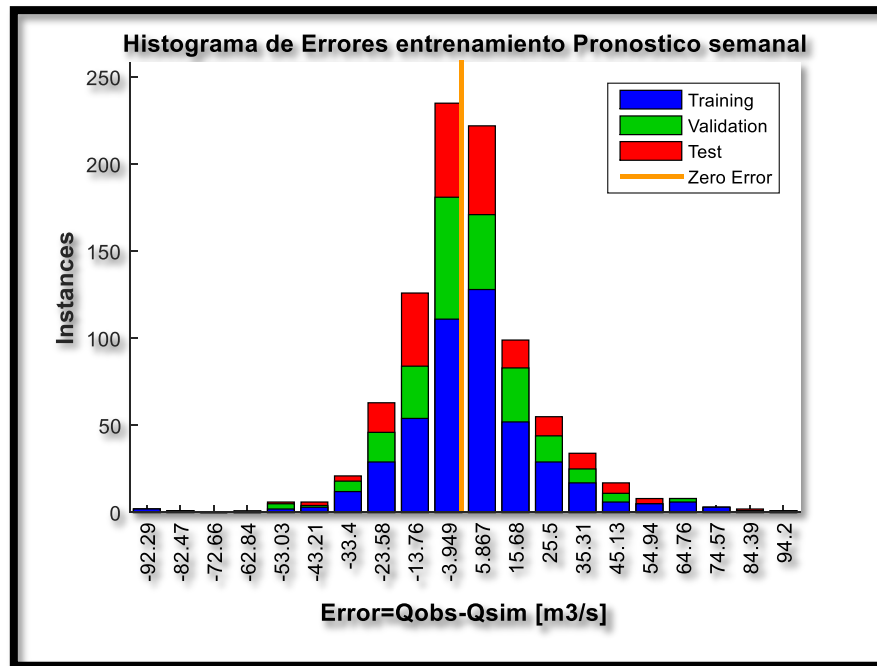


Figura 5.19: Histograma de Errores RNA semanal modelo 2

Tabla 5.10: Indicadores de desempeño entrenamiento RNA Semanal

Modelo	NSE [%]	KGE [%]	NMSE [%]	AIC [-]
1	96,5	98,0	2,1	8110
2	96,5	96,7	2,1	8109

Los resultados del pronóstico a nivel semanal correspondiente a los 4 GCM y el set de entrenamiento elegido tanto para el modelo 1 como el modelo 2 se presentan en las figuras 5.20 a 5.27. Se grafican 5 series en cada gráfico, cada una corresponde al set de datos disponible para el entrenamiento. Adicionalmente se grafica el caudal modelado por VIC en el periodo de pronóstico (2020-2050) para cada GCM:

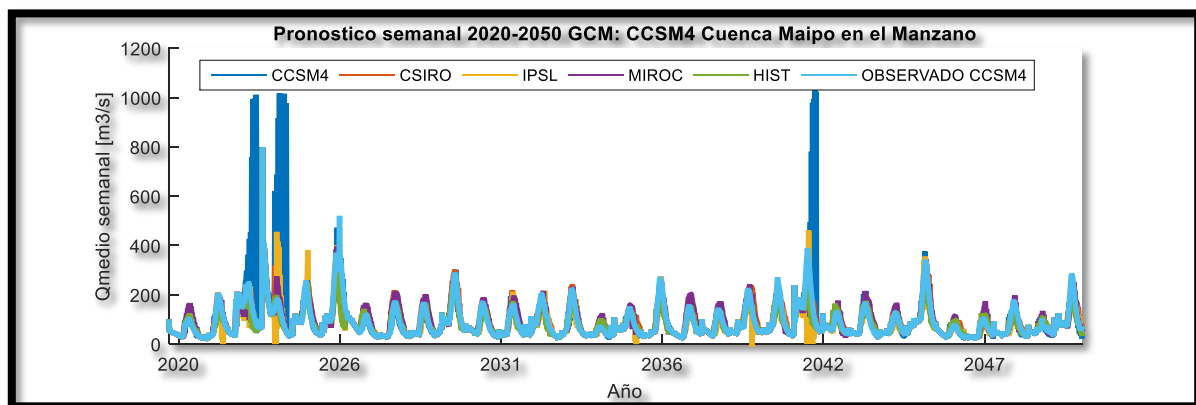


Figura 5.20: Proyección a nivel semanal GCM CCSM4 M1

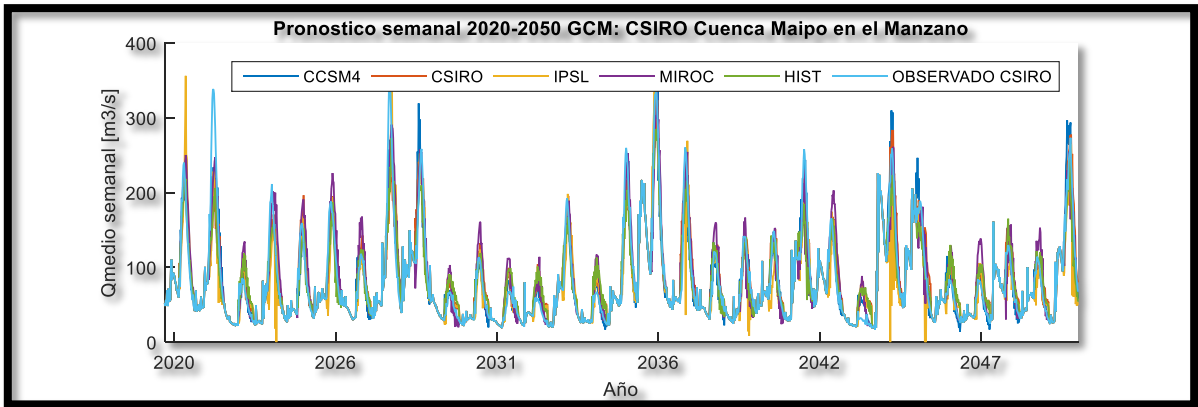


Figura 5.21: Proyección semanal GCM CSIRO M1

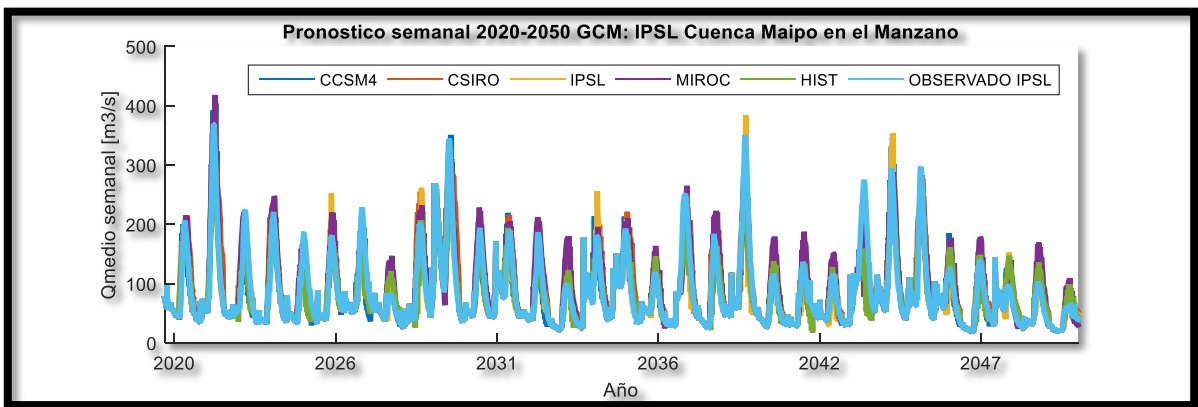


Figura 5.22: Proyección semanal GCM IPSL M1

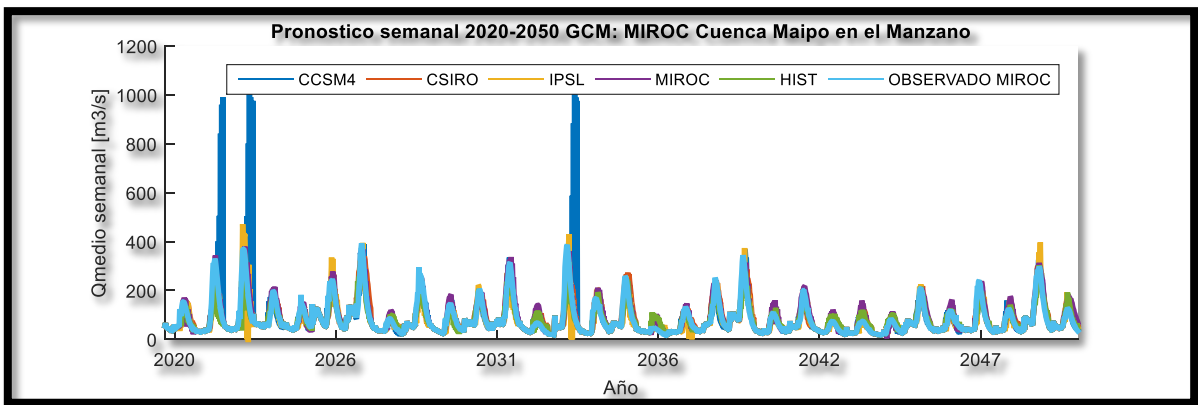


Figura 5.23: Proyección semanal GCM MIROC M1

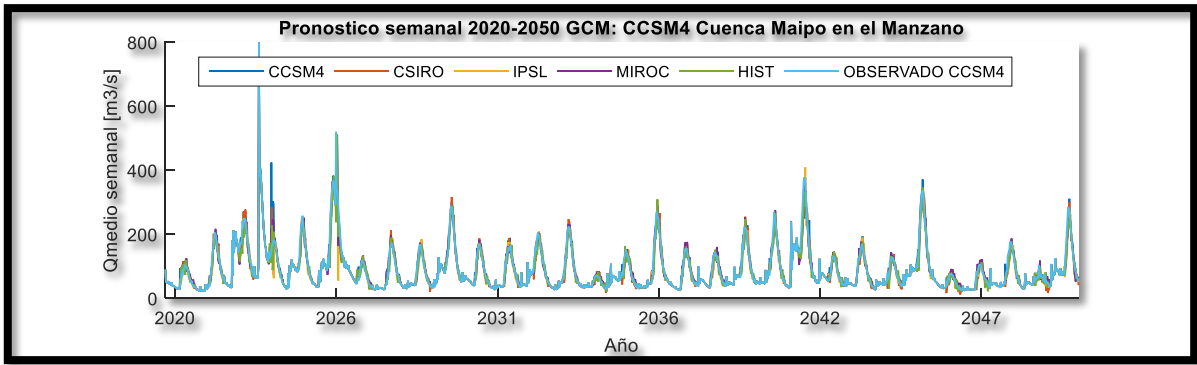


Figura 5.24: Pronóstico semanal GCM CCSM4 modelo 2

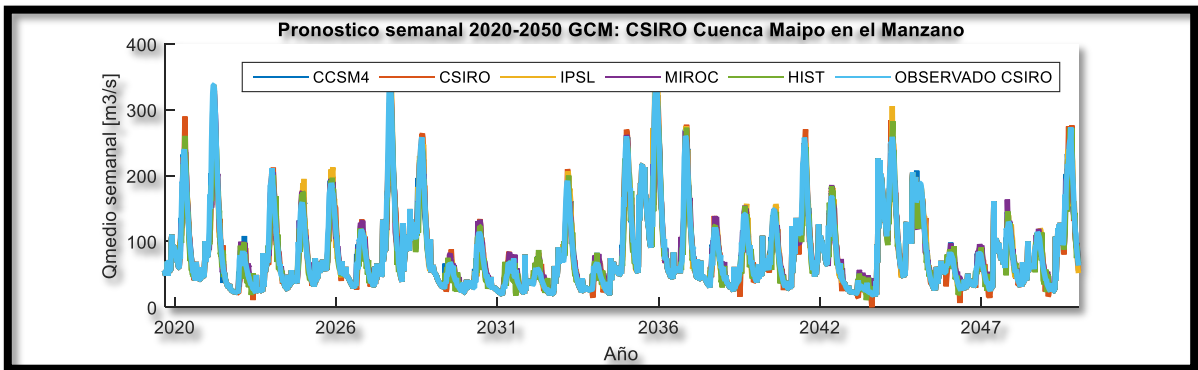


Figura 5.25: Pronóstico semanal GCM CSIRO modelo 2

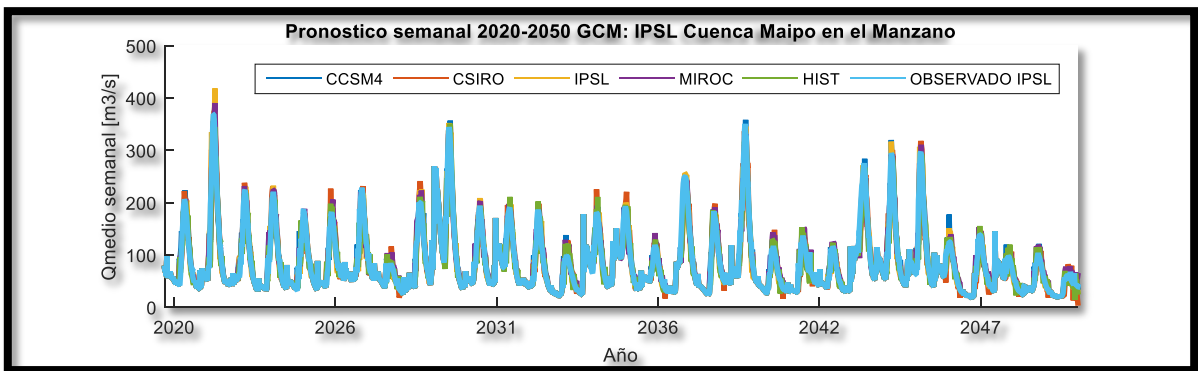


Figura 5.26: Pronóstico semanal GCM IPSL modelo 2

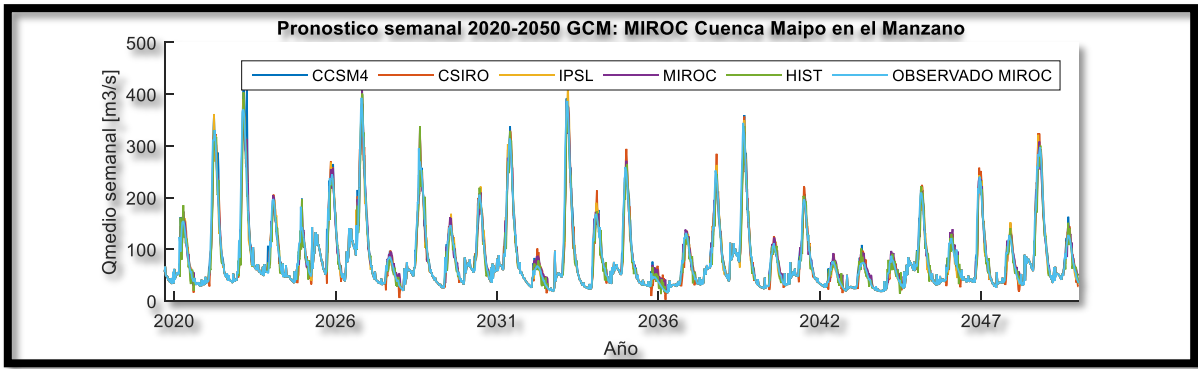


Figura 5.27: Pronóstico semanal GCM MIROC modelo 2

5.3.3.- Pronóstico Mensual

De la selección de predictores, las estructuras elegidas para cada modelo se muestran en las tablas 5.11 y 5.12:

Tabla 5.11: Variables RNA Mensual escenario 1

Lag	Variable
t-2	Q
t-2	SWE

Tabla 5.12: Variables RNA mensual modelo 2

Lag	Variable
t-1	Q
t-2	
t-2	SWE
t-4	Pp
t-5	

Dado que en el modelo 2 el menor lag es Q(t-1), el pronóstico debe realizarse mensualmente para obtener la temporada completa. Bajo esta configuración el resumen del porcentaje de temporadas con error bajo el 20% (DGA, 2018) se muestran en la Tabla 5.13. Se observa que en los resultados del modelo 1, tanto el set CCSM4 como MIROC entregan los mejores resultados en términos del mínimo error por temporada.

Tabla 5.13: Porcentaje de temporadas con error menor al 20% modelo 1

Porcentaje de temporadas con error total de la temporada bajo el 20%					
Set	CCSM4	CSIRO	IPSL	MIROC	HIST
1	81%	90%	77%	84%	52%
2	87%	84%	77%	87%	42%
3	84%	81%	74%	90%	74%
4	84%	84%	77%	81%	52%

Por otro lado, en el modelo 2 (Tabla 5.14) los resultados son mejores que el modelo 1 para todos los sets de datos. Es importante notar la mejora del pronóstico al entrenar el set HIST bajo los supuestos del modelo 2, donde el número de temporadas con error bajo el 20% aumenta considerablemente (de un 50% a un 95% promedio de temporadas con error bajo el 20%).

Tabla 5.14: Porcentaje de temporadas con error menor al 20% modelo 2

Porcentaje de temporadas con error total de la temporada bajo el 20%					
Set	CCSM4	CSIRO	IPSL	MIROC	HIST
1	90%	100%	100%	100%	94%
2	81%	90%	97%	84%	90%
3	97%	100%	100%	100%	97%
4	90%	94%	97%	90%	97%

En vista de los resultados, para ambos modelos el set de datos elegido es CSIRO. Los resultados de la calibración se muestran en la Figura 5.28 y en la Figura 5.29. Contrastando los resultados obtenidos, el modelo 2 entrega mejores resultados con valores de R por sobre 0.95. En cambio, el modelo 1 solo entrego valores de R del orden de 0.88.

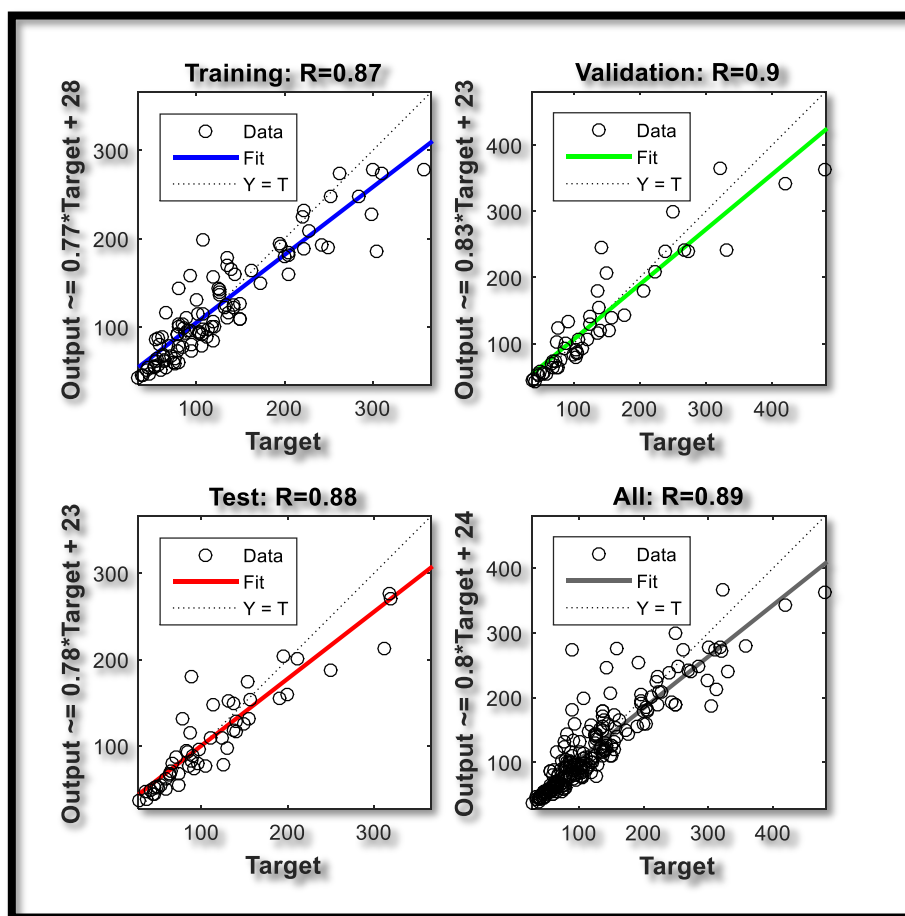


Figura 5.28: Regresión entrenamiento RNA Mensual modelo 1

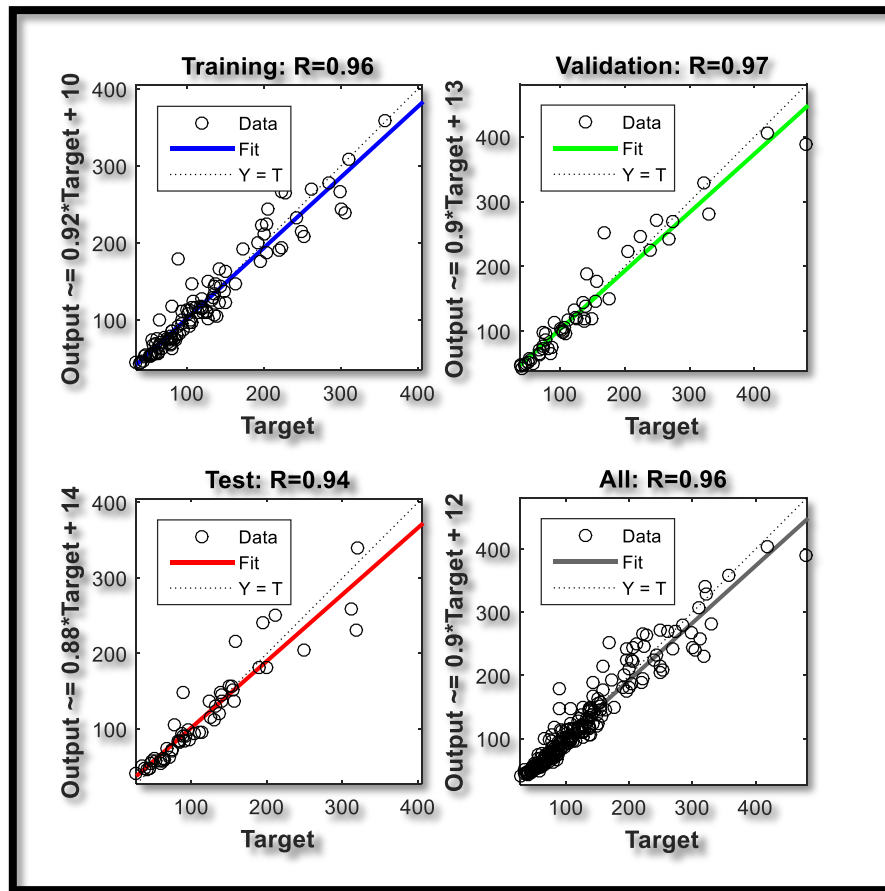


Figura 5.29: Regresión entrenamiento RNA Mensual. Proyecciones GCM CSIRO. Modelo 2

En general, los histogramas de errores del entrenamiento tienen la misma forma. Cabe mencionar, que en el modelo 1 la magnitud de los errores es mayores que en el modelo 2.

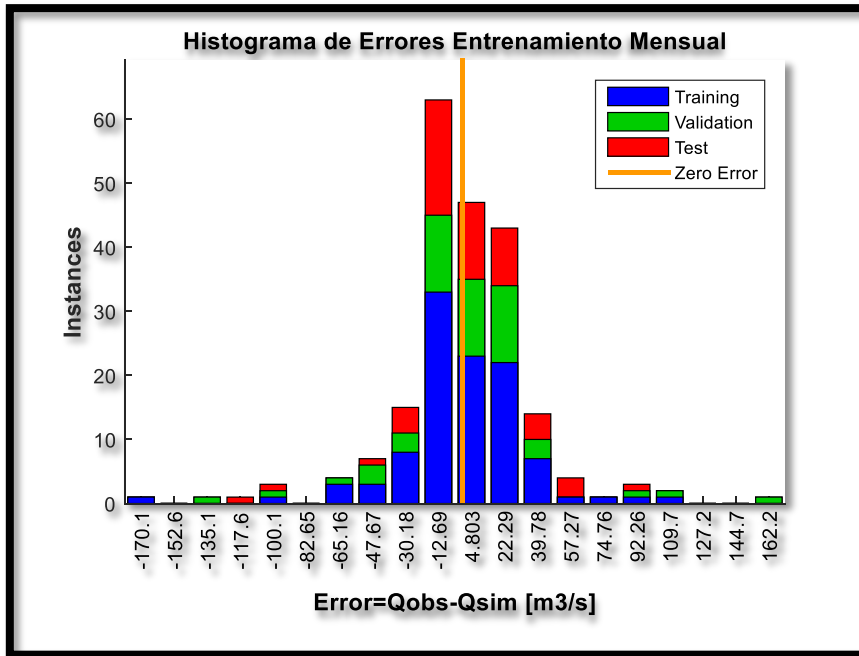


Figura 5.30: Histograma de errores entrenamiento Mensual modelo 1

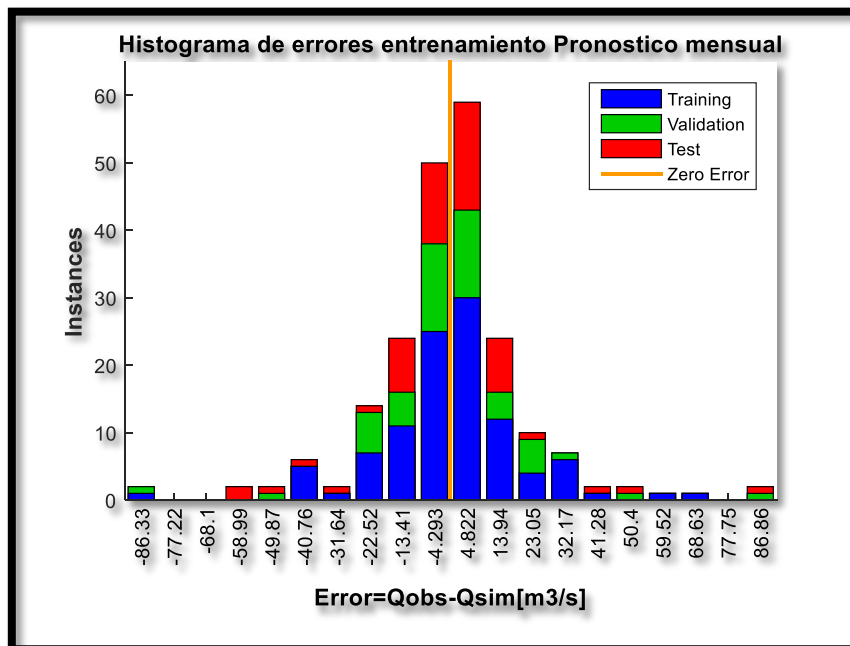


Figura 5.31: Histograma de errores entrenamiento RNA Mensual modelo 2

Los indicadores de desempeño del entrenamiento y la parsimonia (AIC) para ambos modelos se pueden ver en la Tabla 5.15.

Tabla 5.15: Indicadores de desempeño del entrenamiento RNA Mensual modelo 1

Modelo	NSE [%]	KGE [%]	NMSE [%]	AIC [-]
1	77	81	9	1550
2	92	93	3	1551

Los resultados del pronóstico a nivel semanal correspondiente a los 4 GCM y el set de entrenamiento elegido tanto para el modelo 1 como el modelo 2 se presentan en las Figuras 5.32 a 5.39. Se grafican 5 series en cada gráfico, cada una corresponde al set de datos disponible para el entrenamiento. Adicionalmente se grafica el caudal modelado por VIC en el periodo de pronóstico (2020-2050) para cada GCM:

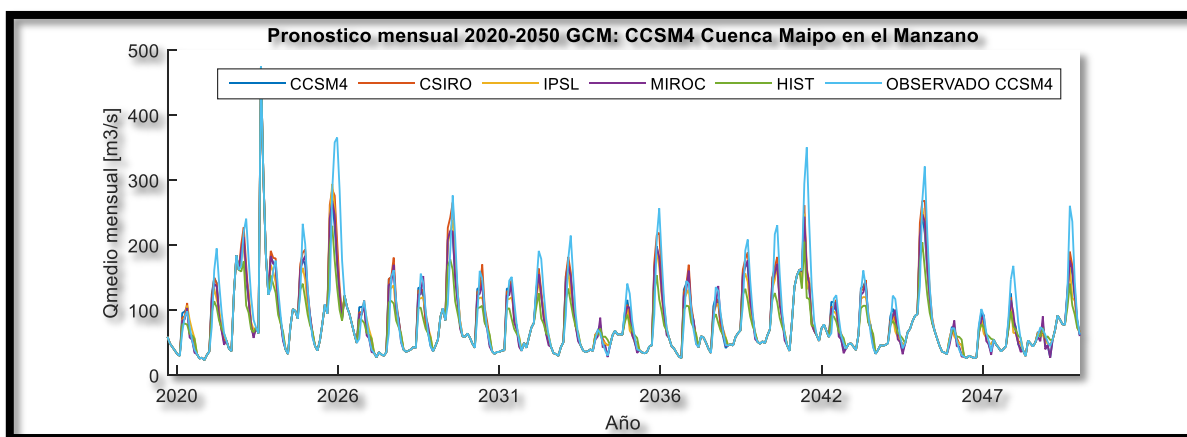


Figura 5.32: Pronóstico mensual GCM CCSM4 modelo 1

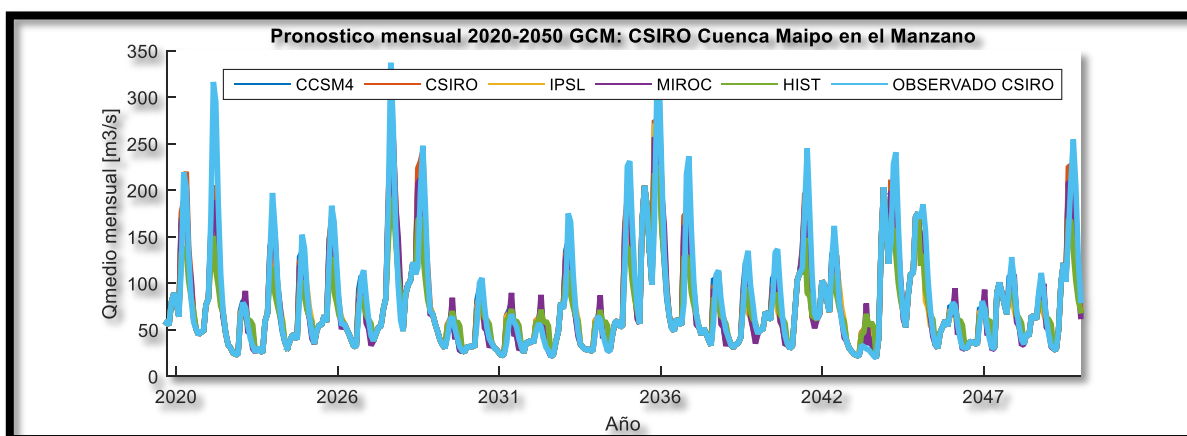


Figura 5.33: Pronóstico mensual GCM CSIRO modelo 1

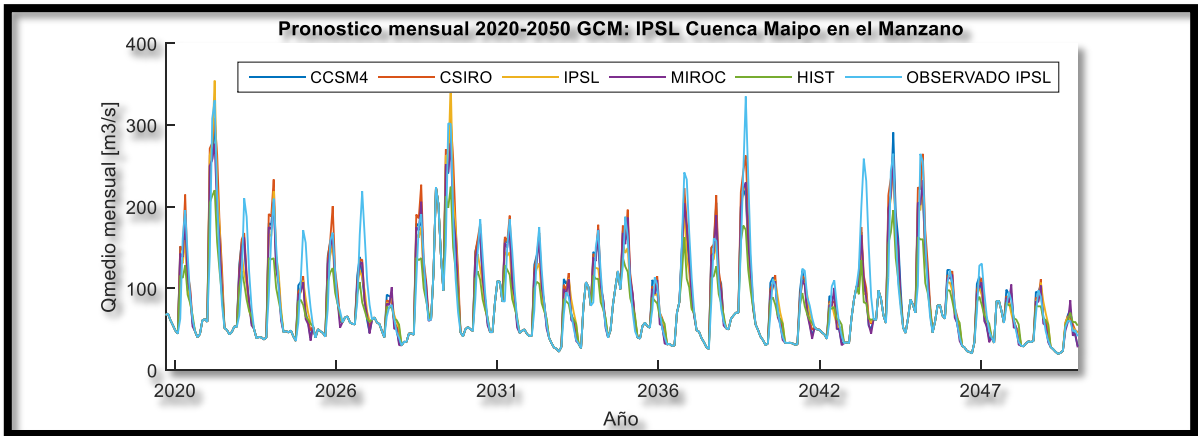


Figura 5.34: Pronóstico mensual GCM IPSL modelo 1

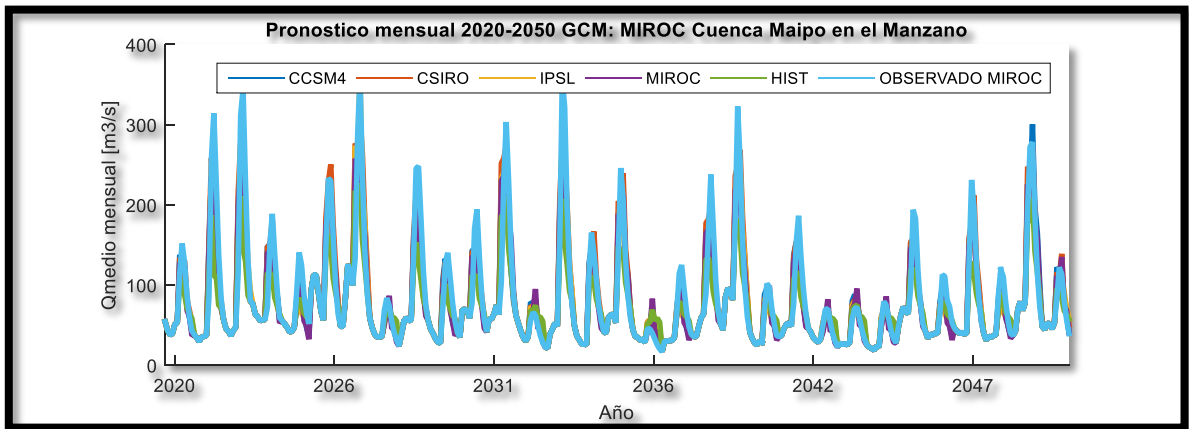


Figura 5.35: Pronóstico mensual GCM MIROC modelo 1

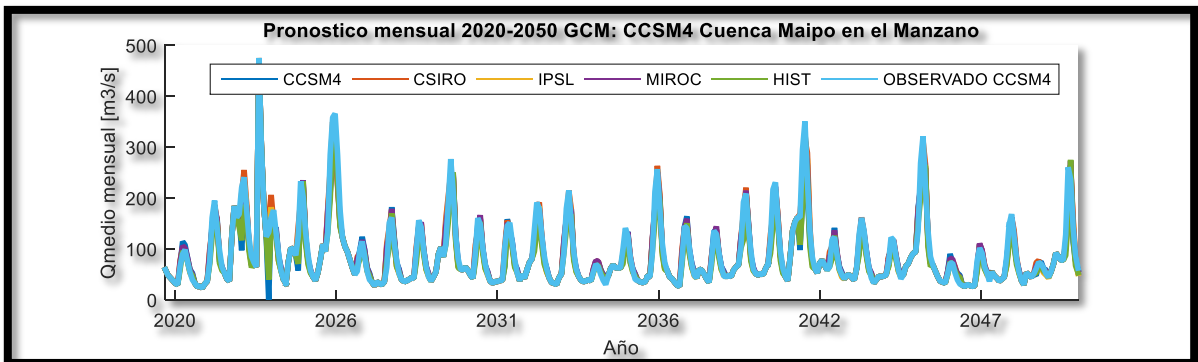


Figura 5.36: Pronóstico mensual GCM CCSM4 modelo 2

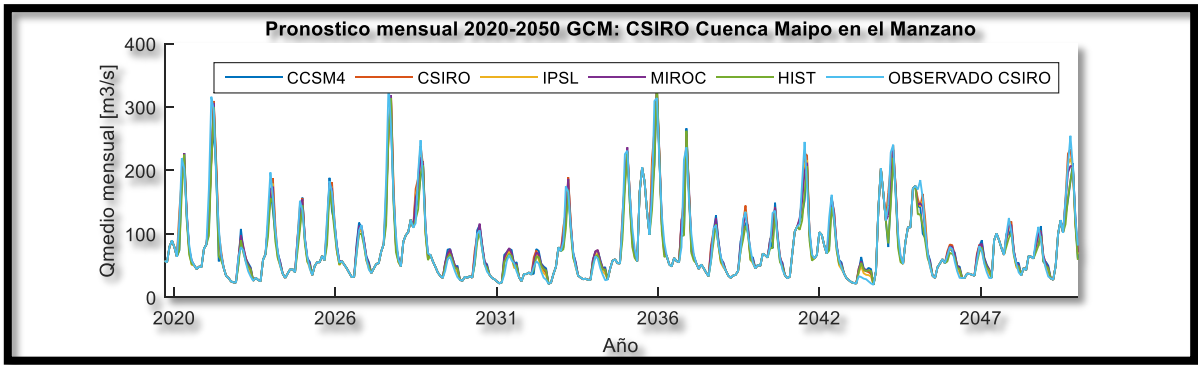


Figura 5.37: Pronóstico mensual GCM CSIRO modelo 2

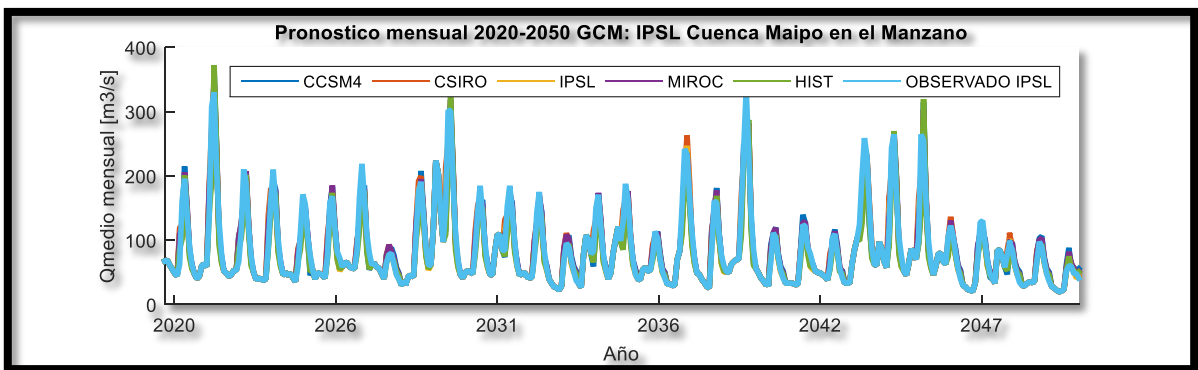


Figura 5.38: Pronóstico mensual GCM IPSL modelo 2

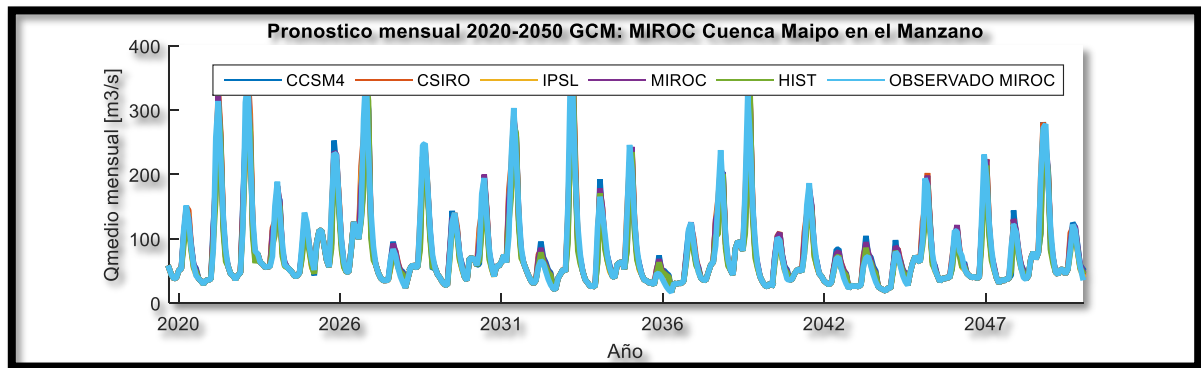


Figura 5.39: Pronóstico mensual GCM MIROC modelo 2

5.3.4.- Pronóstico Semestral

De las sucesivas evaluaciones de los conjuntos de predictores se seleccionó la siguiente configuración de variables (Tabla 5.16):

Tabla 5.16: Arquitectura RNA Semestral

Lag	Variable
t-1	Q
t-1	SWE
t-1	P

El porcentaje de temporadas con error bajo 20% se muestra en la Tabla 5.17:

Tabla 5.17: Porcentaje de temporadas con error bajo el 20%

Porcentaje de temporadas con error total de la temporada bajo el 20%					
Set	CCSM4	CSIRO	IPSL	MIROC	HIST
1	97%	87%	81%	55%	87%
2	87%	77%	65%	35%	61%
3	84%	90%	77%	65%	81%
4	94%	90%	84%	65%	90%

De manera que el set de entrenamiento elegido será CCSM4. Los resultados de la calibración se observan en la Figura 5.40 y Figura 5.41:

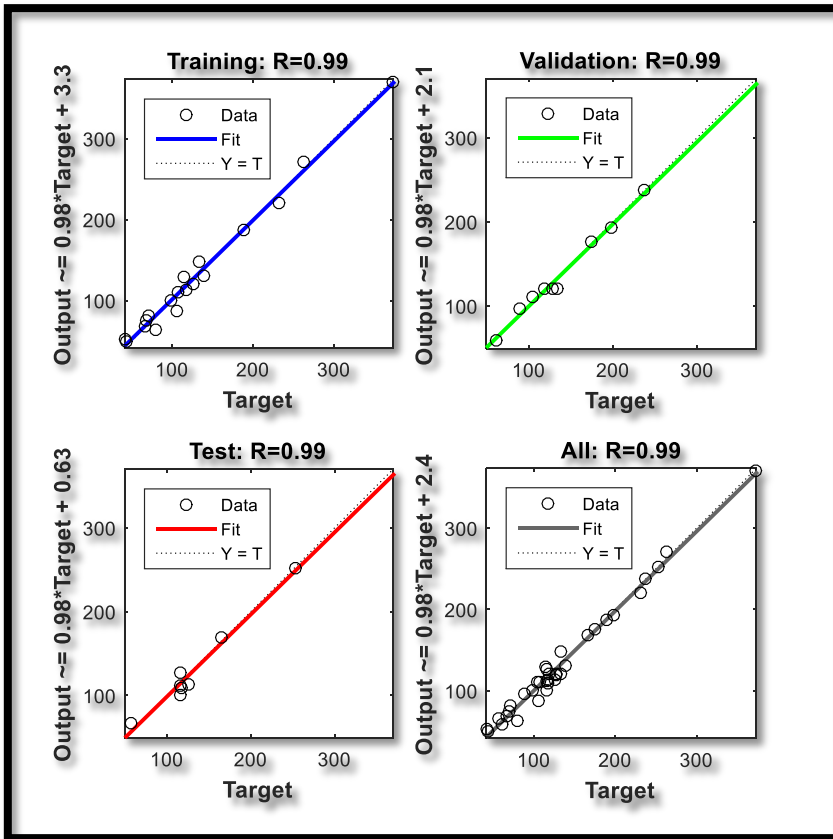


Figura 5.40: Regresión del entrenamiento RNA semestral

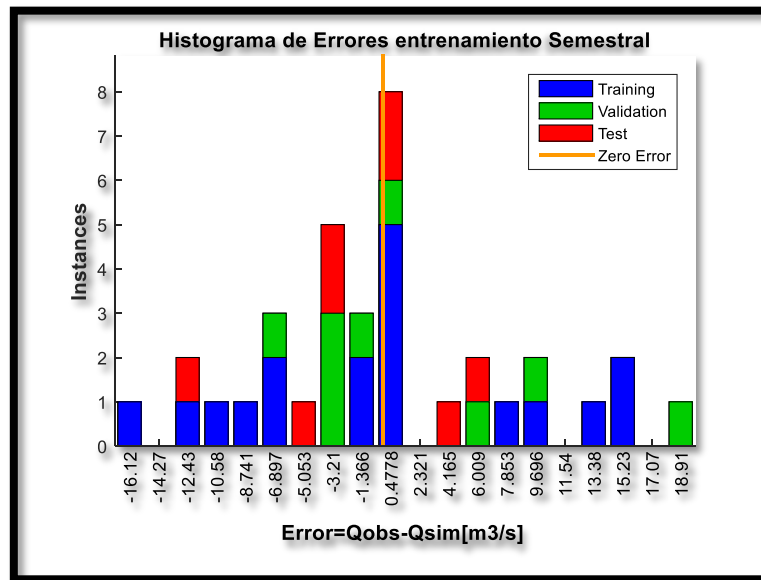


Figura 5.41: Histograma de errores entrenamiento semestral

Los indicadores de desempeño del entrenamiento se ven en la Tabla 5.18:

Tabla 5.18: Indicadores de Desempeño de entrenamiento semestral

NSE [%]	KGE [%]	NMSE [%]	AIC [-]
98	98	2	219

Los resultados de la proyección a nivel semestral utilizando el set de datos CCSM4 se presentan en las figuras 5.42 a 5.45:

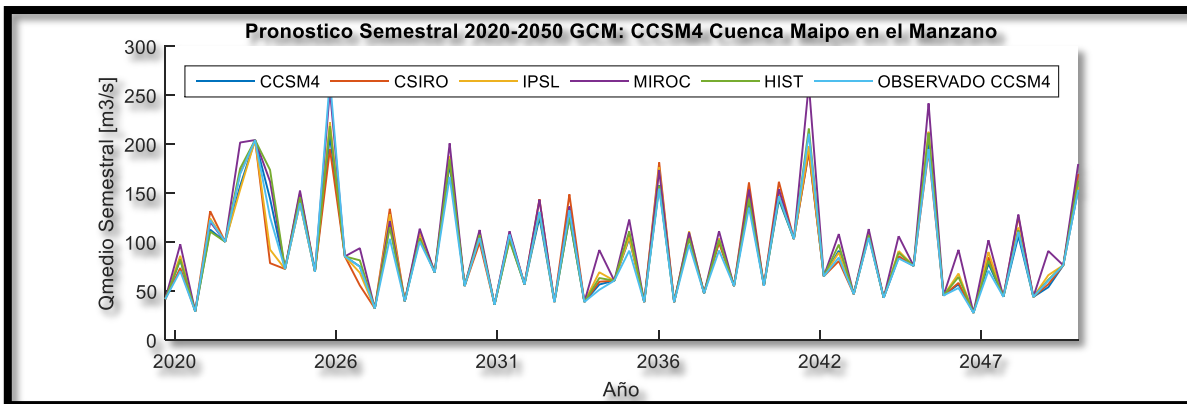


Figura 5.42: Pronóstico semestral GCM CCSM4

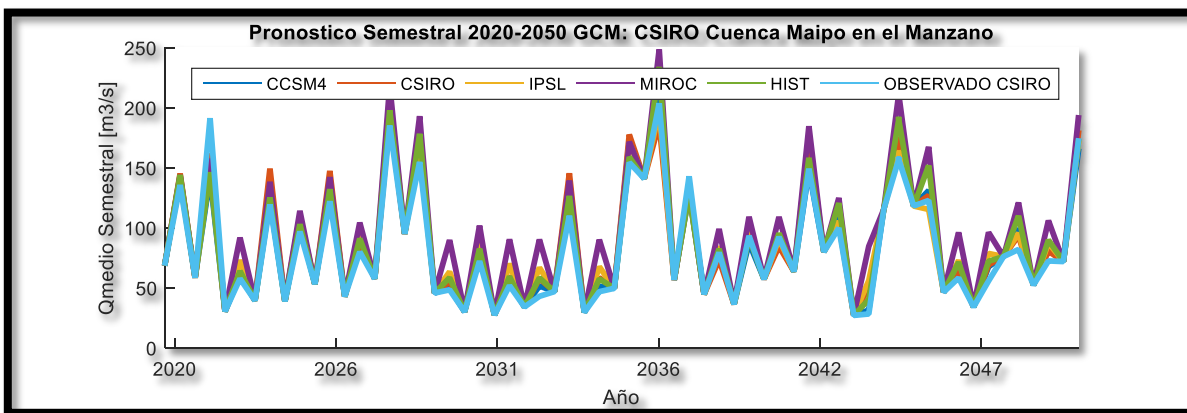


Figura 5.43: Pronóstico semestral GCM CSIRO

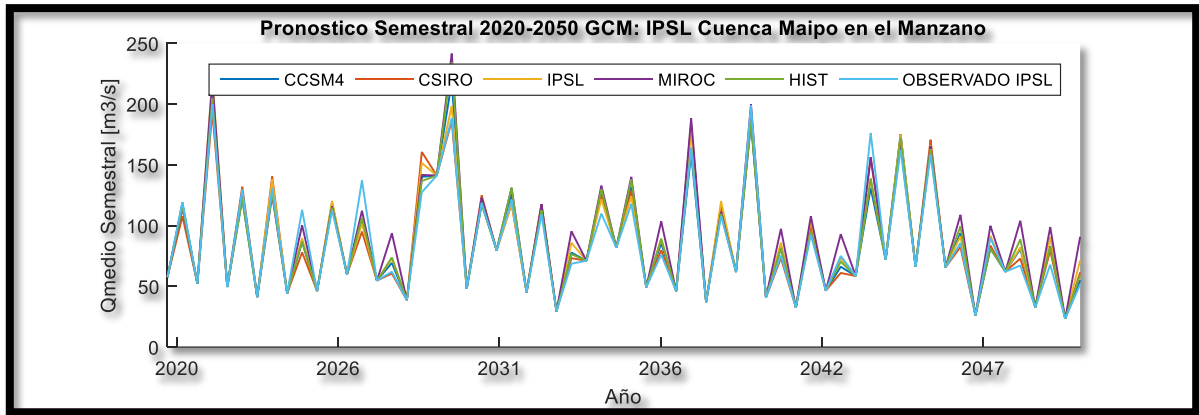


Figura 5.44: Pronóstico semestral GCM IPSL

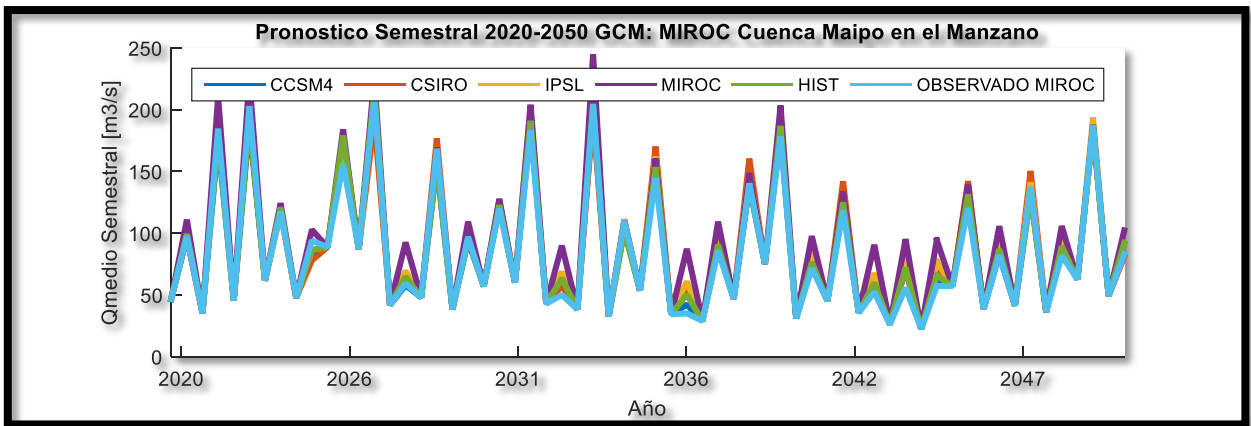


Figura 5.45: Pronóstico semestral GCM MIROC

5.4.- Desempeño de las RNA por GCM

5.4.1.- Nivel Semanal

En las figuras 5.46 y 5.47 se presenta la distribución de errores promedio por semana del periodo de deshielo (desde la semana 27 a la 52) de los años pronosticados (2020-2050). Se puede notar que en el modelo 1 no existe ningún tipo de comportamiento marcado, en efecto el error medio se mantiene en el rango de 10% a 25%. Por otra parte, en el modelo 2 se puede notar que el error disminuye a medida que avanza la temporada hasta llegar a mediados de esta. Pasado este punto el error aumenta del 5% al 20% hasta finalizar la temporada.

Los histogramas de errores se encuentran en el Anexo E.

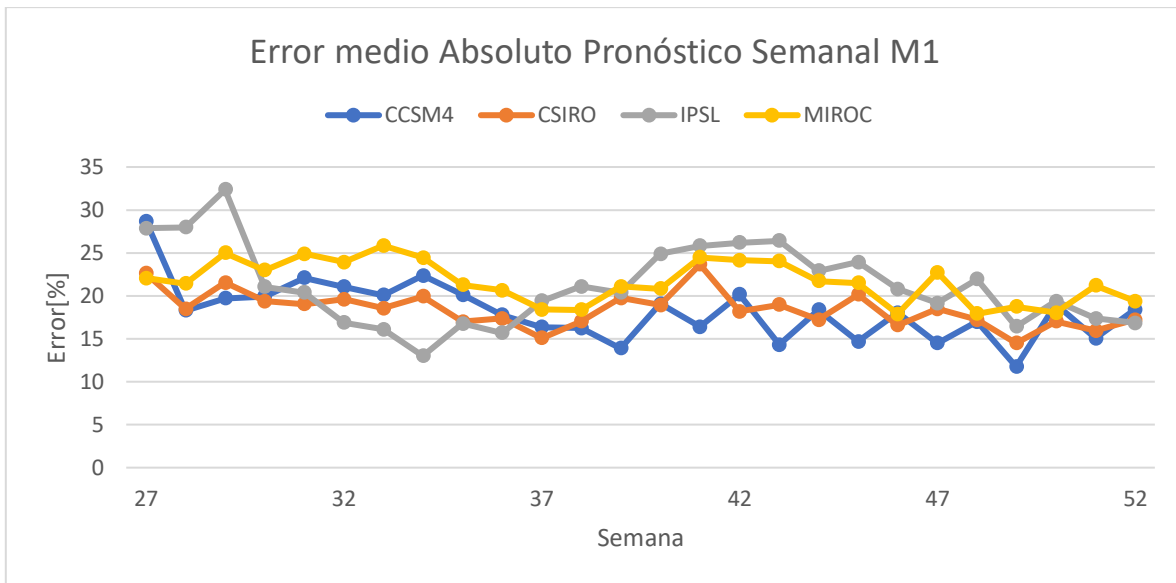


Figura 5.46: Error medio porcentual pronóstico semanal M1

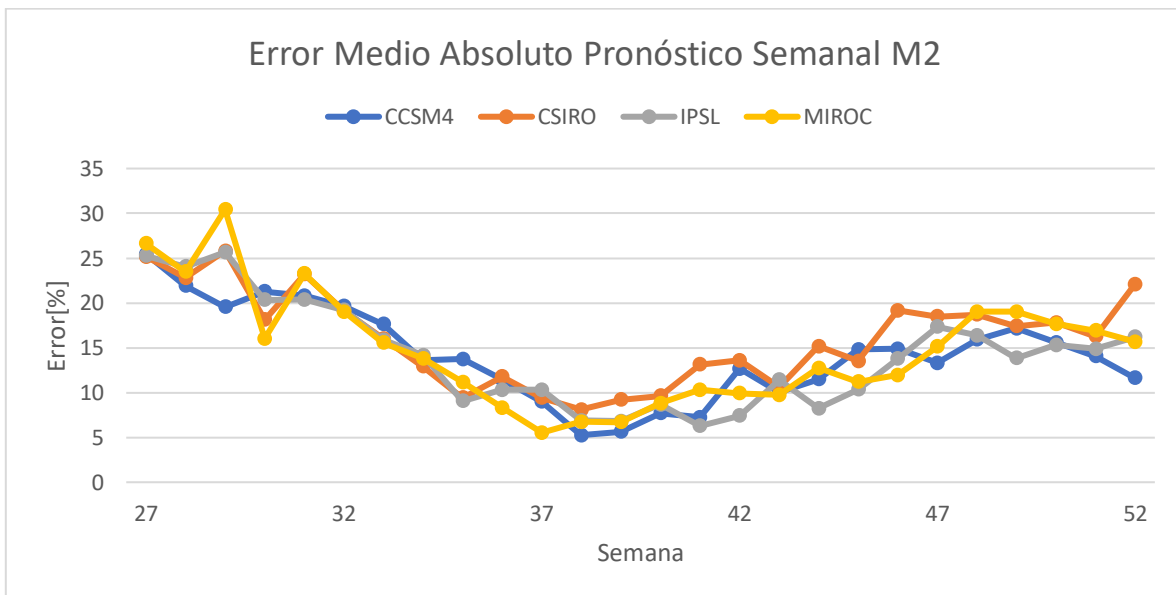


Figura 5.47: Error medio porcentual pronóstico semanal M2

En los pronósticos anuales para la temporada de deshielo realizados por la DGA (DGA, 2018) se indica como aceptable un 20% error en el volumen total pronóstico para la temporada. Dicho lo anterior se presenta el histograma de errores por temporada (Figura 5.48 y Figura 5.49) para el modelo 1 y modelo 2. En el modelo 1 los errores se distribuyen uniformemente con la

mayoría de las temporadas con error entre -10% a 20%. Por otra parte, en el modelo 2 los errores se encuentran todos salvo algunas excepciones en torno al 10%.

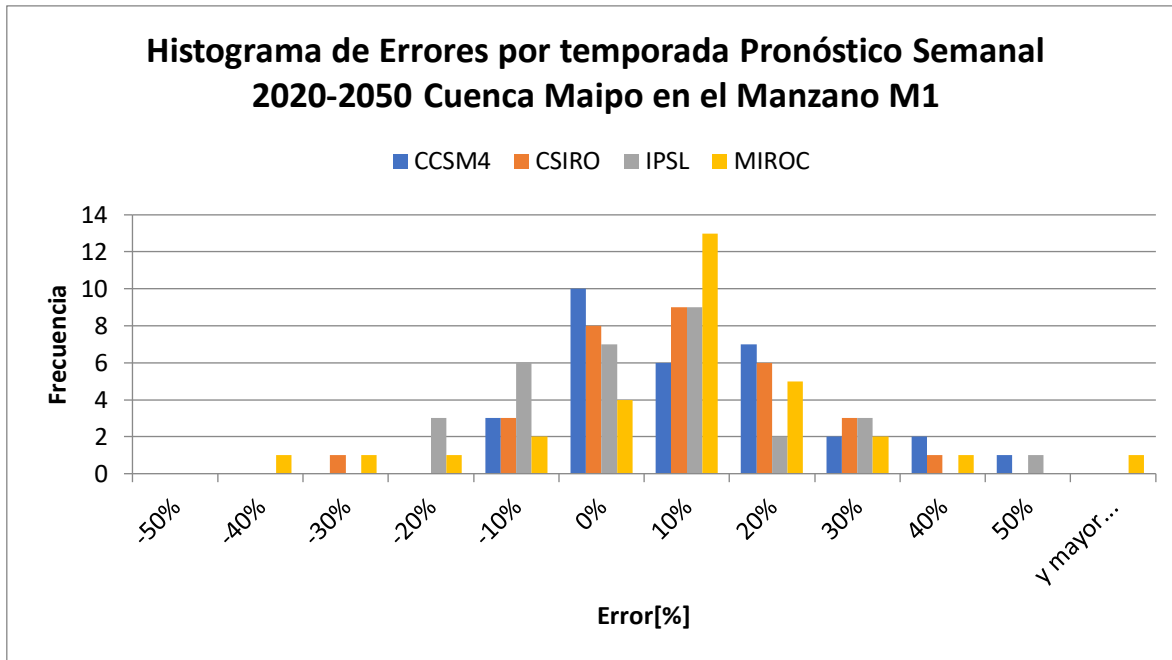


Figura 5.48: Histograma de errores por temporada pronóstico semanal modelo 1

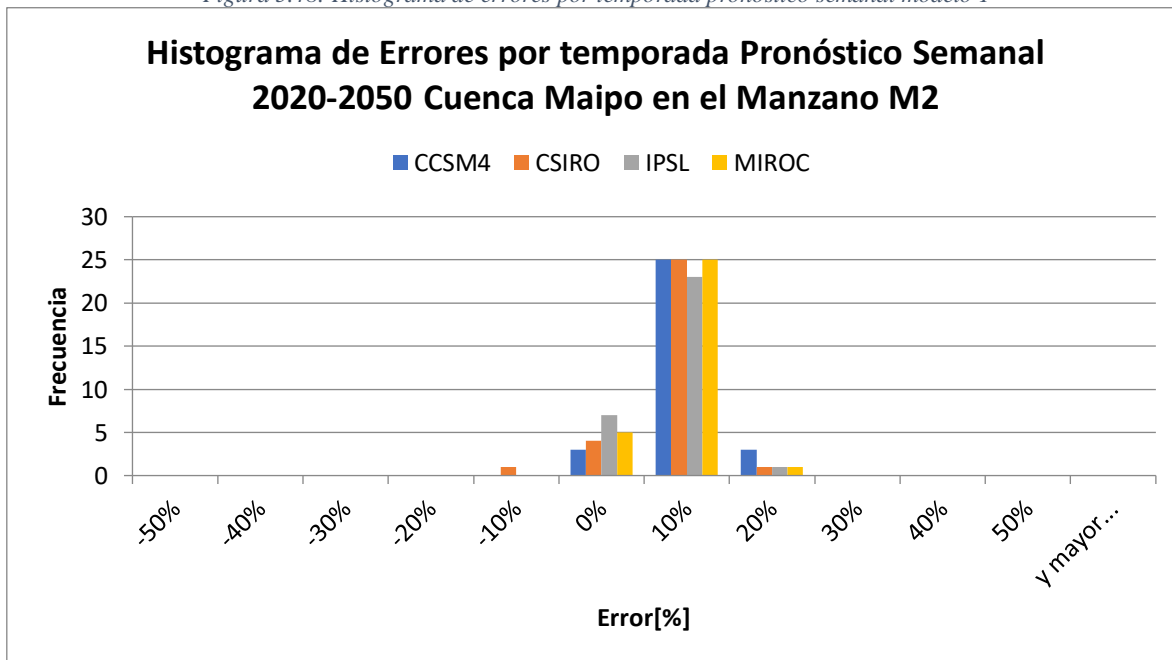


Figura 5.49: Histograma de errores pronóstico semanal modelo 2

5.4.2.- Nivel Mensual

En las figuras 5.50 y 5.51 se presenta la distribución de errores promedio absoluto por mes del periodo de deshielo de los años pronosticados (2020-2050). Se puede notar que en el modelo 1 no existe ningún tipo de comportamiento marcado, más aún, el error medio se mantiene en el rango de 15% a 30%. Por otra parte, en el modelo 2 se puede notar que el error disminuye a medida que avanza la temporada hasta llegar a mediados de esta. Pasado este punto el error aumenta del 5% al 20% hasta finalizar la temporada.

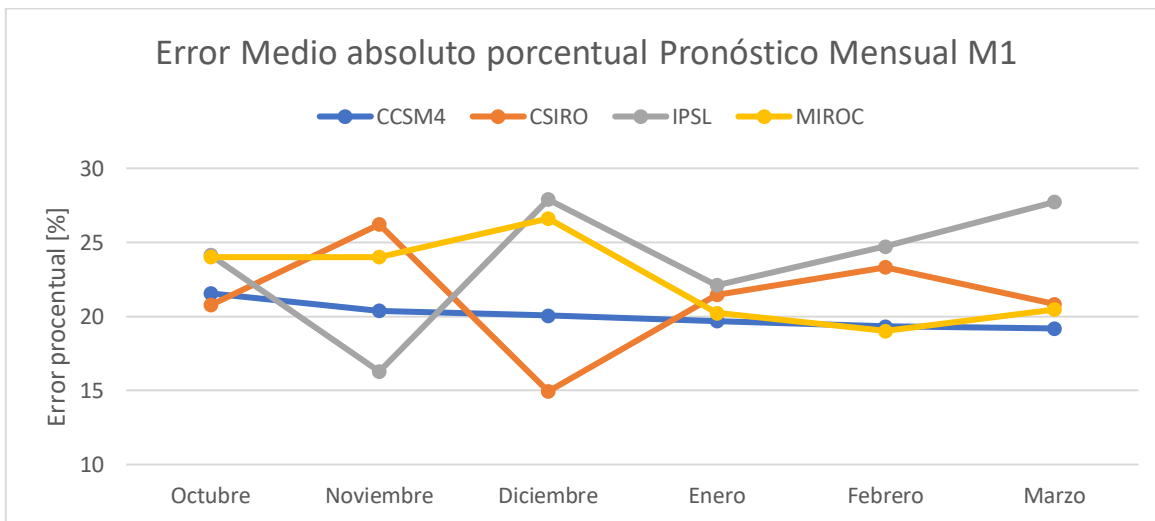


Figura 5.50: Error medio porcentual pronóstico mensual M1

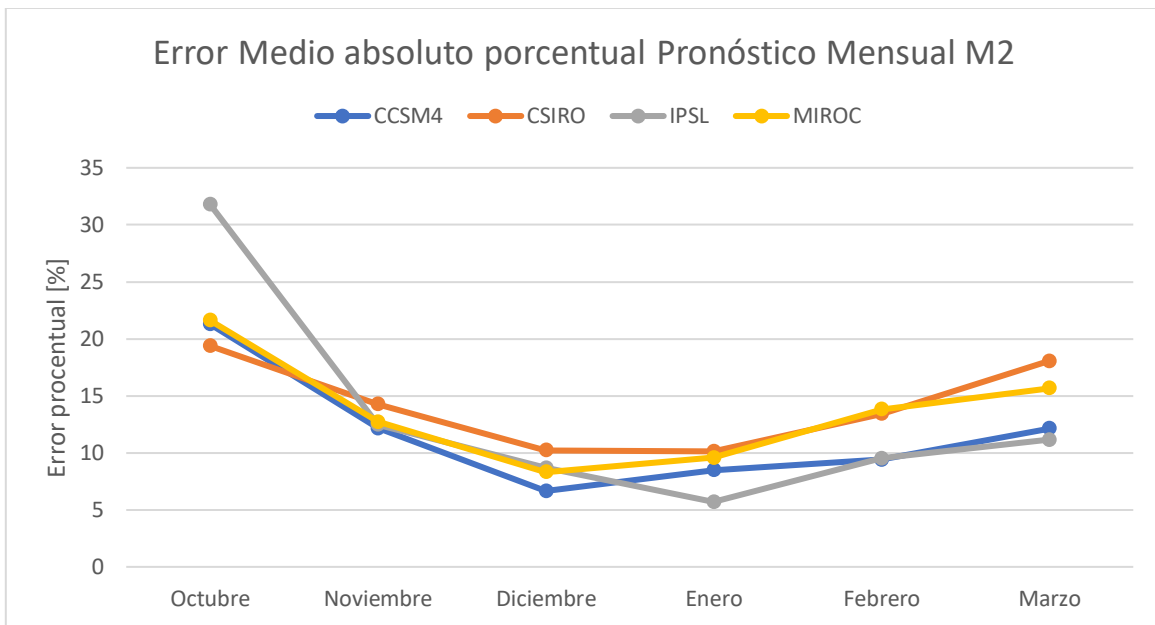


Figura 5.51: Error medio porcentual pronóstico mensual M2

En los pronósticos anuales para la temporada de deshielo realizados por la DGA (DGA, 2018) se indica como aceptable un 20% error en el volumen total pronóstico para la temporada. Análogamente al análisis semanal, se presenta el histograma de errores por temporada (figuras 5.52 y 5.53) para el modelo 1 y modelo 2. En el modelo 1 los errores se distribuyen uniformemente con

la mayoría de las temporadas con error entre -20% a 20%. Por otra parte, en el modelo 2 la mayoría de los errores se encuentran en torno al 10%.

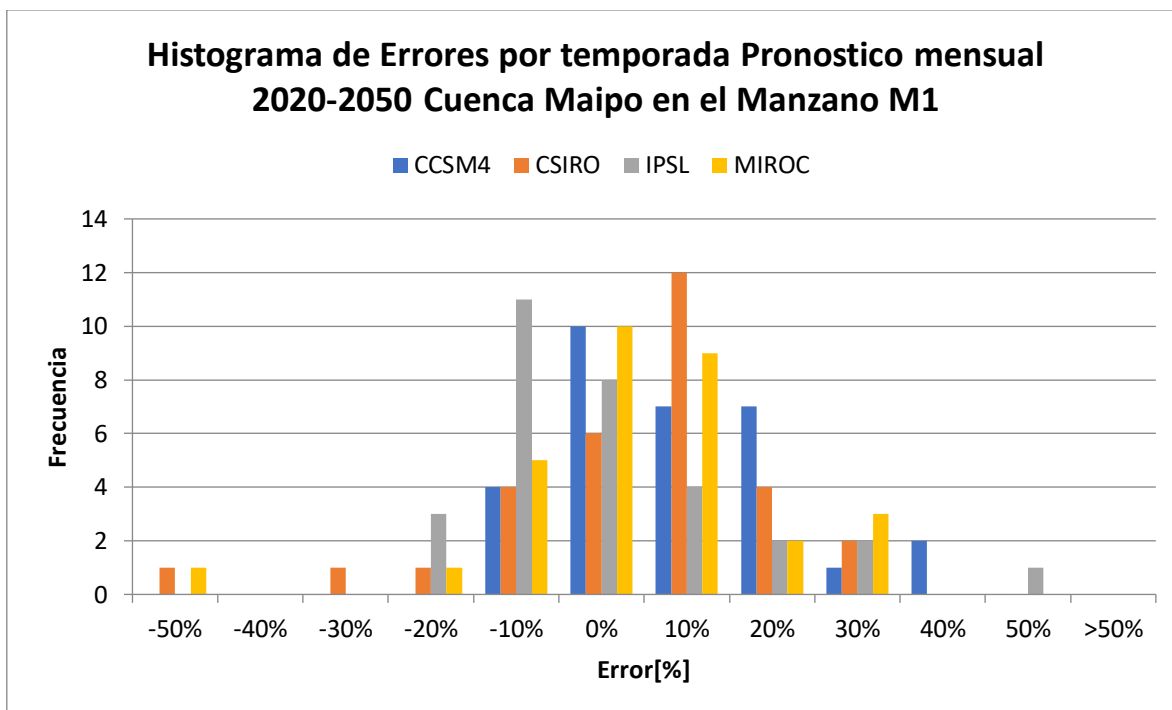


Figura 5.52: Histograma de errores por temporada pronóstico mensual modelo 1

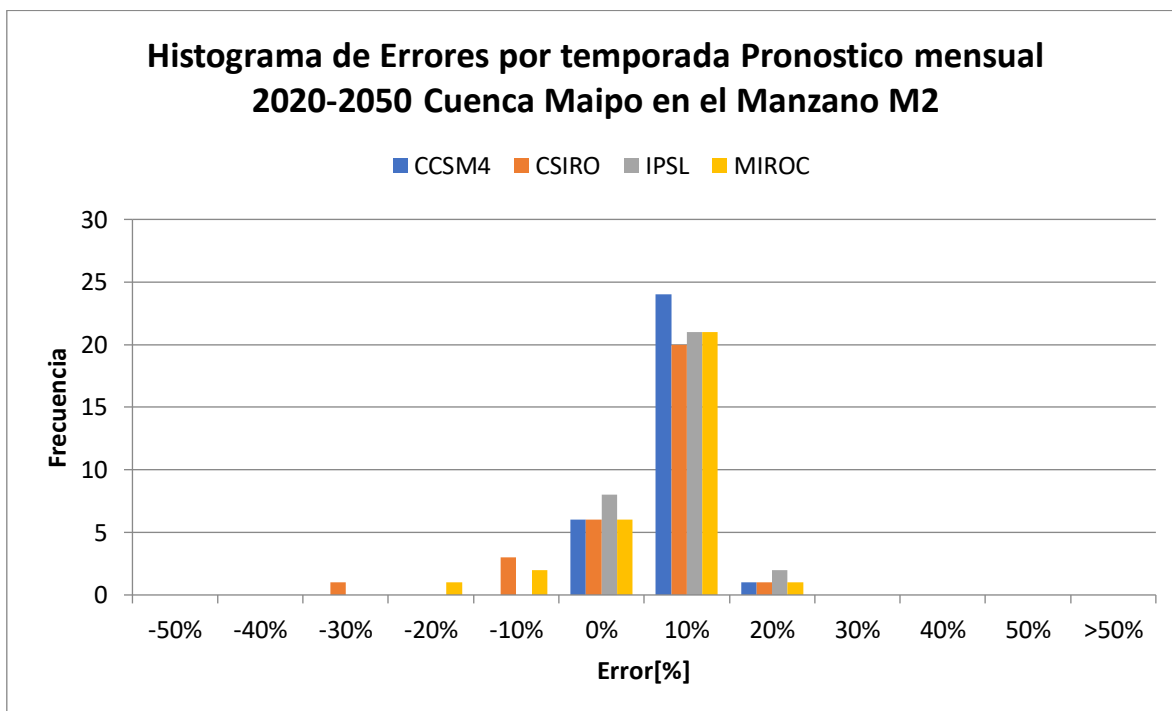


Figura 5.53: Histograma de errores pronóstico mensual modelo 2

5.4.3.- Nivel Semestral

El histograma respectivo se observa en la Figura 5.54. Se observa que los errores del pronóstico semestral son menores en comparación a los otros pasos temporales (solo M1).

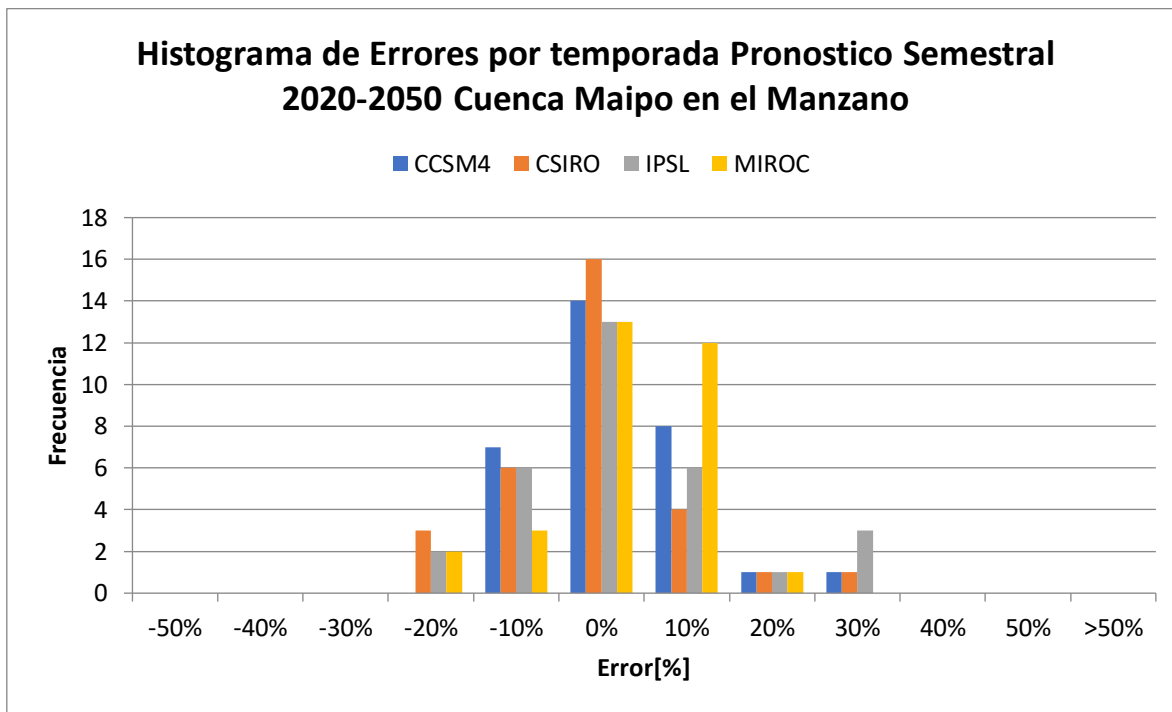


Figura 5.54: Histograma de errores por temporada pronóstico semestral

5.5.- Bandas de Incertidumbre

Luego de haber realizado el pronóstico se tienen las siguientes estructuras neuronales (Tabla 5.19):

Tabla 5.19: Estructuras RNA según modelo.

Modelo	Pronóstico
M1	Semanal
	Mensual
	Semestral
M2	Semanal
	Mensual

Con los resultados de cada red es posible construir la banda de incertidumbre asociada a los 2 modelos definidos para cada GCM utilizado en el pronóstico. De esta forma, se acumulan los volúmenes semanales y mensuales a nivel semestral para poder compararlas. En la Figura 5.55 se muestra una banda incertidumbre involucrando todas las RNA de la Tabla 5.19, donde el área sombreada corresponde a la incertidumbre.

Luego, en la Figura 5.56, se muestra la banda de incertidumbre pero solo considerando las RNA del modelo 2 y la RNA semestral. Se destaca la considerable disminución de la incertidumbre al considerar solo las RNA del modelo 2. En otras palabras, los pronósticos que utilizan la información observada reducen la incertidumbre en contraste con aquellos que utilizan suposiciones en los valores de los predictores. El análisis anterior tuvo los mismos resultados al observar las bandas de incertidumbre para los demás GCM (Anexo digital).

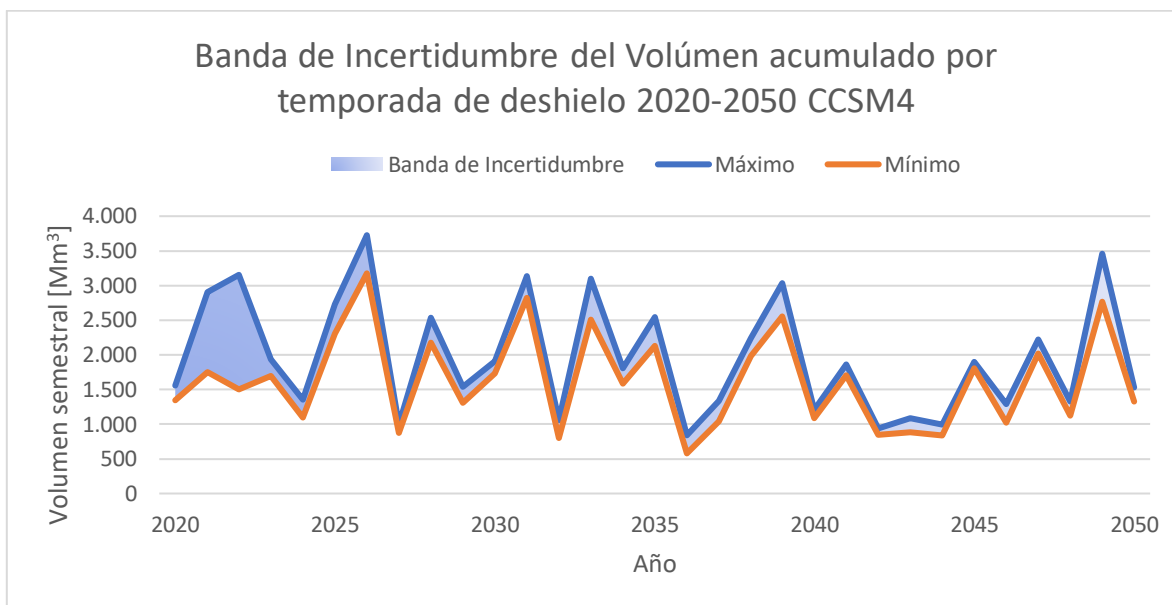


Figura 5.55: Banda de Incertidumbre de los Volúmenes acumulados en la temporada de deshielo 2020-2050

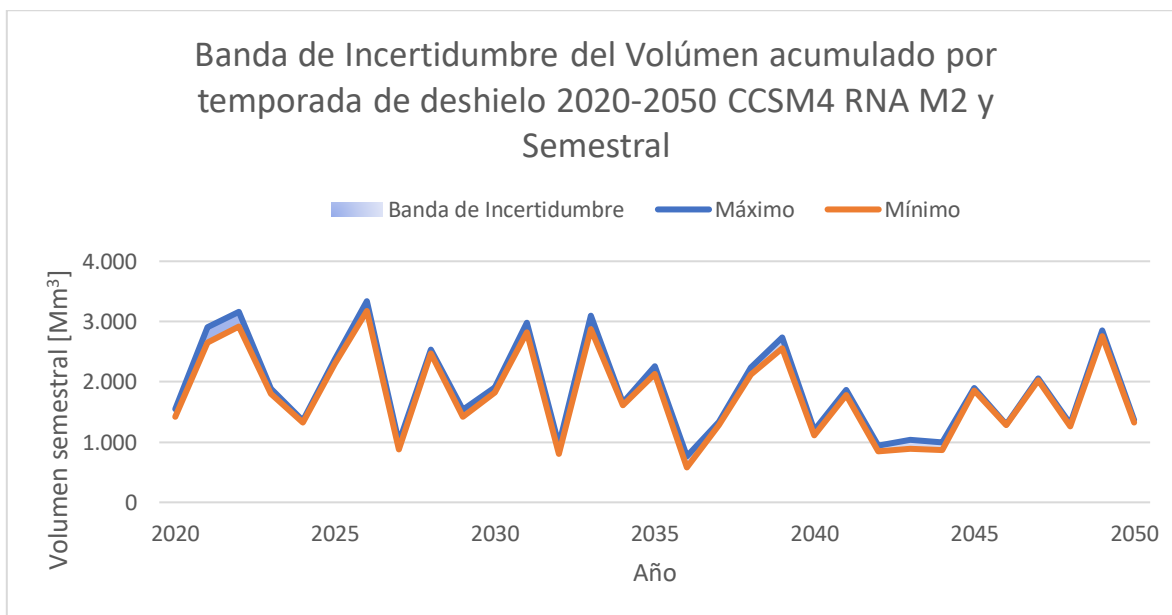


Figura 5.56: Banda de Incertidumbre de los Volúmenes acumulados en la temporada de deshielo 2020-2050 M2

6.-Análisis de los resultados

El análisis realizado a las variables disponibles para el pronóstico arrojó una tendencia a la disminución de los caudales de deshielo, lo que se debe principalmente al déficit de precipitaciones y equivalente de agua en nieve en el periodo de estudio y el hecho que la cuenca del río Maipo en el manzano tiene comportamiento nival, es decir, la principal fuente de caudales en deshielos es la precipitación y acumulación de nieve en invierno.

El análisis de las componentes principales aplicado mostró un patrón repetitivo para cada paso temporal (semanal, mensual y semestral) de variables que explican de mejor manera la varianza. Estas variables son Q, P, T, ET y SWE (solo caso semestral), las cuales explican en promedio un 97% de la varianza con las 4 primeras componentes. Del mismo modo, las funciones de autocorrelación y correlación cruzada lograron entregar una idea de los posibles predictores para cada RNA construida. Al mismo tiempo, se encontró una dependencia entre las estaciones de invierno y verano, debido a que en casi todas las variables las correlaciones más altas se encontraron con un retardo (lag) que indicaban mediciones de las variables en invierno. Sin embargo, las matrices de correlación entregaron la mejor idea de los predictores que se debían utilizar, esto se debe a que era posible buscar la correlación entre todas las semanas y meses del año (52 y 12 respectivamente) con las semanas y meses del periodo de deshielo (Anexo G). El análisis semestral solo indicó que las mejores correlaciones se encuentran con las variables del invierno anterior.

El hecho de haber entrenado múltiples conjuntos aleatorios de predictores (60 semanales, 300 mensuales y 60 semestrales) hizo que el análisis fuese más robusto, de manera que la elección de predictores no fuera totalmente deliberada. Esto provocó que se pudiera elegir la mejor combinación (métricas de evaluación y criterio de la información de Akaike) y que, por lo tanto, se redujera el error total del pronóstico de las temporadas de deshielo. El entrenamiento se realizó con el método Levenberg-Maquardt, lo que significó buenos resultados en la mayoría de las estructuras simuladas (NSE, KGE y NMSE), se obtuvo coeficientes de regresión (R) mayores a 0.9 e histograma de errores con aparente distribución normal. Adicionalmente a lo anterior se obtuvo un buen desempeño en la validación cruzada de cada uno de ellos (NSE, KGE y NMSE).

Los resultados se vieron potenciados por la existencia de 4 set de datos adicionales provenientes del periodo LB de los GCM. Con ellos fue posible encontrar la configuración de pesos y bias más adecuada.

El análisis de 2 modelos de RNA donde lo que cambiaba era la disponibilidad de la información y tiempo de adelanto, entregó valiosa información para calificar la calidad de los pronósticos. El objetivo de la creación de dos 2 modelos era ver el comportamiento de las RNA bajo bases que utiliza la DGA en sus pronósticos anuales en contraste con un modelo que no tuviera tales restricciones.

En general el modelo 2 (información perfecta) obtuvo los mejores resultados (i.e errores menores) tanto en entrenamiento como en pronóstico. Cabe destacar que solo en el pronóstico semanal del modelo 2 fue posible la elección del set HIST para entrenamiento.

Al indagar sobre qué semanas o meses son más complicados de pronosticar (magnitud del error), se encontró que en las RNA pertenecientes al modelo 1 no existe un comportamiento definido en la magnitud de los errores, es más, los errores medios de cada semana y/o mes varían levemente en torno a cierto nivel de error. Lo anterior podría tener como causa el haber tomado valores conocidos de variables (en particular SWE, Pp y T) y caudales pronosticados en tiempos anteriores. En cambio, en las RNA del modelo M2 se ve claramente un comportamiento común, donde el error medio disminuye a medida que avanza la temporada de deshielo hasta cierto punto donde luego comienza a aumentar para terminar con un error medio menor a la inicial.

Finalmente, sobre el análisis de incertidumbre, se nota una importante reducción de esta cuando solo se consideran las RNA del modelo 2 y la RNA semestral, esto ocurre para cada futuro pronosticado (CCSM4, CSIRO, IPSL y MIROC). Esto indicaría la superioridad del modelo 2 sobre el modelo 1 en cuanto a desempeño y distribución de los errores.

7.-Conclusiones

El análisis de las variables disponibles muestra una tendencia a la disminución de los caudales medios de deshielo, donde las posibles causas son la disminución de precipitaciones en invierno.

Si bien el análisis de componentes principales y las autocorrelaciones parciales mostró una dirección en la búsqueda de posibles predictores, los mejores indicadores resultaron ser las matrices de correlación, las que mostraron que tanto el caudal como el equivalente de agua en nieve y la precipitación fueron en general los mejores predictores.

El desempeño de un pronóstico utilizando RNA queda sujeto a un elemento muy importante: la estructura de red. Por lo tanto, la cantidad de información (N° de ejemplos y Variables) y la calidad de esta juega un rol muy importante.

Las estructuras óptimas se escogieron en base a los criterios mencionados anteriormente y en su desempeño en el pronóstico a nivel de temporada (magnitud del error). La existencia de múltiples escenarios con variables escogidas aleatoriamente (60 semanales, 300 mensuales y 60 semestrales) no solo hizo más robusto el análisis, sino que abrió un abanico de posibilidades (variables y parsimonia), lo que fue de gran utilidad debido a que se esperaba pronosticar con tiempo de adelanto una temporada (26 semanas, 6 meses o un semestre según el análisis temporal). Lo anterior dio pie para encontrar algunas otras estructuras con menos variables y más parsimoniosas con igual o mejor desempeño.

Realizar el pronóstico de caudales de deshielo con pasos de adelanto menores (modelo 2) a una temporada resultó en un mejor desempeño (del 10% al 20% de errores medios respectivamente) que si se pronostica con una temporada completa (26 semanas, 6 meses o un semestre según el análisis temporal) de adelanto (modelo 1). Si bien no se cumplió con la meta DGA en las RNA del modelo 1 los resultados no son desalentadores (aproximadamente 90% de las temporadas con error bajo 20%).

Esto se debe a que en las entradas de las RNA ingresa información “observada” y no datos pronosticados (Q) y/o valores conocidos (SWE, Pp y T), lo que le da mayor precisión al pronóstico. Lo anterior se ve potenciado al comparar las bandas de incertidumbre al considerar todas las RNA en contraste con las RNA del modelo 2, donde se demostró una importante disminución en la incertidumbre al considerar solo las redes del modelo 2 y la RNA semestral.

Dicho lo anterior queda a criterio de los usuarios la aplicación de distintos modelos (RNA M1 y M2) y pasos temporales (semanales, mensuales y semestrales) según los objetivos y disponibilidad de la información que se tenga. No obstante, se deja clara la ventaja del modelo 2 sobre el 1 en cuanto a desempeño general e incertidumbre.

Se verificó que los modelos de redes neuronales artificiales son capaces de pronosticar en periodos futuros en la cuenca de estudio. Se recomienda extender el análisis a otras cuencas nivales con el fin de tener una visión más general de la aplicación de las RNA en el futuro.

8.-Bibliografía

- (DGA), D. D. (2018). *PRONÓSTICO DE CAUDALES DE DESHIELO TEMPORADA DE RIEGO 2018-2019*.
- Akaike, H. (1974). *A New Look at the Statistical Model Identification*.IEEE *TRANSACTIONS ON AUTOMATIC CONTROL*, VOL. AC-19, KO. 6, DECEMBER 1974 .
- Artigas, M. (2015). PRONÓSTICO DE CAUDALES DE DESHIELO A DISTINTAS ESCALAS TEMPORALES, USANDO REDES NEURONALES ARTIFICIALES EN LA CUENCA DEL RÍO MAPOCHO EN LOS ALMENDROS. *Memoria Para Optar al Título de Ingeniero Civil, Universidad de Chile*.
- Barría, P. (2010). PRONÓSTICO DE CAUDALES MEDIOS MENSUALES EN LAS CUENCAS DE LOS RIOS BAKER Y PASCUA. *Memoria Para Optar al Título de Ingeniero Civil, Universidad de Chile*.
- Brynjolfsson, E. (2017). What can machine learning do? Workforce implications: Profound change is coming, but roles for humans remain. *Science*, 22 December 2017, Vol.358(6370), pp.1530-1534.
- CADE-IDEPE Consultores en Ingeniería. (2004). Cuenca Del Río Maipo. *Diagnostico y Clasificación de los Cursos y Cuerpos de Agua Según Objetos de Calidad*.
- Center for Climate and Resilience Research (CR)2. (2015). Informe a la Nación: La megasequía 2010-2015: Una lección para el futuro. *Researchgate*.
- Court, S. (2008). PRONÓSTICO DE CAUDALES DE LA CUENCAS DE BETANIA Y BIO-BIO UTILIZANDO MÉTODOS ESTADÍSTICOS. *Memoria Para Optar al Título de Ingeniero Civil Matemático, Universidad de Chile*.
- DGA. (2017). Actualización del Balance Hídrico Nacional, SIT N°417, Ministerio de Obras Públicas, Dirección General de Aguas, División de Estudios y Planificación, Santiago, Chile, Realizado por: Universidad de Chile & Pontificia Universidad Católica de Chile.
- Easey et.al, J. (2006). *Seasonal forecasting of river flows: a review of the state-of-the-art*. Climate variability and change: hydrological impacts. Proceedings of the fifth FRIEND World Conference, Havana, Cuba, November 2006. IAHS, 158-162.
- Espinoza, A. (1973). *Pronóstico de crecidas en tiempo real. Uso de Redes Neuronales*. Tesis para optar al grado de Magíster en Ciencias de la Ingeniería Mención Recursos Hídricos y Medio Ambiente. *Universidad de Chile*.
- FAO. (2006). *Evapotranspiración del Cultivo: Guías para la determinación de los requerimientos de agua en los cultivos*.
- G, S. D. (2005). Aplicación de Redes Neuronales MLP a la predicción de un paso en series de tiempo. *Facultad de Ingeniería en Sistemas, Fundación Universitaria Konrad Lorenz, Bogota*.

- Gao et.al. (2009). *Water Budget Record from Variable Infiltration Capacity (VIC) Model Algorithm Theoretical Basis Document*.
- Garreaud et.al, R. (2017). *The 2010–2015 megadrought in central Chile: impacts on regional hydroclimate and vegetation*.
- Gomez, P. (1999). *Notas del Curso Fundamentos de Redes Neuronales. Universidad de las Americas. Puebla*.
- Gonzalez, R. (2003). PRONÓSTICO DE CRECIDAS EN TIEMPO REAL: USO DE REDES NEURONALES. *Tesis Para Optar al Título de Magíster en Ciencias De la Ingeniería Mencion Recursos Hídricos y Medio Ambiente, Universidad de Chile*.
- Hudson, M. (2015). *Neural Network Toolbox User Guide. Matlab*.
- Iniciativa Científica Milenio. (2017). *El Agua Como Recurso Vital: Bases Científicas para la legislación del recurso Hídrico en Chile*.
- IPCC. (1990). *Working Group I: Scientific Assesment of Climate Change*.
- IPCC. (2007). *CLIMATE CHANGE 2007: The Physical Science Basis*.
- IPCC. (2013). *CLIMATE CHANGE 2013: The Physical Science Basis*.
- Jolliffe et.al. (2016). *Principal component analysis: a review and recent developments. Philos Trans A Math Phys Eng Sci. 2016 Apr 13; 374(2065): 20150202*.
- Kling, & Gupta. (s.f.). *Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology, 2009, Vol.377(1), pp.80-91*.
- Lagos, M. (2012). IMPACTO DEL CAMBIO CLIMÁTICO EN EVENTOS EXTREMOS Y ANÁLISIS DE LA VULNERABILIDAD DE ALGUNAS OBRAS HIDRÁULICAS EN CHILE. *Tesis para optar al grado de Magíster en Ciencias de la Ingeniería Mencion Recursos y Medioambiente Hídrico, Universidad de Chile*.
- Lagos, M. (2013). *Diseño Hidrológico En Cuencas de Régimen Mixto, Incertidumbre Y Cambio Climático: El Caso Del Embalse Puclaro*.
- Larson, S. (1931). *The shrinkage of the coefficient of multiple correlation. J. Educat. Psychol., 22:45–55, 1931*.
- Liang et.al. (1994). *A Simple hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs. Journal of Geophysical Research Atmospheres 99(14)*.
- Linnainmaa, S. (1970). *The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors. Master's Thesis (in Finnish), Univ. Helsinki, 6–7*.
- Mendoza, P. (2010). PRONÓSTICO DE CAUDALES DE CRECIDA MEDIANTE UN MODELO HIDROLÓGICO DISTRIBUIDO Y LA ASIMILACIÓN DE DATOS OBSERVADOS. *Memoria Para Optar al Título de Ingeniero Civil, Universidad de Chile*.

- MOP. (2003). Evaluación de los Recursos Hídricos en la Cuenca del río Maipo.
- Muñoz, A. (1996). *Aplicación de Tecnicas de Redes Neuronales al Diagnostico de Procesos Industriales. Tesis Doctoral Escuela Tecnica Superior de Ingenieros Industriales. Universidad Pontificia Comilla, Madrid, España.*
- Muñoz, M. (2008). PRONÓSTICO DE CRECIDAS EN TIEMPO REAL USANDO REDES NEURONALES RECURRENTE. *Memoria Para Optar al Título de Ingeniero Civil, Universidad de Chile.*
- Nash, & Sutcliffe. (1970). *RIVER FLOW FORECASTING THROUGH CONCEPTUAL MODELS. Journal of Hydrology 10 (1970) 282-290.*
- Papert, & Minsky. (1972). *Perceptrons: an introduction to computational geometry.* The MIT Press; Expanded.
- Prakash, O., Sudheer, K., & Srinivasan, K. (2014). IMPROVED HIGHER LEAD TIME RIVER FLOW FORECAST USING SEQUENTIAL NEURAL NETWORK WITH ERROR UPDATING. *J. Hydrol. Hydromech, 60-74.*
- Purves, D. (2001). *Neuroscience.* Panamericana.
- Stone, M. (1974). *Cross-Validatory Choice and Assessment of Statistical Predictions. Journal of the Royal Statistical Society. Series B (Methodological).*
- Tapia, C. (2012). ADAPTACIÓN AL CAMBIO CLIMÁTICO DE LA GESTIÓN HÍDRICA PARA EL SECTOR RIEGO EN LA TERCERA SECCIÓN DEL RÍO MAIPO. *TÉISIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN Y POLÍTICAS PÚBLICAS, Universidad de Chile.*
- Troncoso, M. (1999). *Uso de redes neuronales para el pronóstico de crecidas en cuencas pluvionivales. Memoria para optar al título de Ingeniero Civil, Universidad de Chile.*
- Videla, Y. (2013). MODELACIÓN HIDROLÓGICA DE LA CUENCA ALTA DEL RÍO MAIPO MEDIANTE LA APLICACIÓN DEL SOFTWARE "COLD REGION HYDROLOGICAL MODEL". *Memoria Para Optar al título de Geógrafo, Universidad de Chile.*
- Wedgbrow et.al. (2002). *Prospects for seasonal forecasting of summer drought and low river flow anomalies in England and Wales. Int. J. Climatol. 22, 219-236.*
- Widrow, & Hoff. (1960). *Adaptive Switching Circuits. Stanford Electronics Laboratories. Stanford University.*
- Wood et.al. (2002). *Long-range experimental hydrologic forecasting for the eastern United States. JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 107, NO. D20, 4429.*

9.-Anexos

Anexo A. Análisis de las forzantes y Variables VIC

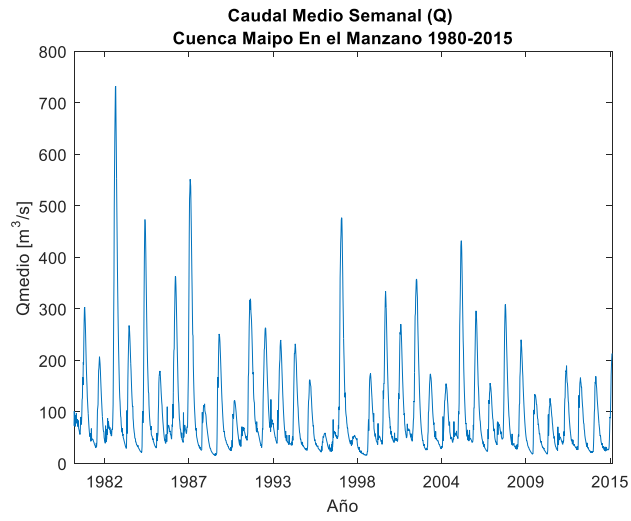


Figura 9.1: Caudal Semanal Cuenca Maipo en el Manzano 1980-2015

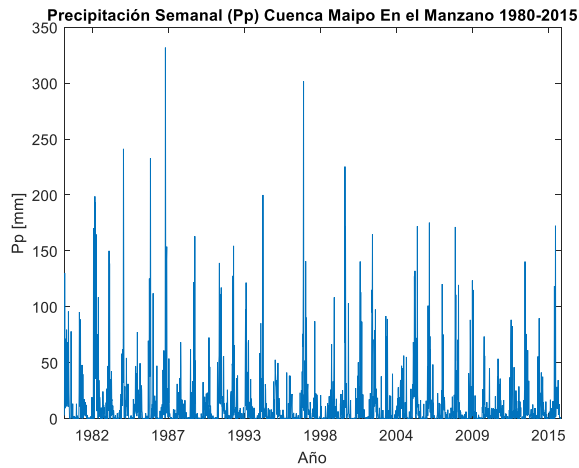


Figura 9.2: Precipitación Semanal Maipo en el Manzano 1980-2015

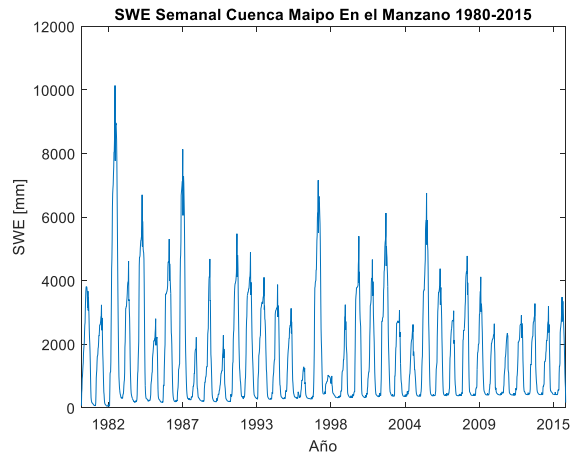


Figura 9.3: SWE Semanal Cuenca Maipo en el Manzano 1980-2015

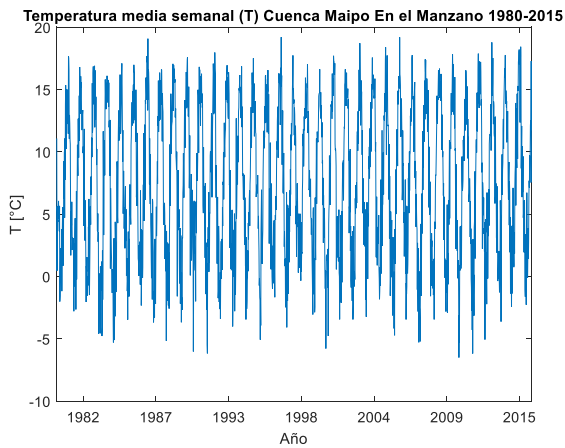


Figura 9.4: Temperatura Media Semanal Maipo en el Manzano 1980-2015

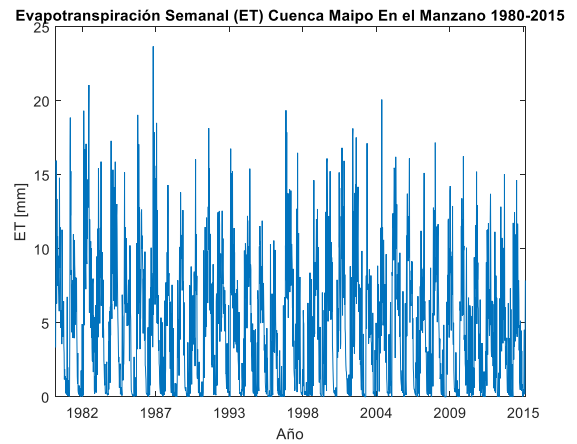


Figura 9.5: Evaporación Semanal Maipo en el Manzano 1980-2015

Tabla 9.1: Estadísticos Variables periodo de deshielo nivel Semanal

	Q[m ³ /s]	SWE [mm]	Pp [mm]	T [°C]	Q [mm]
Promedio	150,0	1154,9	3,1	12,2	1,5
Std	107,9	1486,0	6,5	3,7	1,4
Max	732,1	8990,1	55,1	19,2	8,0
Min	22,4	56,6	0,0	-0,2	0,0

Tabla 9.2: PAC semanal

	PC1	PC2	PC3	PC4	PC5
Q	0,48	0,41	-0,09	0,63	-0,45
SWE	-0,45	0,34	-0,65	0,31	0,41
Pp	-0,39	0,35	0,75	0,33	0,22
T	0,62	-0,05	0,08	0,11	0,77
E	0,16	0,77	-0,01	-0,62	0,01

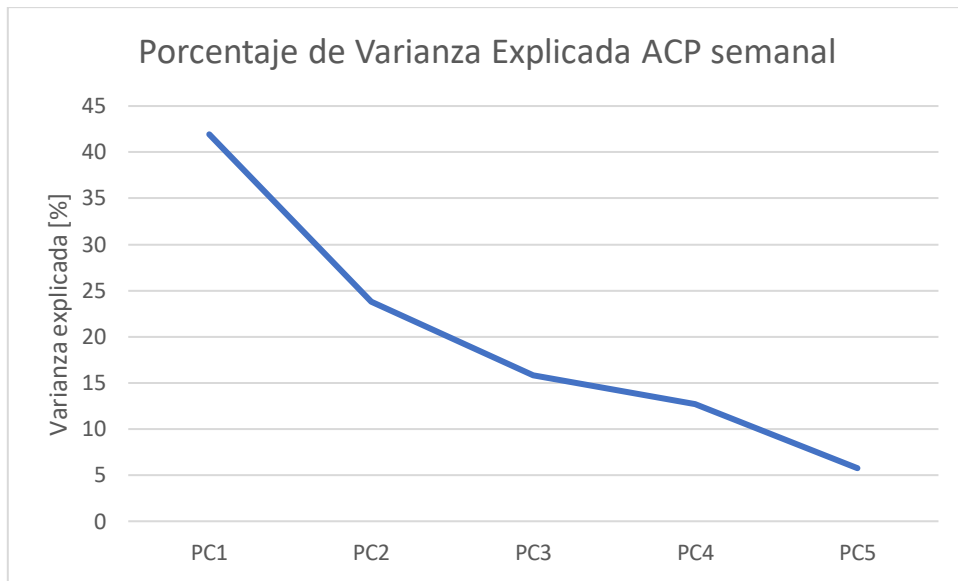


Figura 9.6: Porcentaje de Varianza explicada por cada componente principal nivel semanal

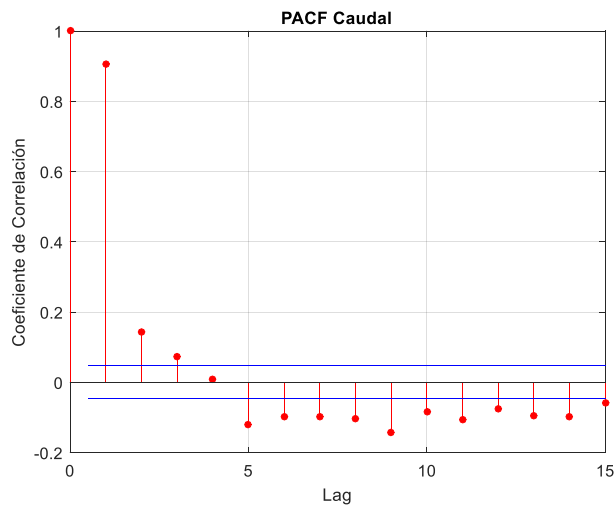


Figura 9.7: PACF Caudal Medio Semanal

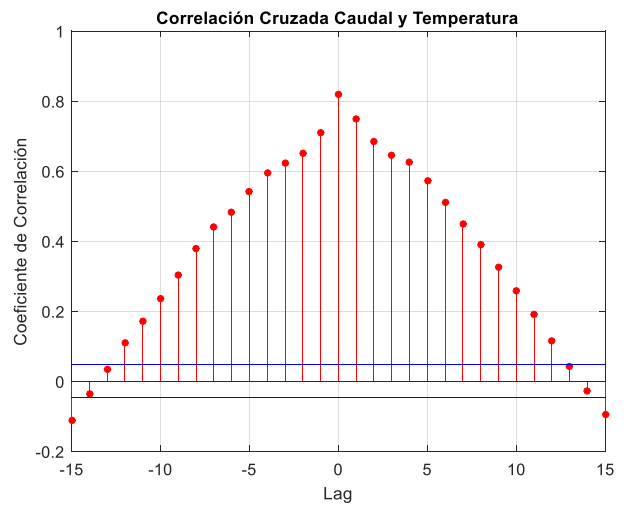


Figura 9.8: Correlacion Caudal-Temperatura Semanal

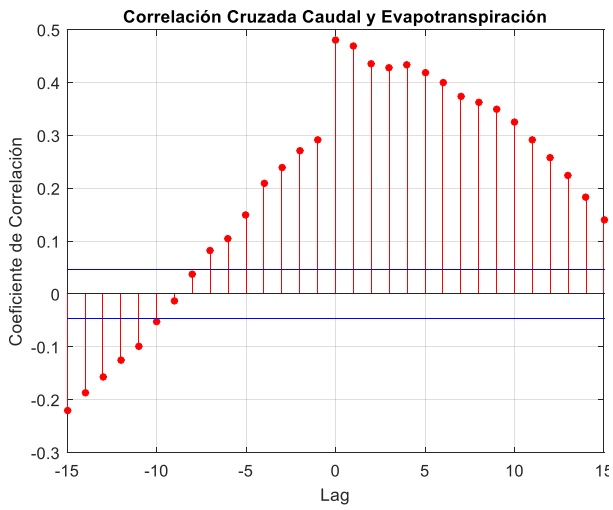


Figura 9.9: Correlación cruzada caudal- evaporación nivel semanal

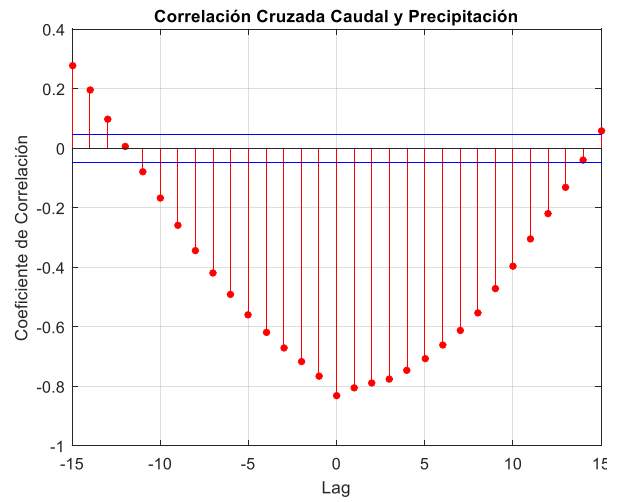


Figura 9.10: Correlación caudal-precipitación nivel semanal

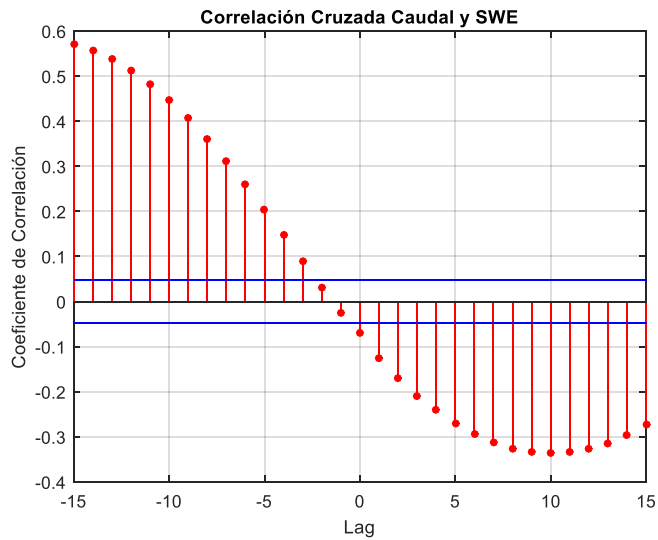


Figura 9.11: Correlación caudal-swe nivel semanal

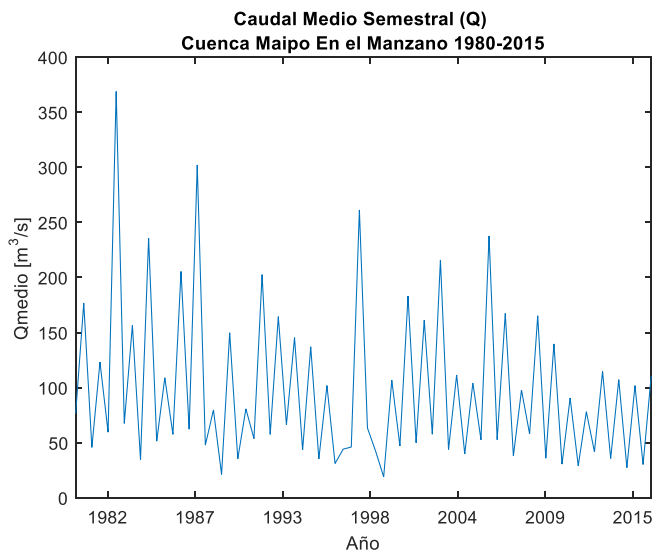


Figura 9.12: Caudales Medios Semestrales

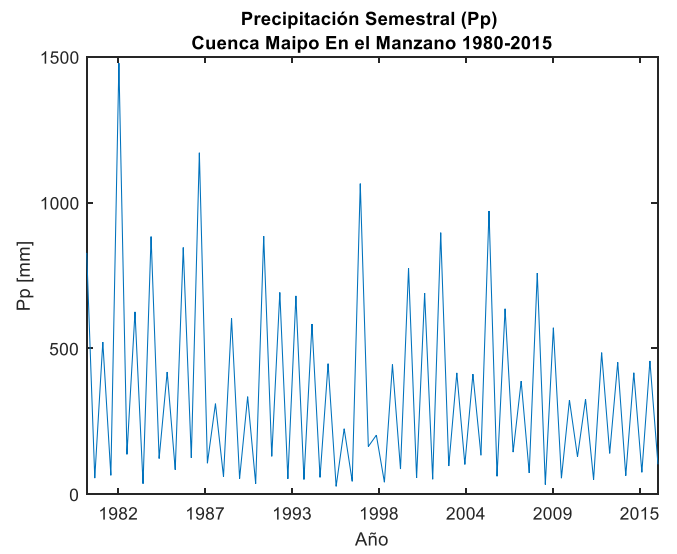


Figura 9.13: Precipitación Semestral

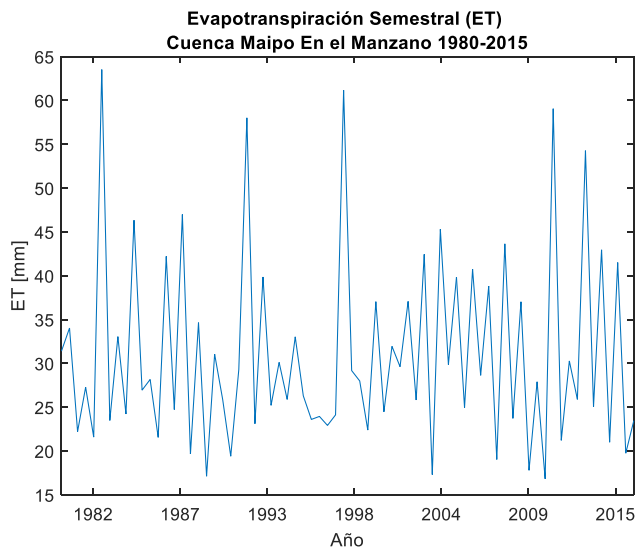


Figura 9.14: Evaporación Semestral

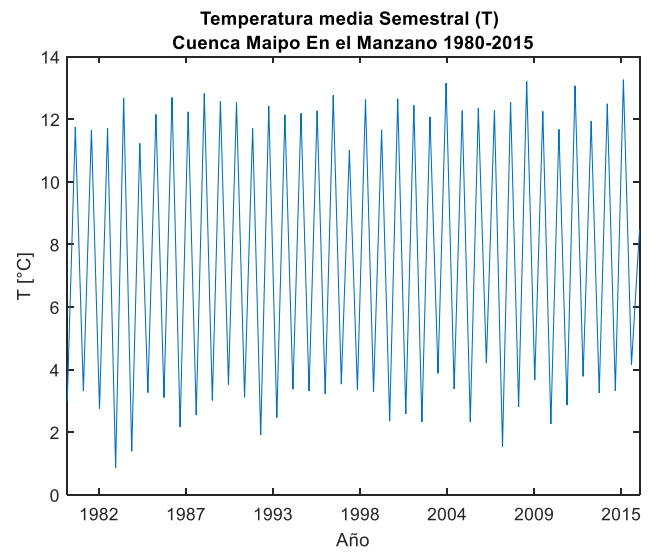


Figura 9.15: Temperatura Media semestral

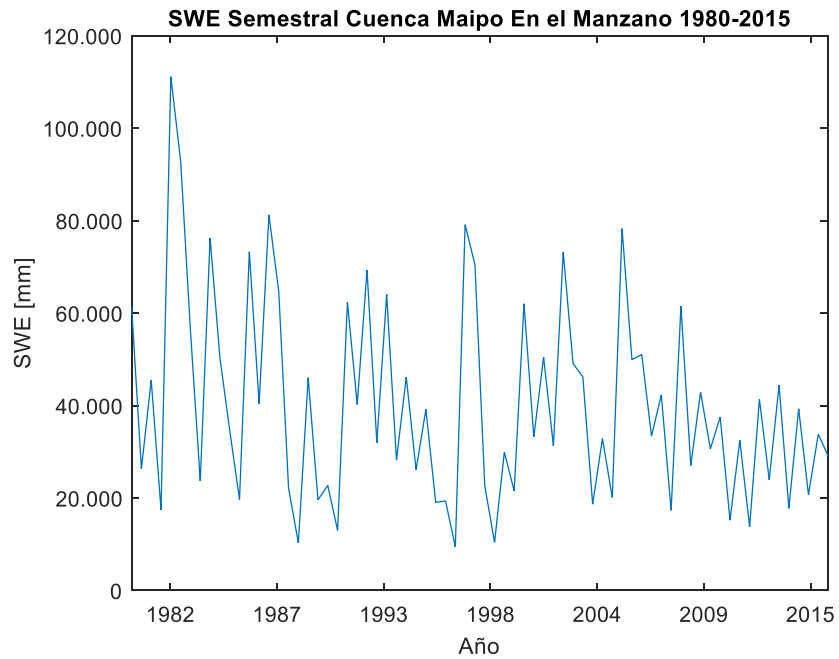


Figura 9.16: SWE medio semestral

Tabla 9.3: Estadísticos Variable Semestral

	Q[m ³ /s]	SWE [mm]	Pp [mm]	T [°C]	Q [mm]
Promedio	135,1	14234,9	85,9	12,0	41,0
Std	41,3	6918,8	44,9	0,6	11,7
Max	241,5	36485,1	191,1	13,2	69,0
Min	77,0	5447,1	25,1	10,8	20,1

Tabla 9.4: PCA Semestral

	PC1	PC2	PC3	PC4	PC5
Q	0,43	0,49	-0,48	-0,30	0,50
SWE	-0,31	0,69	-0,24	0,06	-0,60
Pp	-0,51	0,31	0,18	0,50	0,60
T	0,53	-0,03	-0,21	0,81	-0,14
E	0,43	0,42	0,79	-0,06	-0,05

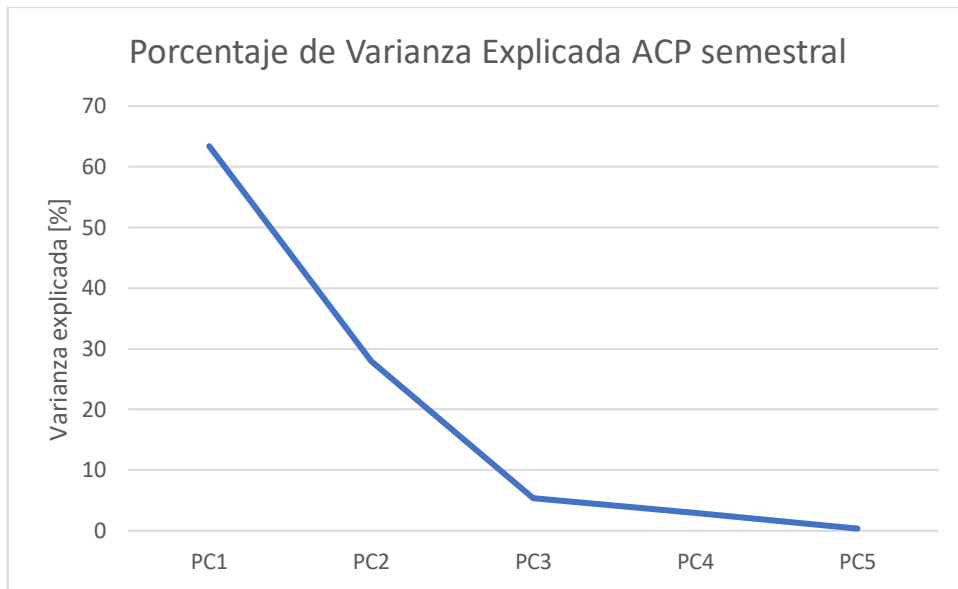


Figura 9.17: Porcentaje de Varianza explicada por cada componente principal nivel semestral

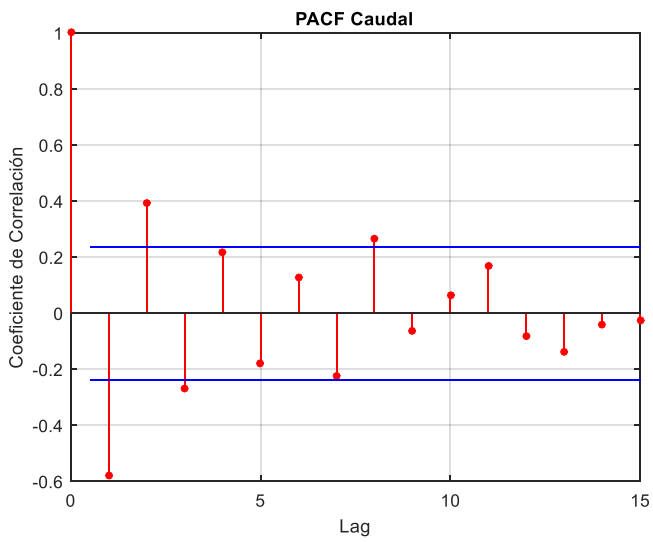


Figura 9.18: PACF Q medio semestral

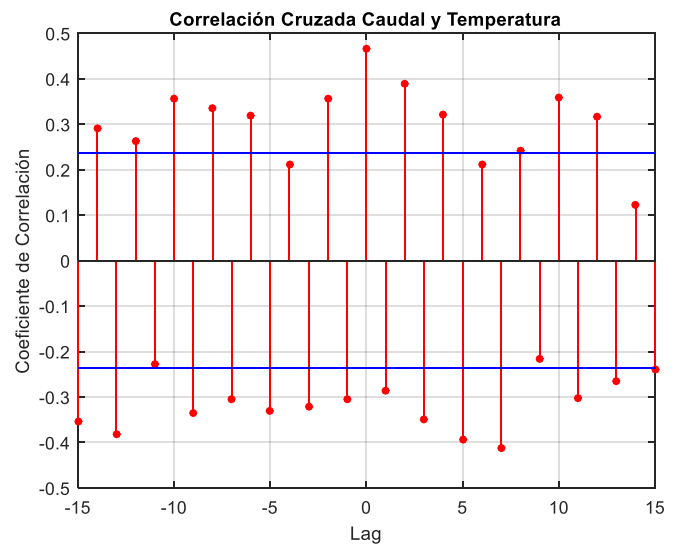


Figura 9.19: Correlación Q-T Semestral

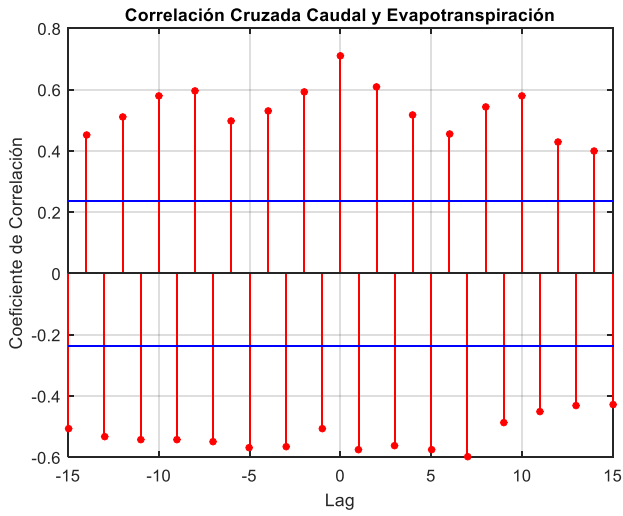


Figura 9.20: Correlación Q-E semestral

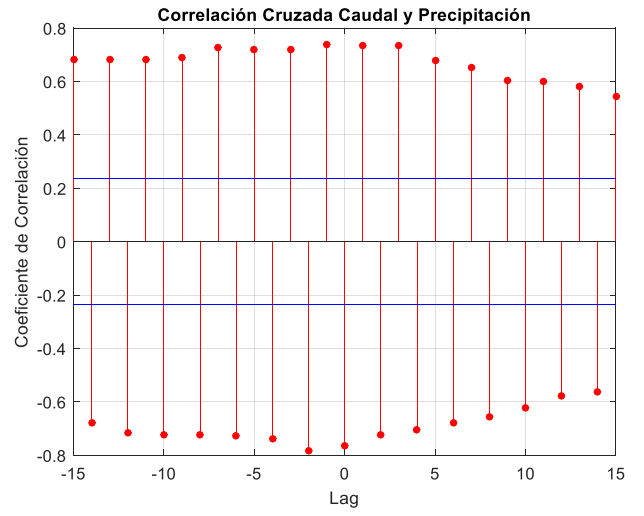


Figura 9.21: Correlación Q-P semestral

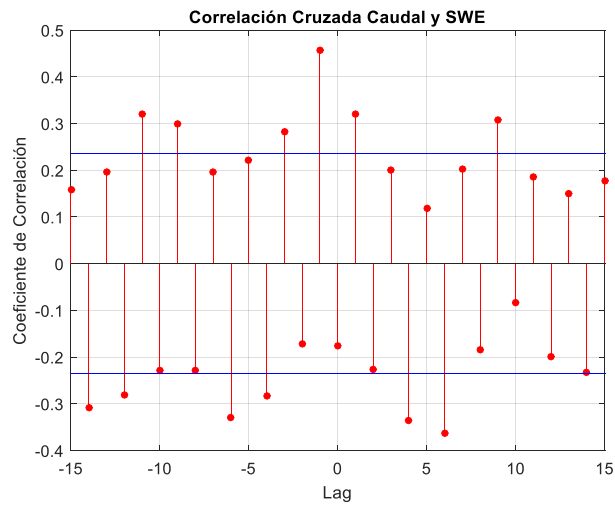


Figura 9.22: Correlación Q-SWE semestral

Anexo B. Ranking de los conjuntos de variables aleatorias

Tabla 9.5: Conjuntos de predictores RNA semanal

Conjunto	AIC	AICc	CALIBRACION			VALIDACIÓN CRUZADA		
			NSE	KGE	NMSE	NSE	KGE	NMSE
25	7971	8453	90%	92%	5%	91%	93%	5%
54	7987	8477	93%	94%	4%	90%	93%	5%
11	8006	8576	53%	71%	25%	55%	73%	22%
57	8012	8511	91%	94%	4%	91%	93%	5%
30	8042	8594	75%	85%	12%	74%	83%	13%
45	8053	8575	92%	94%	4%	90%	93%	5%
13	8062	8584	96%	97%	2%	91%	95%	5%
43	8063	8584	96%	98%	2%	95%	97%	2%
15	8065	8587	98%	98%	1%	97%	98%	1%
5	8067	8603	93%	94%	4%	93%	95%	4%
50	8069	8604	91%	94%	4%	90%	93%	5%
41	8070	8606	89%	92%	6%	87%	91%	7%
38	8072	8607	93%	94%	4%	92%	94%	4%
29	8073	8625	93%	92%	4%	86%	93%	7%
3	8083	8618	95%	97%	3%	94%	96%	-10%
40	8084	8620	96%	97%	2%	96%	97%	2%
16	8085	8621	97%	98%	1%	94%	97%	2%
59	8087	8623	97%	98%	1%	94%	97%	3%
28	8092	8643	95%	96%	2%	95%	97%	2%
44	8096	8647	91%	93%	5%	89%	91%	5%
12	8102	8693	74%	86%	14%	75%	85%	20%
27	8104	8674	89%	91%	6%	86%	88%	7%
39	8107	8658	97%	98%	1%	94%	97%	3%
31	8107	8659	97%	99%	1%	97%	98%	1%
46	8116	8686	92%	94%	4%	90%	92%	5%
4	8118	8708	85%	90%	8%	83%	87%	7%
7	8118	8709	92%	94%	4%	86%	92%	0%
18	8121	8691	97%	98%	1%	96%	97%	2%
37	8122	8692	94%	97%	3%	92%	95%	4%

Tabla 9.6: Conjunto de predictores RNA mensual

Conjunto	Aic	AICc	CALIBRACION			VALIDACIÓN CRUZADA		
			NSE	KGE	NMSE	NSE	KGE	NMSE
97	1656	1770	94%	96%	3%	90%	92%	5%
237	1653	1769	94%	92%	3%	88%	90%	5%
194	1657	1772	94%	95%	3%	90%	91%	5%
295	1658	1773	94%	96%	3%	88%	90%	6%
66	1662	1780	94%	95%	3%	85%	86%	7%
166	1656	1773	94%	93%	3%	91%	92%	4%
176	1656	1770	94%	96%	3%	88%	90%	6%
296	1670	1792	94%	96%	3%	88%	91%	6%
136	1662	1779	94%	95%	3%	88%	89%	6%
140	1659	1777	94%	94%	3%	88%	89%	6%
236	1666	1785	94%	96%	3%	86%	89%	6%
217	1669	1789	94%	96%	3%	89%	90%	5%
228	1660	1777	94%	94%	3%	86%	88%	6%
243	1660	1776	94%	96%	3%	87%	88%	6%
59	1656	1770	94%	96%	3%	88%	90%	5%
284	1666	1787	94%	95%	3%	88%	89%	6%
253	1658	1773	94%	97%	3%	89%	92%	5%
108	1673	1796	94%	95%	3%	88%	89%	6%
96	1663	1780	94%	96%	3%	89%	91%	5%
39	1667	1786	94%	96%	3%	89%	90%	5%
84	1664	1783	94%	96%	3%	89%	91%	5%
268	1657	1773	94%	95%	3%	88%	90%	5%
12	1666	1786	94%	95%	3%	87%	88%	6%
53	1648	1759	94%	96%	3%	91%	92%	4%
259	1668	1789	94%	96%	3%	85%	87%	7%

Tabla 9.7: Conjuntos de predictores RNA semestral

Conjunto	AIC	AICc	CALIBRACION			VALIDACIÓN CRUZADA		
			NSE	KGE	NMSE	NSE	KGE	NMSE
52	241,3	272,5	49%	66%	10%	95%	93%	1%
10	241,8	269,5	94%	93%	1%	95%	93%	1%
53	241,9	284,4	31%	38%	14%	84%	82%	3%
8	242,6	278,6	42%	59%	12%	25%	50%	15%
31	242,6	273,8	9%	54%	17%	15%	34%	17%
42	242,7	270,4	99%	99%	0%	98%	99%	0%
27	242,8	274,0	23%	58%	15%	15%	46%	16%
45	243,6	279,6	24%	57%	16%	-2%	29%	21%
5	243,9	326,3	33%	37%	14%	69%	74%	6%
38	243,9	286,4	50%	72%	10%	89%	90%	2%
47	244,3	275,5	98%	96%	0%	93%	94%	1%
49	244,4	275,6	93%	94%	1%	88%	89%	2%
7	244,5	280,5	93%	89%	2%	89%	91%	2%
19	244,6	275,8	98%	98%	0%	95%	97%	1%
37	244,6	275,8	98%	99%	0%	90%	92%	2%
18	244,7	275,9	96%	94%	1%	94%	91%	1%
1	244,9	276,1	99%	99%	0%	92%	94%	2%
33	244,9	276,1	99%	99%	0%	84%	88%	3%
58	245,0	276,2	99%	99%	0%	98%	98%	0%

Anexo C. Valores conocidos de variables para el pronóstico

Tabla 9.8: Valores conocidos Pp pronóstico mensual

Pexcd [%]	Pp Total invierno[mm]	Octubre	Noviembre	Diciembre	Enero	Febrero	Marzo
>10	1086,4	21,2	28,9	8,1	7,9	0,8	16,2
20	834,9	35,9	9,2	2,9	14,2	14,1	12,5
30	774,9	58,0	28,7	1,1	1,9	18,8	3,0
40	655,4	58,5	22,1	4,3	4,0	6,1	4,6
50	567,3	19,7	3,7	2,9	6,3	5,2	34,5
60	486,8	22,2	11,6	4,3	6,2	1,1	1,1
70	468,0	13,9	7,8	25,2	1,4	4,3	16,1
80	398,2	22,6	6,7	7,1	6,6	20,7	8,6
90	348,3	25,0	27,3	25,5	10,8	6,5	19,4
100	280,0	22,8	18,4	0,4	5,3	7,6	3,0

Tabla 9.9: Valores Conocidos SWE pronóstico Mensual

Pexcd [%]	SWE Total Invierno [mm]	Octubre	Noviembre	Diciembre	Enero	Febrero	Marzo
10	56370	18825	10315	2583	423	93	55
20	41581	12326	5196	883	140	51	39
30	37116	10603	4293	586	30	6	2
40	33398	10115	4081	711	56	8	0
50	28276	8423	3746	493	66	23	22
60	25420	6887	1895	157	27	8	2
70	22539	6089	1415	146	1	0	0
80	19405	6116	1831	433	69	18	11
90	15724	3492	816	165	10	0	2
100	11236	2668	1081	131	3	0	0

Tabla 9.10: Temperaturas medias Históricas para el pronóstico mensual.

Temperaturas Medias Históricas [°C]					
Octubre	Noviembre	Diciembre	Enero	Febrero	Marzo
6,7	10,3	13,1	15,2	14,5	12,9

Anexo D. Resultados del pronóstico

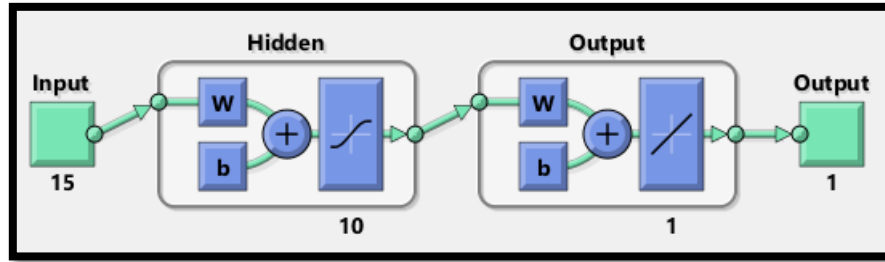


Figura 9.23: Estructura RNA Semanal modelo 1

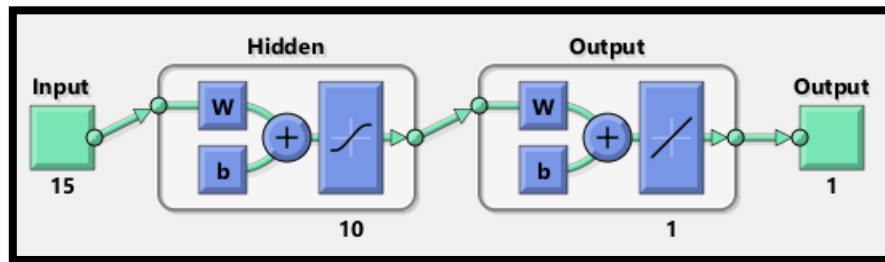


Figura 9.24: Estructura RNA Semanal modelo 2

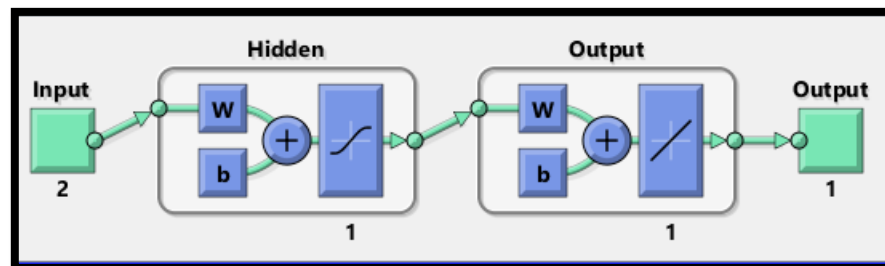


Figura 9.25: Estructura RNA mensual modelo 1

Tabla 9.11: Pesos Capa De Entrada RNA Mensual modelo 1

	Q(t-2)	SWE(t-2)
Neurona 1	-0,82	-1,31

Tabla 9.12: Pesos y Bias Capa Oculta modelo 1

	Peso Capa Oculta	Bias Capa Entrada	Bias Capa Oculta
Neurona 1	-0,671	-0,922	-0,380

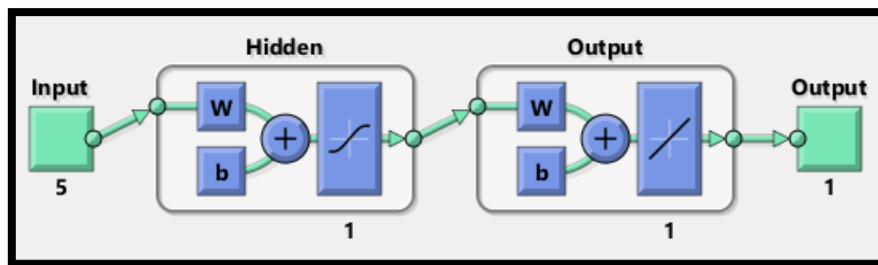


Figura 9.26: Estructura RNA mensual modelo 2

Tabla 9.13: Pesos Capa De Entrada RNA Mensual modelo 2

Variable	Q		SWE	Pp	
Lag	t-1	t-2	t-2	t-4	t-5
Neurona 1	-0,6099	0,15478	-0,337	-0,049	-0,0628

Tabla 9.14: Pesos de la capa de entrada RNA semanal. Modelo 1

Variable	Q	SWE					Pp					T			
		t-2	t-12	t-13	t-14	t-15	t-16	t-22	t-23	t-24	t-25	t-26	t-1	t-3	t-24
Neurona 1	0,46	1,92	-0,65	0,42	-0,09	-1,07	0,58	-0,15	0,35	-0,04	1,36	0,15	-0,87	0,32	-0,40
Neurona 2	-5,05	1,33	-0,82	3,53	1,42	-2,16	-1,73	0,66	0,93	-0,35	0,47	-2,94	1,15	1,21	0,11
Neurona 3	-12,25	2,59	-4,07	-4,03	-2,15	3,65	4,27	4,55	1,75	1,02	-1,37	6,97	5,71	1,13	-4,53
Neurona 4	2,16	5,89	-3,26	-0,30	1,57	-2,91	-1,67	-1,77	2,38	-2,95	2,65	-0,37	1,03	-0,23	-2,23
Neurona 5	7,64	-2,44	11,62	-8,84	-1,72	-1,85	-1,40	-0,45	-5,12	-1,37	2,66	-3,14	-0,86	1,39	0,95
Neurona 6	1,73	-4,63	6,16	-0,42	-3,03	0,30	-0,47	0,57	1,29	-1,45	-0,64	0,28	0,03	-0,63	0,23
Neurona 7	3,61	-0,56	8,19	-0,35	-1,02	1,45	1,49	1,72	1,16	-2,50	-1,35	-1,82	8,89	0,69	-7,17
Neurona 8	-1,24	2,90	-3,95	-0,80	3,35	1,79	-0,22	-0,09	1,69	0,83	-0,63	0,80	1,03	-0,47	0,43
Neurona 9	0,73	-1,22	0,03	0,20	0,09	0,05	-0,03	0,21	0,05	0,01	0,10	0,77	0,38	-0,21	-0,08
Neurona 10	0,88	-0,74	-0,05	0,06	0,22	0,14	-0,04	0,15	-0,01	0,08	-0,01	0,51	0,30	-0,14	0,01

Tabla 9.15: Pesos y biases neuronas capa oculta y capa de salida. RNA semanal. Modelo 1

Neurona	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	Bias Capa de Salida
Pesos	0,22	-0,67	-0,06	0,04	-0,05	0,23	0,02	-0,09	-1,34	1,87	
Bias	1,67	3,66	1,79	1,31	-2,05	-2,83	-1,70	2,66	0,13	0,56	0,19

Tabla 9.16: Pesos de la capa de entrada RNA semanal modelo 2

Variable	Q		SWE					Pp					T		
	t-2	t-4	t-12	t-13	t-14	t-15	t-16	t-22	t-23	t-24	t-25	t-26	t-3	t-24	t-25
Neurona 1	9,49	-11,24	-9,54	3,86	1,45	-0,60	9,78	-0,21	2,51	-4,80	5,13	-2,31	2,63	-0,58	-1,07
Neurona 2	-1,77	2,92	3,00	-3,95	-0,79	0,04	1,48	-0,74	-1,73	0,49	2,83	0,30	-0,15	3,81	-0,41
Neurona 3	-7,16	2,11	6,38	2,56	-1,43	-4,93	0,49	0,74	-1,66	2,34	1,96	1,20	1,94	-1,68	2,02
Neurona 4	1,10	-0,75	-0,17	-0,61	0,07	0,49	-0,45	-0,03	0,18	0,03	-0,12	0,15	0,02	-0,09	-0,28
Neurona 5	-3,01	1,98	0,08	-4,22	4,61	3,15	-2,55	-0,87	-0,90	-0,36	-0,46	0,01	-1,06	0,94	1,01
Neurona 6	-0,60	-2,35	-2,59	-0,89	-0,37	1,42	3,38	1,70	2,83	-0,53	-2,12	0,96	1,83	-0,37	0,46
Neurona 7	-10,08	3,40	-5,38	7,75	-5,43	-7,62	2,95	2,66	0,96	-0,61	3,27	0,20	-2,46	3,61	-2,02
Neurona 8	0,66	-0,25	0,14	0,44	0,03	-0,30	0,28	0,07	-0,04	-0,04	0,00	-0,06	-0,04	-0,07	0,13
Neurona 9	-5,12	8,66	-3,20	-4,74	3,17	0,04	-6,96	5,41	9,56	-6,16	-8,63	5,23	-1,16	3,81	-4,46
Neurona 10	-7,31	0,64	-7,87	-4,08	-5,04	8,09	-4,91	1,41	2,36	-7,35	-1,22	3,47	6,64	-1,32	-5,98

Tabla 9.17: Pesos y biases neuronas capa oculta y capa de salida modelo 2

	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	Bias Capa Salida
Pesos	-0,012	0,057	-0,013	2,545	0,015	0,023	-0,029	1,095	-0,010	-0,054	
Bias	3,799	1,579	1,614	1,517	-1,355	-1,460	-4,569	0,076	1,767	0,930	-2,413

Tabla 9.18: Pesos y Bias Capa Oculta modelo 2

	Peso Capa oculta	Bias Capa de entrada	Bias Capa oculta
Neurona 1	-1,2437	-0,193	-0,1996

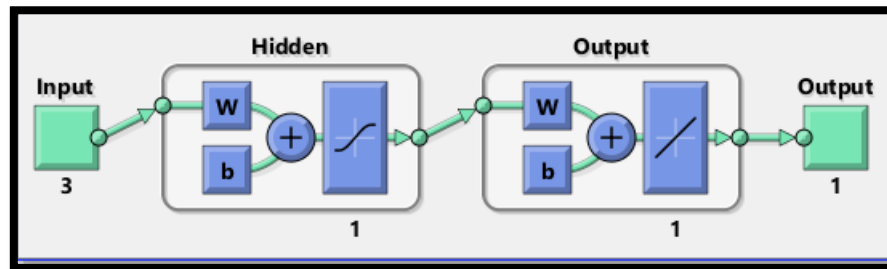


Figura 9.27: Estructura RNA Semestral

Tabla 9.19: Pesos Capa de Entrada calibración semestral

	Q(t-1)	SWE(t-1)	P(t-1)
Neurona 1	0,09	-0,14	-0,51

Tabla 9.20: Pesos Capa oculta y salida calibración semestral

	Pesos	Bias Capa oculta	Bias Capa de salida
Neurona 1	-1,83	-0,05	-0,07

Anexo E. Desempeño del Pronóstico

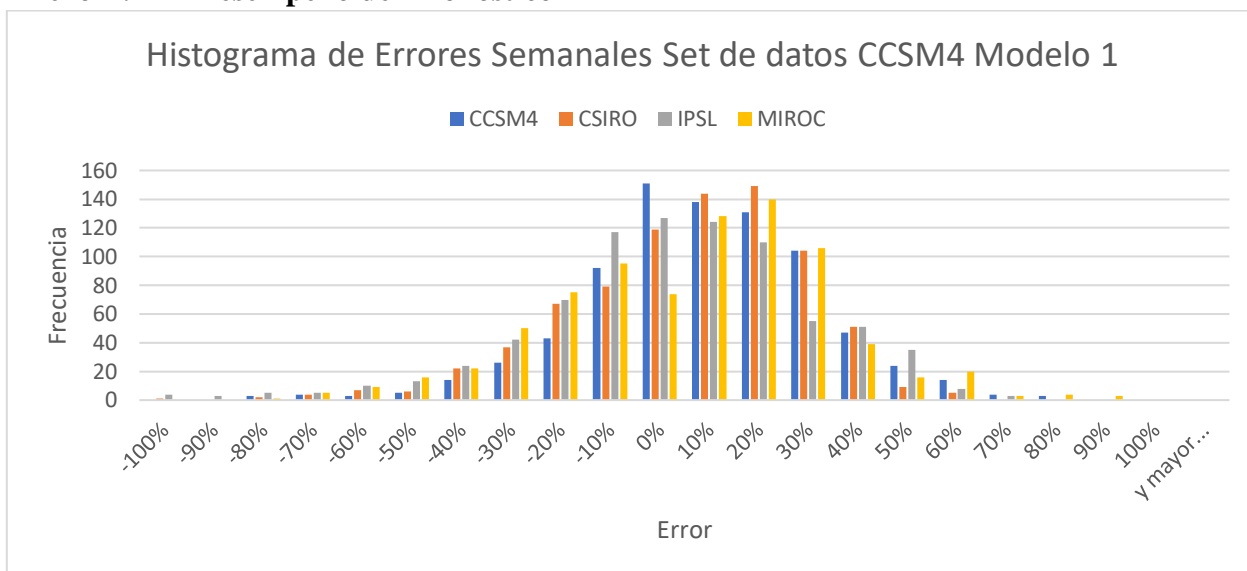


Figura 9.28: Histograma de errores semanales modelo 1. RNA calibrada con GCM CCSM4

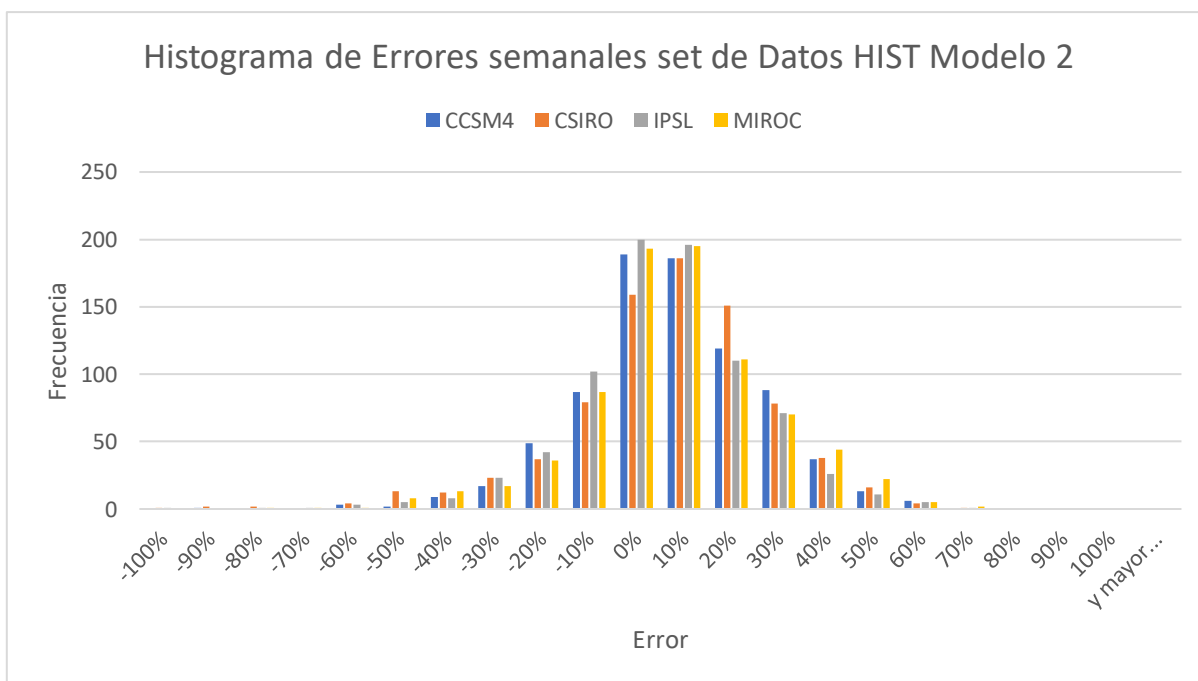


Figura 9.29: Histograma de errores semanales modelo 2 RNA calibrada con Set HIST

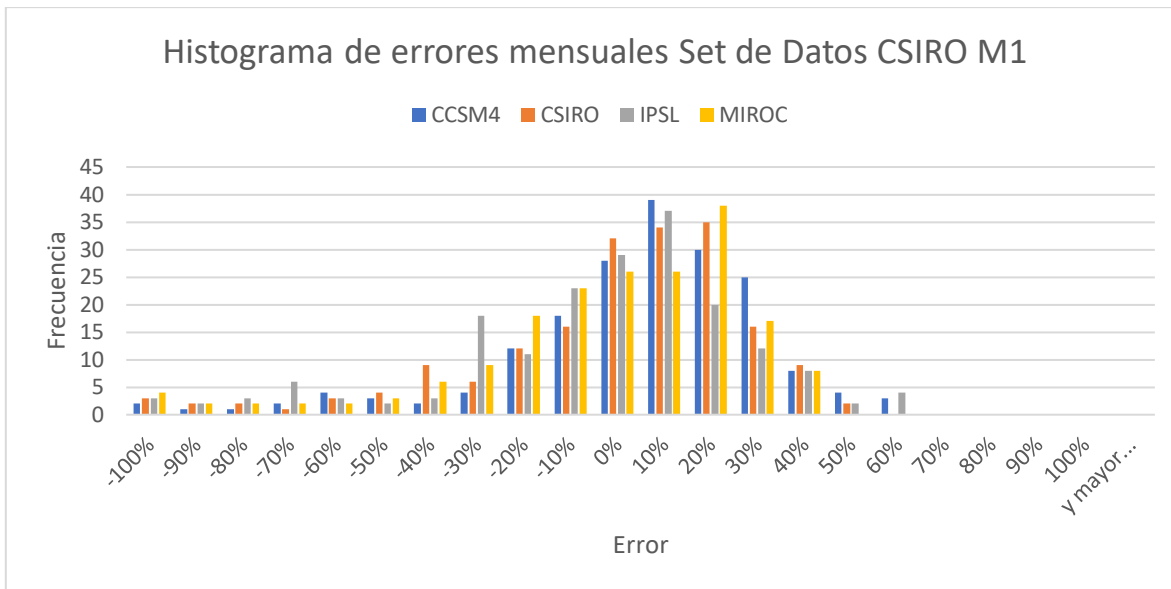


Figura 9.30: Histograma de errores mensuales modelo 1

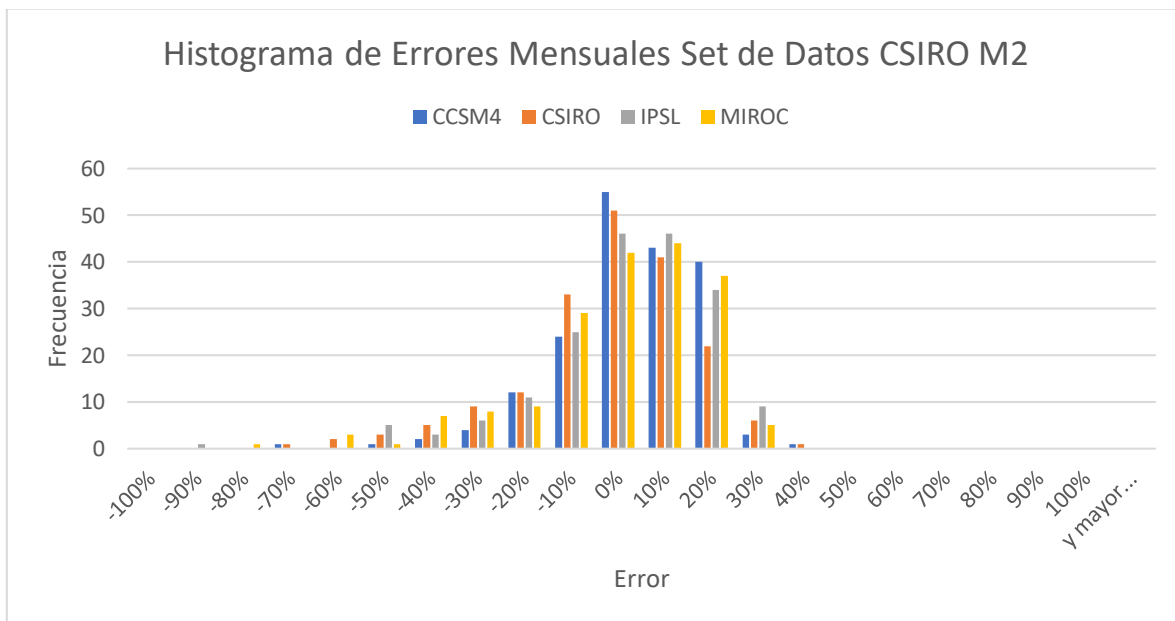


Figura 9.31: Histograma de errores mensuales modelo 2

Tabla 9.21: Frecuencia de errores de pronóstico semanales para las proyecciones de cada GCM con RNA Modelo 1

Error	CCSM4	CSIRO	IPSL	MIROC
-100%	0	1	4	0
-90%	0	0	3	0
-80%	3	2	5	1
-70%	4	4	5	5
-60%	3	7	10	9
-50%	5	6	13	16
-40%	14	22	24	22
-30%	26	37	42	50
-20%	43	67	70	75
-10%	92	79	117	95
0%	151	119	127	74
10%	138	144	124	128
20%	131	149	110	140
30%	104	104	55	106
40%	47	51	51	39
50%	24	9	35	16
60%	14	5	8	20
70%	4	0	3	3
80%	3	0	0	4
90%	0	0	0	3
100%	0	0	0	0
Y mayor...	0	0	0	0

Tabla 9.22: Frecuencia de Errores semanales modelo 2

Error	CCSM4	CSIRO	IPSL	MIROC
-100%	0	1	1	0
-90%	1	2	0	0
-80%	0	2	1	1
-70%	0	0	1	1
-60%	3	4	3	1
-50%	2	13	5	8
-40%	9	12	8	13
-30%	17	23	23	17
-20%	49	37	42	36
-10%	87	79	102	87
0%	189	159	200	193
10%	186	186	196	195
20%	119	151	110	111
30%	88	78	71	70
40%	37	38	26	44
50%	13	16	11	22
60%	6	4	5	5
70%	0	1	1	2
80%	0	0	0	0
90%	0	0	0	0
100%	0	0	0	0
Y mayor...	0	0	0	0

Tabla 9.23: Frecuencia de Errores por temporada Pronóstico Semanal modelo 1

Error	CCSM4	CSIRO	IPSL	MIROC
-50%	0	0	0	0
-40%	0	0	0	1
-30%	0	1	0	1
-20%	0	0	3	1
-10%	3	3	6	2
0%	10	8	7	4
10%	6	9	9	13
20%	7	6	2	5
30%	2	3	3	2
40%	2	1	0	1
50%	1	0	1	0
Y mayor...	0	0	0	1

Tabla 9.24: Frecuencia de errores pronóstico semanal modelo 2

Error	CCSM4	CSIRO	IPSL	MIROC
-50%	0	0	0	0
-40%	0	0	0	0
-30%	0	0	0	0
-20%	0	0	0	0
-10%	0	1	0	0
0%	3	4	7	5
10%	25	25	23	25
20%	3	1	1	1
30%	0	0	0	0
40%	0	0	0	0
50%	0	0	0	0
Y mayor...	0	0	0	0

Tabla 9.25: Frecuencia de errores mensuales modelo 1. RNA calibrada con...

Error	CCSM4	CSIRO	IPSL	MIROC
-100%	2	3	3	4
-90%	1	2	2	2
-80%	1	2	3	2
-70%	2	1	6	2
-60%	4	3	3	2
-50%	3	4	2	3
-40%	2	9	3	6
-30%	4	6	18	9
-20%	12	12	11	18
-10%	18	16	23	23
0%	28	32	29	26
10%	39	34	37	26
20%	30	35	20	38
30%	25	16	12	17
40%	8	9	8	8
50%	4	2	2	0
60%	3	0	4	0
70%	0	0	0	0
80%	0	0	0	0
90%	0	0	0	0
100%	0	0	0	0
Y mayor...	0	0	0	0

Tabla 9.26: Frecuencia de Errores mensuales modelo 2

Error [%]	CCSM4	CSIRO	IPSL	MIROC
-100%	0	0	0	0
-90%	0	0	1	0
-80%	0	0	0	1
-70%	1	1	0	0
-60%	0	2	0	3
-50%	1	3	5	1
-40%	2	5	3	7
-30%	4	9	6	8
-20%	12	12	11	9
-10%	24	33	25	29
0%	55	51	46	42
10%	43	41	46	44
20%	40	22	34	37
30%	3	6	9	5
40%	1	1	0	0
50%	0	0	0	0
60%	0	0	0	0
70%	0	0	0	0
80%	0	0	0	0
90%	0	0	0	0
100%	0	0	0	0
Y mayor...	0	0	0	0

Tabla 9.27: Frecuencias de error por temporada pronóstico mensual modelo 1

Error	CCSM4	CSIRO	IPSL	MIROC
-50%	0	1	0	1
-40%	0	0	0	0
-30%	0	1	0	0
-20%	0	1	3	1
-10%	4	4	11	5
0%	10	6	8	10
10%	7	12	4	9
20%	7	4	2	2
30%	1	2	2	3
40%	2	0	0	0
50%	0	0	1	0
>50%	0	0	0	0

Tabla 9.28: Frecuencia de errores pronóstico mensual modelo 2

Error	CCSM4	CSIRO	IPSL	MIROC
-50%	0	0	0	0
-40%	0	0	0	0
-30%	0	1	0	0
-20%	0	0	0	1
-10%	0	3	0	2
0%	6	6	8	6
10%	24	20	21	21
20%	1	1	2	1
30%	0	0	0	0
40%	0	0	0	0
50%	0	0	0	0
>50%	0	0	0	0

Tabla 9.29: Frecuencia de errores por temporada pronóstico semestral

Error	CCSM4	CSIRO	IPSL	MIROC
-50%	0	0	0	0
-40%	0	0	0	0
-30%	0	0	0	0
-20%	0	3	2	2
-10%	7	6	6	3
0%	14	16	13	13
10%	8	4	6	12
20%	1	1	1	1
30%	1	1	3	0
40%	0	0	0	0
50%	0	0	0	0
>50%	0	0	0	0

Anexo F. Errores del Pronóstico por temporada y modelo.

Tabla 9.30: Error por temporada pronóstico semanal GCM CCSM4 escenario 1

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	0%	-6%	8%	-22%	-15%
2021	31%	-17%	12%	10%	38%
2022	-8%	17%	16%	19%	-1%
2023	12%	-18%	3%	2%	13%
2024	-4%	25%	26%	2%	-20%
2025	8%	-1%	-8%	24%	1%
2026	2%	28%	-8%	-12%	-2%
2027	0%	-17%	10%	-59%	12%
2028	2%	-21%	12%	-27%	4%
2029	6%	-17%	8%	-17%	17%
2030	5%	3%	9%	-5%	-2%
2031	-6%	8%	-10%	-32%	-9%
2032	-19%	0%	-34%	21%	-20%
2033	10%	-32%	6%	-1%	-10%
2034	-14%	6%	-6%	12%	11%
2035	12%	1%	-13%	6%	9%
2036	-7%	-9%	-46%	3%	1%
2037	4%	20%	19%	-15%	8%
2038	-3%	-23%	4%	0%	-7%
2039	11%	1%	-10%	-14%	4%
2040	12%	-17%	-6%	0%	18%
2041	25%	3%	3%	33%	19%
2042	4%	4%	1%	-13%	-3%
2043	-48%	43%	-25%	-29%	10%
2044	13%	-32%	8%	4%	8%
2045	22%	-8%	-3%	-7%	17%
2046	5%	-5%	14%	8%	21%
2047	-8%	-3%	0%	10%	6%
2048	7%	-7%	-2%	15%	12%
2049	8%	-19%	-17%	26%	18%
2050	10%	1%	-13%	15%	0%

Tabla 9.31: Error por temporada pronóstico semanal GCM CSIRO escenario 1

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	-5%	-3%	18%	11%	16%
2021	3%	19%	43%	7%	29%
2022	-24%	38%	0%	25%	23%
2023	3%	27%	27%	11%	10%
2024	20%	19%	14%	30%	30%
2025	5%	40%	30%	15%	16%
2026	12%	5%	5%	33%	16%
2027	4%	2%	19%	-5%	1%
2028	12%	18%	19%	-2%	24%
2029	2%	16%	1%	15%	16%
2030	6%	7%	20%	12%	16%
2031	2%	20%	2%	16%	10%
2032	-38%	27%	-21%	14%	-17%
2033	9%	18%	25%	-12%	25%
2034	-6%	2%	-10%	20%	10%
2035	3%	19%	21%	14%	17%
2036	-25%	21%	12%	7%	-35%
2037	22%	11%	24%	27%	18%
2038	9%	27%	24%	-7%	18%
2039	11%	16%	11%	27%	18%
2040	-3%	28%	29%	-8%	6%
2041	2%	43%	39%	27%	21%
2042	1%	11%	21%	20%	-1%
2043	-26%	-1%	-39%	46%	-5%
2044	3%	20%	28%	-2%	12%
2045	6%	16%	34%	13%	4%
2046	12%	-3%	15%	15%	20%
2047	4%	12%	-5%	15%	22%
2048	-3%	31%	15%	2%	19%
2049	-2%	13%	15%	-9%	5%
2050	-26%	35%	27%	-3%	-8%

Tabla 9.32: Error por temporada pronóstico semanal GCM IPSL escenario 1

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	-5%	2%	4%	4%	-11%
2021	13%	34%	-6%	23%	9%
2022	23%	4%	17%	14%	20%
2023	12%	12%	-18%	1%	-5%
2024	11%	-4%	29%	28%	7%
2025	30%	6%	-5%	3%	24%
2026	-12%	11%	31%	0%	-7%
2027	-15%	9%	-23%	7%	-19%
2028	-3%	5%	-13%	18%	-6%
2029	-2%	7%	-2%	16%	-18%
2030	4%	17%	5%	13%	1%
2031	5%	-1%	3%	1%	-6%
2032	14%	-27%	1%	-28%	17%
2033	7%	3%	-27%	14%	0%
2034	-6%	-5%	8%	6%	-2%
2035	10%	13%	0%	3%	10%
2036	11%	5%	-9%	-25%	1%
2037	-6%	2%	20%	26%	-14%
2038	1%	1%	-9%	8%	6%
2039	8%	16%	17%	1%	4%
2040	14%	14%	-21%	8%	14%
2041	33%	23%	5%	4%	33%
2042	-3%	-4%	12%	10%	-2%
2043	-9%	-28%	38%	-7%	-17%
2044	16%	13%	-5%	7%	11%
2045	12%	13%	0%	0%	-4%
2046	4%	-7%	-8%	16%	0%
2047	16%	-3%	9%	3%	-4%
2048	29%	2%	-1%	8%	27%
2049	10%	-2%	-21%	0%	10%
2050	23%	14%	2%	-13%	23%

Tabla 9.33: Error por temporada pronóstico semanal GCM MIROC escenario I

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	-5%	-3%	18%	11%	16%
2021	3%	19%	43%	7%	29%
2022	-24%	38%	0%	25%	23%
2023	3%	27%	27%	11%	10%
2024	20%	19%	14%	30%	30%
2025	5%	40%	30%	15%	16%
2026	12%	5%	5%	33%	16%
2027	4%	2%	19%	-5%	1%
2028	12%	18%	19%	-2%	24%
2029	2%	16%	1%	15%	16%
2030	6%	7%	20%	12%	16%
2031	2%	20%	2%	16%	10%
2032	-38%	27%	-21%	14%	-17%
2033	9%	18%	25%	-12%	25%
2034	-6%	2%	-10%	20%	10%
2035	3%	19%	21%	14%	17%
2036	-25%	21%	12%	7%	-35%
2037	22%	11%	24%	27%	18%
2038	9%	27%	24%	-7%	18%
2039	11%	16%	11%	27%	18%
2040	-3%	28%	29%	-8%	6%
2041	2%	43%	39%	27%	21%
2042	1%	11%	21%	20%	-1%
2043	-26%	-1%	-39%	46%	-5%
2044	3%	20%	28%	-2%	12%
2045	6%	16%	34%	13%	4%
2046	12%	-3%	15%	15%	20%
2047	4%	12%	-5%	15%	22%
2048	-3%	31%	15%	2%	19%
2049	-2%	13%	15%	-9%	5%
2050	-26%	35%	27%	-3%	-8%

Tabla 9.34: Errores por temporada pronóstico Mensual GCM CCSM4 escenario 1

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	-9%	-9%	-16%	0%	1%
2021	13%	15%	19%	19%	27%
2022	21%	16%	20%	20%	33%
2023	-2%	-6%	3%	-2%	12%
2024	6%	0%	15%	5%	25%
2025	30%	21%	25%	28%	42%
2026	-8%	-8%	-14%	0%	3%
2027	-14%	-19%	-3%	-11%	10%
2028	3%	3%	3%	-2%	19%
2029	1%	-14%	-4%	1%	20%
2030	0%	-8%	6%	6%	15%
2031	3%	6%	10%	10%	25%
2032	20%	20%	20%	23%	33%
2033	9%	8%	15%	12%	25%
2034	-7%	-12%	-10%	-4%	-20%
2035	15%	17%	24%	19%	19%
2036	10%	-1%	15%	11%	31%
2037	-15%	-18%	-3%	-10%	9%
2038	6%	7%	3%	2%	18%
2039	6%	0%	15%	7%	25%
2040	15%	13%	19%	18%	34%
2041	33%	30%	34%	32%	42%
2042	4%	-2%	5%	8%	11%
2043	-1%	2%	6%	5%	15%
2044	13%	15%	21%	18%	16%
2045	11%	-2%	9%	9%	28%
2046	-1%	-7%	-5%	2%	-14%
2047	6%	9%	12%	12%	6%
2048	31%	32%	38%	35%	35%
2049	11%	6%	8%	13%	-1%
2050	21%	17%	29%	21%	36%

Tabla 9.35: Errores por temporada pronóstico Mensual GCM CSIRO escenario 1

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	1%	-10%	7%	-2%	18%
2021	31%	26%	37%	30%	45%
2022	-6%	-6%	-3%	0%	-10%
2023	14%	13%	14%	8%	28%
2024	2%	2%	1%	7%	17%
2025	4%	4%	11%	8%	26%
2026	2%	2%	6%	12%	12%
2027	13%	0%	6%	1%	22%
2028	1%	-13%	1%	0%	21%
2029	-2%	-13%	-11%	-1%	-22%
2030	7%	6%	12%	12%	7%
2031	-9%	-13%	-10%	-4%	-19%
2032	-21%	-31%	-28%	-19%	-40%
2033	9%	9%	9%	13%	23%
2034	-12%	-20%	-17%	-9%	-28%
2035	14%	5%	20%	12%	30%
2036	9%	-3%	-4%	-2%	16%
2037	8%	4%	12%	11%	29%
2038	4%	-3%	4%	9%	10%
2039	15%	18%	25%	20%	22%
2040	15%	18%	24%	19%	20%
2041	25%	23%	31%	26%	37%
2042	2%	3%	2%	7%	17%
2043	-35%	-78%	-72%	-46%	-91%
2044	9%	0%	15%	7%	26%
2045	17%	20%	29%	17%	25%
2046	-4%	-4%	-2%	0%	-11%
2047	-7%	-8%	-5%	-2%	-14%
2048	4%	7%	14%	8%	9%
2049	4%	8%	11%	9%	5%
2050	12%	3%	11%	11%	30%

Tabla 9.36: Errores por temporada pronóstico Mensual GCM IPSL escenario 1

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	-1%	-11%	7%	-2%	16%
2021	-9%	-19%	-17%	-3%	16%
2022	11%	8%	21%	13%	29%
2023	-12%	-19%	-13%	-8%	15%
2024	28%	28%	24%	33%	36%
2025	-7%	-15%	2%	-1%	12%
2026	32%	29%	31%	32%	42%
2027	-12%	-11%	-16%	-6%	-11%
2028	-7%	-20%	-1%	-10%	13%
2029	-13%	-24%	-20%	-8%	11%
2030	-1%	-5%	10%	0%	20%
2031	-1%	-7%	8%	-1%	18%
2032	-1%	-4%	9%	1%	19%
2033	-21%	-21%	-26%	-12%	-6%
2034	-5%	-8%	7%	-4%	15%
2035	-8%	-15%	2%	-9%	13%
2036	-3%	-9%	-2%	2%	3%
2037	19%	13%	25%	18%	35%
2038	-22%	-24%	-4%	-14%	8%
2039	18%	6%	18%	16%	33%
2040	-11%	-11%	-16%	-3%	3%
2041	10%	11%	6%	17%	21%
2042	9%	9%	15%	13%	9%
2043	45%	42%	47%	46%	51%
2044	-19%	-13%	-6%	-11%	12%
2045	0%	-15%	0%	-3%	20%
2046	-1%	-7%	2%	3%	8%
2047	12%	6%	12%	16%	18%
2048	-8%	-2%	2%	-2%	-5%
2049	-12%	-13%	-20%	-3%	-2%
2050	-2%	-9%	-7%	1%	-16%

Tabla 9.37: Errores por temporada pronóstico Mensual GCM MIROC escenario 1

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	4%	-2%	8%	7%	16%
2021	22%	12%	22%	23%	39%
2022	15%	7%	9%	16%	34%
2023	-1%	-5%	11%	0%	20%
2024	25%	26%	31%	29%	27%
2025	2%	-11%	5%	1%	23%
2026	-5%	-14%	-13%	0%	18%
2027	4%	1%	3%	8%	-5%
2028	19%	10%	24%	21%	34%
2029	12%	7%	14%	17%	21%
2030	8%	0%	15%	13%	23%
2031	5%	-9%	-5%	-7%	15%
2032	-23%	-28%	-20%	-18%	-29%
2033	17%	5%	11%	17%	34%
2034	3%	-4%	10%	10%	19%
2035	-4%	-12%	7%	-2%	20%
2036	-32%	-53%	-49%	-33%	-64%
2037	23%	23%	28%	28%	23%
2038	7%	0%	15%	7%	26%
2039	6%	-8%	0%	6%	24%
2040	3%	3%	10%	8%	5%
2041	6%	8%	13%	11%	22%
2042	5%	-5%	-3%	6%	-13%
2043	-16%	-18%	-11%	-11%	-18%
2044	3%	-1%	1%	7%	-7%
2045	4%	4%	10%	8%	20%
2046	18%	21%	23%	23%	18%
2047	6%	-3%	10%	10%	28%
2048	12%	13%	19%	16%	14%
2049	-8%	-4%	0%	-2%	19%
2050	-8%	-13%	-10%	-8%	7%

Tabla 9.38: Error total por temporada pronóstico semestral GCM CCSM4

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	-16%	-4%	-22%	-39%	-17%
2021	8%	-7%	-1%	1%	10%
2022	7%	-1%	9%	-19%	-3%
2023	-15%	37%	26%	-29%	-38%
2024	0%	-7%	1%	-9%	-4%
2025	22%	27%	17%	4%	18%
2026	0%	26%	8%	-25%	-9%
2027	-15%	-30%	-24%	-18%	-11%
2028	-3%	-9%	-9%	-13%	-3%
2029	-8%	-11%	-13%	-21%	-11%
2030	-2%	5%	-1%	-8%	-3%
2031	5%	0%	-1%	-3%	6%
2032	4%	-10%	1%	-10%	0%
2033	6%	-12%	-3%	-3%	5%
2034	-11%	-16%	-35%	-79%	-23%
2035	-14%	-18%	-15%	-35%	-22%
2036	-2%	-17%	-14%	-12%	-2%
2037	-10%	-9%	-14%	-13%	-6%
2038	-11%	-9%	-13%	-22%	-12%
2039	-5%	-19%	-10%	-14%	-7%
2040	2%	-10%	-2%	-5%	1%
2041	8%	9%	6%	-24%	-2%
2042	-8%	4%	-6%	-28%	-16%
2043	-3%	-2%	-5%	-8%	-2%
2044	-1%	-3%	-9%	-27%	-7%
2045	-4%	1%	-9%	-24%	-8%
2046	-7%	-10%	-28%	-73%	-21%
2047	-10%	-19%	-27%	-44%	-14%
2048	4%	-14%	-4%	-16%	-1%
2049	11%	6%	-9%	-50%	-2%
2050	0%	-11%	-2%	-17%	-6%

Tabla 9.39: Error total por temporada pronóstico semestral GCM CSIRO

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	-5%	-7%	-3%	-6%	-6%
2021	26%	15%	22%	18%	24%
2022	-4%	-8%	-27%	-59%	-11%
2023	-3%	-25%	-13%	-16%	-5%
2024	-5%	-13%	-12%	-18%	-6%
2025	-6%	-21%	-10%	-16%	-8%
2026	-12%	-7%	-18%	-30%	-14%
2027	-6%	-4%	-14%	-18%	-7%
2028	-8%	-11%	-5%	-25%	-15%
2029	-3%	-14%	-31%	-86%	-21%
2030	-11%	-17%	-26%	-41%	-14%
2031	-6%	-16%	-35%	-74%	-15%
2032	-20%	-34%	-56%	-110%	-36%
2033	-12%	-32%	-19%	-27%	-15%
2034	-9%	-25%	-44%	-92%	-23%
2035	-2%	-15%	-11%	-11%	-3%
2036	-6%	8%	3%	-22%	-15%
2037	10%	6%	9%	8%	10%
2038	-1%	8%	-6%	-25%	-4%
2039	6%	-2%	-4%	-17%	2%
2040	4%	9%	2%	-18%	-3%
2041	1%	-17%	-6%	-24%	-6%
2042	-13%	-3%	-6%	-25%	-21%
2043	-11%	-80%	-99%	-200%	-43%
2044	-10%	-10%	-3%	-34%	-21%
2045	-7%	-3%	6%	-36%	-24%
2046	-9%	-10%	-25%	-65%	-22%
2047	-21%	-24%	-41%	-73%	-29%
2048	-21%	-14%	-17%	-48%	-34%
2049	-15%	-11%	-21%	-46%	-24%
2050	5%	-4%	1%	-11%	0%

Tabla 9.40: Error total por temporada pronóstico semestral GCM IPSL

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	0%	10%	4%	4%	3%
2021	-4%	3%	-11%	-10%	-2%
2022	6%	-2%	3%	1%	6%
2023	-2%	-8%	-6%	4%	3%
2024	23%	31%	21%	11%	23%
2025	-5%	-4%	-7%	-3%	-2%
2026	24%	31%	26%	18%	22%
2027	-12%	2%	-18%	-51%	-19%
2028	-10%	-26%	-19%	-11%	-7%
2029	-16%	1%	-5%	-28%	-24%
2030	1%	-5%	-2%	-4%	1%
2031	-4%	3%	3%	-8%	-8%
2032	-5%	-8%	-8%	-8%	-3%
2033	-13%	-7%	-25%	-39%	-11%
2034	-14%	-14%	-11%	-21%	-18%
2035	-12%	-9%	-5%	-19%	-17%
2036	-12%	-5%	-17%	-36%	-17%
2037	3%	-12%	-8%	-15%	0%
2038	-8%	-7%	-11%	-3%	-2%
2039	7%	6%	2%	0%	7%
2040	-8%	3%	-14%	-29%	-9%
2041	-4%	-9%	-13%	-17%	-3%
2042	12%	19%	3%	-24%	6%
2043	26%	12%	22%	11%	21%
2044	-8%	-7%	-8%	-2%	-6%
2045	-3%	-8%	-5%	-5%	-3%
2046	-10%	4%	-7%	-27%	-16%
2047	10%	7%	-2%	-11%	10%
2048	-20%	-9%	-22%	-55%	-32%
2049	-23%	-17%	-33%	-46%	-22%
2050	-6%	-19%	-38%	-76%	-15%

Tabla 9.41: Error total por temporada pronóstico semestral GCM MIROC

Temporada	CCSM4	CSIRO	IPSL	MIROC	HIST
2020	1%	1%	-3%	-14%	-2%
2021	5%	-4%	-9%	-14%	4%
2022	7%	5%	-4%	-11%	5%
2023	-1%	0%	0%	-6%	-3%
2024	13%	17%	8%	-9%	8%
2025	-11%	-11%	-9%	-18%	-14%
2026	-2%	8%	-2%	-8%	-4%
2027	2%	-1%	-17%	-57%	-10%
2028	7%	-5%	-1%	-1%	7%
2029	3%	-1%	-4%	-13%	1%
2030	-1%	-5%	-2%	-6%	-2%
2031	-4%	-3%	-11%	-11%	-4%
2032	-19%	-14%	-37%	-80%	-27%
2033	2%	4%	-11%	-20%	0%
2034	5%	1%	-1%	0%	8%
2035	-6%	-17%	-12%	-11%	-6%
2036	-20%	-53%	-74%	-150%	-43%
2037	-2%	-17%	-16%	-27%	-6%
2038	1%	-14%	-6%	-6%	1%
2039	-1%	-4%	-5%	-14%	-5%
2040	-4%	-4%	-18%	-36%	-7%
2041	-5%	-20%	-11%	-13%	-6%
2042	-2%	-11%	-29%	-73%	-14%
2043	-27%	-28%	-47%	-70%	-29%
2044	-7%	-19%	-35%	-65%	-15%
2045	-6%	-18%	-9%	-16%	-9%
2046	0%	-5%	-10%	-28%	-6%
2047	4%	-9%	-3%	2%	5%
2048	-6%	-10%	-16%	-29%	-9%
2049	1%	1%	-3%	0%	2%
2050	-11%	2%	-12%	-23%	-11%

Mes	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Enero	Febrero	Marzo
Octubre	0,24	0,43	0,62	0,62	0,65	0,74	1,00					
Noviembre	0,15	0,33	0,67	0,62	0,62	0,63	0,82	1,00				
Diciembre	0,19	0,38	0,62	0,65	0,67	0,59	0,63	0,88	1,00			
Enero	0,18	0,38	0,59	0,63	0,57	0,53	0,44	0,66	0,92	1,00		
Febrero	0,20	0,37	0,58	0,59	0,53	0,48	0,37	0,57	0,86	0,99	1,00	
Marzo	0,18	0,36	0,57	0,59	0,52	0,48	0,38	0,56	0,85	0,98	0,99	1,00

Figura 9.37: Autocorrelación Q mensual

Mes	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Enero	Febrero	Marzo
Octubre	0,18	0,19	0,29	0,56	0,13	-0,13	0,24					
Noviembre	0,80	0,23	0,27	-0,35	-0,56	-0,29	0,12	0,60				
Diciembre	0,21	0,24	0,12	-0,19	-0,26	-0,35	-0,14	0,59	0,21			
Enero	0,65	0,27	0,23	-0,24	-0,35	-0,32	-0,11	0,49	0,35	0,73		
Febrero	0,83	0,28	-0,19	-0,23	-0,38	-0,29	-0,14	0,44	0,41	0,90	0,43	
Marzo	-0,19	0,28	-0,73	-0,26	-0,39	-0,29	-0,12	0,43	0,46	0,88	0,39	0,48

Figura 9.38: Correlación Q-T mensual

Mes	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Enero	Febrero	Marzo
Octubre	0,13	0,27	0,45	0,57	0,62	0,60	0,45					
Noviembre	-0,22	0,36	0,66	0,75	0,82	0,82	0,75	0,52				
Diciembre	-0,13	0,37	0,74	0,88	0,95	0,96	0,94	0,84	0,55			
Enero	-0,14	0,32	0,73	0,88	0,93	0,95	0,97	0,98	0,83	0,64		
Febrero	-0,13	0,31	0,75	0,85	0,89	0,99	0,94	0,99	0,89	0,72	0,28	
Marzo	-0,14	0,28	0,69	0,84	0,88	0,98	0,93	0,98	0,89	0,72	0,28	0,93

Figura 9.39: Correlación Q-SWE mensual

Mes	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Enero	Febrero	Marzo
Octubre	0,13	0,27	0,45	0,57	0,62	0,60	0,45					
Noviembre	-0,22	0,36	0,66	0,75	0,82	0,82	0,75	0,52				
Diciembre	-0,13	0,37	0,74	0,88	0,95	0,96	0,94	0,84	0,55			
Enero	-0,14	0,32	0,73	0,88	0,93	0,95	0,97	0,98	0,83	0,64		
Febrero	-0,13	0,31	0,75	0,85	0,89	0,99	0,94	0,99	0,89	0,72	0,28	
Marzo	-0,14	0,28	0,69	0,84	0,88	0,98	0,93	0,98	0,89	0,72	0,28	0,93

Figura 9.40: Correlación Q-PP mensual

Mes	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Enero	Febrero	Marzo
Octubre	-0,38	-0,94	-0,29	-0,22	-0,94	-0,30	0,15					
Noviembre	-0,34	0,14	-0,23	0,17	-0,24	-0,18	0,48	-0,77				
Diciembre	-0,78	-0,64	-0,35	0,21	-0,44	-0,24	-0,28	-0,40	-0,15			
Enero	-0,11	-0,14	-0,38	-0,12	-0,46	-0,25	-0,35	-0,54	-0,25	-0,35		
Febrero	-0,14	-0,16	-0,36	-0,17	-0,47	-0,26	-0,42	-0,57	-0,34	-0,32	-0,34	
Marzo	-0,13	-0,17	-0,39	-0,21	-0,46	-0,27	-0,41	-0,56	-0,33	-0,33	-0,34	-0,43

Figura 9.41: Correlación Q-E Mensual

Tabla 9.42: Correlaciones cruzadas Semestrales de todas las variables.

Variable	Correlación cruzada		
	Deshielo (t=0)	Invierno Pasado (t-1)	Deshielo Pasado (t-2)
Caudal	1	0,51	0,43
Temperatura Media	-0,13	-0,03	-0,2
Precipitación Acumulada	0,22	0,57	0,26
Evapotranspiración	0,48	0,51	0,46
Equivalente en agua de nieve	0,59	0,63	0,65

Anexo H. Anexo Digital

En el anexo digital se encuentran los siguientes archivos:

Archivo	Descripción
ConglomeradoCCEneroconmacro.xlsx	Base de datos de todas las series históricas ocupadas en esta memoria, además contiene CVE y valores supuestos en el pronóstico
BANDAS DE INCERTIDUMBRE.xlsx	Contiene las bandas de incertidumbre de los modelos y GCM
DIAGNOSTICO RNA MENSUAL.xlsx	Errores y series pronosticadas a nivel mensual del modelo 1
DIAGNOSTICO RNA MENSUAL_INFO PERFECTA.xlsx	Errores y series pronosticadas a nivel mensual del modelo 2
Diagnostico RNA_semanal_infoperfecta.xlsx	Errores y series pronosticadas a nivel semanal del modelo 2
Diagnostico RNA_SEMESTRAL_v2.xlsx	Errores y series pronosticadas a nivel semestral modelo 1
Diagnostico RNA_V1.xlsx	Errores y series pronosticadas a nivel semanal del modelo 1
ERRORES VF.xlsx	Gráficos de errores vs caudal observado
Matrices mensuales.xlsx	Matrices de Correlación
Matrices semanales.xlsx	Matrices de Correlación
PCA.xlsx	Análisis de las componentes principales
MaipoMensual.xlsx	Resultados validación cruzada
Maiposemanal.xlsx	Resultados validación cruzada
Maiposemestrall.xlsx	Resultados validación cruzada

Apéndice A. Algoritmos de Entrenamiento de RNA

Perceptrón

Con vectores de entrada y salida observados (p y t) se calcula la salida como la suma ponderada de los pesos (w) y entradas más el valor del bias (b) luego de aplicarle la función de transferencia que es binaria (0 o 1), se define el error como $e=t-a$ donde “ t ” es la salida deseada y “ a ” la salida calculada. De esta manera se tienen 3 casos:

1. Si la entrada es presentada y la salida del perceptrón (a) es correcta, entonces el error $e=0$ y el vector de pesos w no se altera.
2. Si la respuesta es 0 y debió haber sido 1 ($e=1$), el vector de entrada p se agrega al vector de pesos w .
3. Si la respuesta es 1 y debió haber sido 0 ($e=-1$), vector de entrada p se le resta al vector de pesos w .

De esta manera los pesos varían según la ecuación 2 y 3:

$$\Delta w = (t - a)p^t = e * p^t \quad (15)$$

$$\Delta b = e \quad (16)$$

Donde b corresponde al bias, luego la actualización de los pesos y bias es de la siguiente manera:

$$W_{new} = W_{old} + e * p^t \quad (17)$$

$$b_{new} = b_{old} + e \quad (18)$$

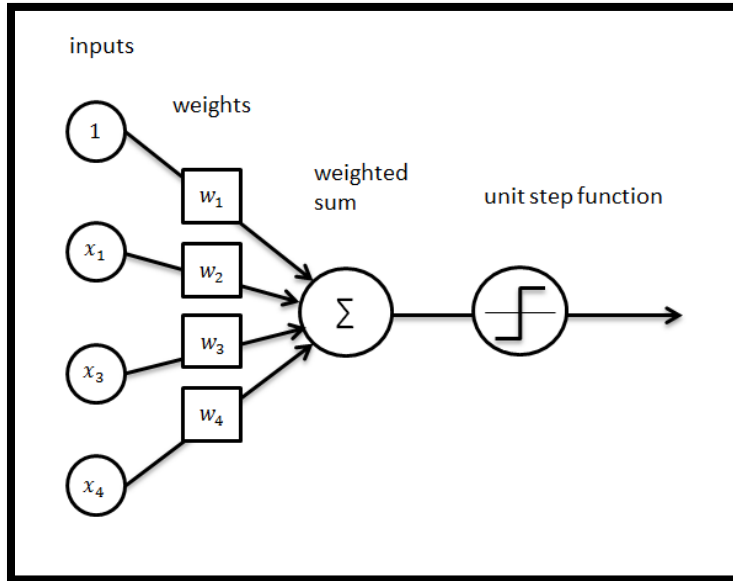


Figura 9.42: Perceptrón (Fuente: <http://ataspinar.com/2016/12/22/the-perceptron/>)

Regla LMS (Least Mean Square) o Mínimos cuadrados

En esta regla se utilizan neuronas con salida lineal (σ), la cual se conoce como ADALINE (Adaptive Linear Neuron Network). Para realizar el entrenamiento se cuenta con el algoritmo LMS o Regla de Aprendizaje Widrow-Hoff (Widrow and Hoff, 1960), la cual está basada en el método de optimización del gradiente o del paso descendente.

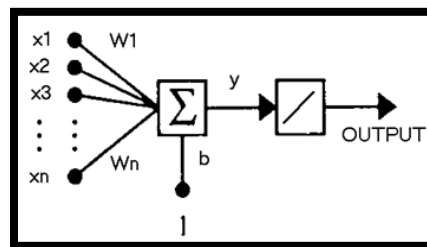


Figura 9.43: Neurona Lineal (<http://www.hugo-inc.com/RNA/Unidad%202/2.2.5.html>)

Conociendo los patrones de entrada y salida (p_i, t_i) y sea a la salida calculada, se define el error cuadrático medio (ECM):

$$ECM = \frac{1}{Q} \sum_{k=1}^Q (t(k) - a(k))^2 \quad (19)$$

La formulación general del algoritmo del gradiente tiene la siguiente ecuación:

$$W_{k+1} = W_k - \alpha \nabla_k \quad (20)$$

Donde k es el índice de la iteración y α es la tasa de aprendizaje. Luego de determinar el gradiente $(\frac{\partial e^2(k)}{\partial w_{i,j}})$ la actualización de los pesos es la siguiente:

$$W_{k+1} = w_k + 2 * \alpha * e(k) * p^t(k) \quad (21)$$

$$b_{k+1} = b_k + 2 * \alpha * e(k) \quad (22)$$

Retro propagación o backpropagation

El proceso de entrenamiento consta de 2 fases (Linnainmaa, 1970):

- Hacia adelante (forward): con los pesos fijos se calculan las repuestas de las unidades ocultas y las salidas de las entradas. Se determina el error.
- Hacia atrás (Backward): La señal de error es propagada hacia atrás usando los pesos de la red. Se ajustan los pesos.

La modificación de los pesos en este algoritmo se basa en minimizar la función objetivo, definida como el error o diferencia entre las salidas observadas y calculadas. Para esto se determina el gradiente negativo del error con respecto a los pesos.

Definido el error total cometido por la red como “Ep”, se determina el gradiente del error según los pesos “w”, obteniéndose las siguientes ecuaciones:

$$E_p = \frac{1}{2} * \sum_{a=1}^T (y_a - S'_a)^2 \quad (23)$$

Donde y_a es la salida observada, S'_a la salida simulada y T es el número de pares de datos entrada salida.

$$-\nabla E_p = (y_a - S'_a) * f'(S'_a) * i_j \quad (24)$$

La actualización de los pesos se considera proporcional al gradiente negativo, luego, los pesos entre la capa oculta y la capa de salida se actualizan según la siguiente ecuación:

$$w'_{ij}(t + 1) = w'_{ij}(t) + \Delta p * w'_{ij}(t) \quad (25)$$

$$\Delta p * w'_{ij}(t) = \eta(y_a - S'_a) * f'(S') * i_j \quad (26)$$

Donde η es la tasa de aprendizaje.

Dado lo anterior, se observa que las funciones de transferencia a utilizar tanto en las neuronas de la capa oculta como en la capa de salida deben ser diferenciables. Por otro lado, se observa que para la determinación de los nuevos pesos entre la capa de entrada y oculta, se debe calcular el gradiente negativo del error con respecto a los pesos correspondientes.

Variaciones al Algoritmo Backpropagation

Considerando una expresión genérica del algoritmo *Backpropagation*:

$$w_{k+1} = w_k + \alpha_k * D_k \quad (27)$$

Donde:

α_k : tasa de aprendizaje

D_k : dirección de búsqueda en la iteración k

De acuerdo con González (2003), “Los distintos métodos de optimización basados en el gradiente se diferencian en la forma de obtener la dirección de búsqueda D_k , a partir de la información de primer orden contenida con el gradiente ∇E_k , y en algunos casos de la estimación de la información adicional de segundo orden”.

Entre las variaciones al algoritmo de backpropagation destacan las siguientes:

- Método del Descenso del gradiente: $D_k = -\nabla E_k$
- Método del gradiente conjugado: $D_k = -\nabla E_k + \gamma_k * D_{k-1}$
- Método Cuasi-Newton: Consiste en determinar la matriz de Hessianos o segunda derivada; éste “converge a menudo más rápidamente que los métodos del gradiente conjugado, pero es complejo y requiere de alta capacidad computacional” (Socha D. 2005). Estos métodos se basan en el método de Newton, pero no requieren cálculo de la segunda derivada, dado que utilizan una aproximación de la matriz Hessiana en cada iteración

$$\alpha_k * D_k = -B_k * \nabla_k \quad (28)$$

Donde B_k es una estimación del inverso de la matriz Hessiana de E dada por:

$$\nabla^2 E = \left| \frac{\partial^2 E}{\partial w_i \partial w_j} \right| \quad (29)$$

Método de Levenberg-Marquardt: Similar a los métodos cuasi-Newton, el cual aproxima el gradiente y la matriz hessiana por las ecuaciones (17) y (18). “La matriz de Jacobianos puede computarse a través de Retro propagación normal, lo cual es mucho menos complejo de computar que la matriz de Hessianos” (Socha D.2005).

$$\nabla E_k = J^T * e \quad (30)$$

$$H = J^T * J \quad (31)$$

Sobre entrenamiento, Regularización y Detención Temprana.

El sobre entrenamiento es uno de los principales problemas a la hora de entrenar una RNA, este problema consiste en que la red se sobreajuste a los datos con los cuales está siendo entrenada, entregando un error muy bajo al simular la red con estos datos. Sin embargo, al simular la red con otros datos, el error aumenta en comparación al entrenamiento. Esto es mejorable mediante la regularización y la detención temprana.

Regularización

La regularización consiste en modificar la función objetivo, la cual es usualmente el error cuadrático medio cometido por la red (6), agregando un término que corresponde al promedio de la suma de los cuadrados de los pesos y bias de la red:

$$mse' = \gamma * mse + (1 - \gamma) * msw \quad (32)$$

$$msw = \frac{1}{n} \sum_{j=1}^n w_j^2 \quad (33)$$

Con gamma la razón de desempeño.

Detención temprana

Consiste en dividir el conjunto de datos en 3 grupos: entrenamiento, validación y prueba. La idea es que al final de cada iteración, la red simule con los datos de validación y se compare el error obtenido versus el error obtenido con los datos de entrenamiento en la última iteración. Normalmente el error de validación decrece durante la fase inicial del entrenamiento al igual, que el error de entrenamiento, sin embargo, cuando la red comienza a ajustar los datos, el error de validación comienza a aumentar, luego, el entrenamiento se detiene.

Apéndice B. MATLAB

Un volvox de Matlab es una “caja de herramientas” o un grupo de funciones que han sido creadas y agrupadas con una finalidad común; en este caso, el toolbox de redes neuronales de Matlab es un conjunto de funciones y/o instrucciones especializadas para crear, entrenar, validar y utilizar una red neuronal. Este toolbox viene incluido gratuitamente en los programas de Matlab más recientes como “Matlab R2014b”. Dentro de la gran cantidad de funciones disponibles en el toolbox de RNA, destacan las siguientes:

Para ingresar el set de datos de entrada y valores de salida observados, se deben ordenar en columnas cada par de datos de “inputs – Salida Observada”, siendo la primera fila el vector de salidas observadas y las filas siguientes los vectores de entrada.

Crear una red prealimentada o feedforward, con una capa oculta:

$$net = feedforwardnet(hiddenSizes,trainFcn)$$

Donde:

trainFcn = 'LM' (Método de Levenberg-Marquardt o alguna otra función de entrenamiento)

hiddenSizes = número de neuronas en la capa oculta

Determinar la función de transferencia para todas las capas:

$$net.layers{:}.transferFcn = 'tansig'$$

Por defecto, el toolbox de RNA divide el set de datos ingresados al azar en 3 grupos, con el 70% de los datos para entrenamiento, 15% para validación y 15% para prueba. De este modo, si se desea cambiar el porcentaje asignado a cada grupo, se escribe el porcentaje en fracción, donde el denominador es el número total de datos disponibles, 1000 por ejemplo, por lo tanto, si deseo que el 50% de los datos sean de entrenamiento, el numerador debe ser 500. Además, es posible modificar la manera de dividir el set de datos, la cual es por defecto aleatoria, sin embargo, se puede cambiar a división por bloques continuos, por intervalos o por un índice.

$$net.divideParam.trainRatio = 0.5$$
$$net.divideParam.valRatio = 0.25$$
$$net.divideParam.testRatio = 0.25$$

Al crear una red neuronal, el toolbox de RNA automáticamente asigna una función de pre procesamiento a las entradas y salidas observadas, la cual es por defecto “mapminmax” y normaliza las entradas y salidas observadas en el rango [-1,1]. Las funciones de pre procesamiento transforman los valores de las entradas y salidas observadas en valores que son mejores para el entrenamiento de la red, por ejemplo “si las entradas son muy grandes, los pesos deberán ser muy pequeños para prevenir que la función de transferencia se sature, por lo que usualmente se normalizan las entradas.” (Hudson et al. 2014). Existen otras funciones de pre procesamiento, como “mapstd” la cual normaliza las entradas y salidas observadas para que tengan media cero y varianza igual a 1.

En la capa oculta sea “tangente sigmoïdal” y por último la función de transferencia en la capa de salida sea “purelin”.

Existen seis parámetros de entrenamiento generales, que utilizan las distintas funciones de entrenamiento, los cuales son:

- Epochs = número de iteraciones máxima, el entrenamiento se detiene si se excede este número de iteraciones.
- Minigrad = magnitud mínima del gradiente, el entrenamiento se detiene si la magnitud del gradiente está por debajo de este valor.
- Goal = decrementos de la función de desempeño, el entrenamiento se detiene si los decrementos de la función de desempeño están por debajo de este valor.
- Time = tiempo máximo de entrenamiento, el entrenamiento se detiene si el tiempo de entrenamiento excede este valor.
- Max_fail = asociado con la técnica de detención temprana.
- Validation checks = representa el número de iteraciones sucesivas que la función de desempeño no decae. El entrenamiento se detiene si se excede el número de validation checks, el cual es por defecto igual a seis.

Para entrenar la red, se escribe: $[net,r] = trainlm(net,input,target)$

Donde:

Net = red creada

input = set de datos de entrada

target = set de datos de salida observados o deseados

Para simular la red con nuevos datos de entrada, una vez que ya ha sido entrenada, se escribe:

$TargetSimulados = sim(net,inputS)$










donde:

net = red ya entrenada

inputs = nuevos datos de entrada

El toolbox de RNA está programado según la configuración que produce el mejor rendimiento en la mayoría de los casos, esta configuración implica que por defecto el número de capas ocultas sea uno, el número de neuronas en la capa oculta sea 10, la función de entrenamiento sea “LM” (Método de Levenberg-Marquardt), la función de transferencia.

Tabla 9.43: Funciones de transferencia RNA (Artigas,2008)

Nombre	Relación Entrada /Salida	Icono	Función
Limitador Fuerte	$a = 0 \quad n < 0$ $a = 1 \quad n \geq 0$		<i>hardlim</i>
Escalón	$a = -1 \quad n < 0$ $a = +1 \quad n \geq 0$		<i>hardlims</i>
Lineal Positiva	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$		<i>poslin</i>
Lineal	$a = n$		<i>purelin</i>
Lineal Saturado	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$		<i>satlin</i>
Lineal Saturado Simétrico	$a = -1 \quad n < -1$ $a = n \quad -1 \leq n \leq 1$ $a = +1 \quad n > 1$		<i>satlins</i>
Sigmoidal Logarítmico	$a = \frac{1}{1 + e^{-n}}$		<i>logaritmo sigmoidal</i>
Tangente Sigmoidal Hiperbólica	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		<i>tangente sigmoidal</i>
Competitiva	$a = 1$ Neurona con n max $a = 0$ El resto de neuronas		<i>compet</i>