

UNIVERSIDAD CATOLICA DE LA SANTISIMA CONCEPCION
Facultad de Ingeniería
Ingeniería Civil



Predicción de la concentración de sedimento en suspensión utilizando redes neuronales artificiales y curva de gasto de sedimento.

MATÍAS OSVALDO HERMOSILLA BARRA

INFORME DE PROYECTO DE TÍTULO PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL

Profesor Guía

Diego Caamaño A.

Profesor Informante

Enrique Muñoz O.

Concepción, Diciembre 2016

RESUMEN

El presente proyecto tiene como fin estimar de manera precisa la concentración de sedimento en suspensión (C_{ss}) diaria en el río Cautín ubicado en la región de la Araucanía, Chile, a través de la aplicación de Redes Neuronales (ANN).

Se propone una arquitectura para la formulación de la ANN que se utilizó en la estimación de la concentración de sedimento en suspensión. Se utilizaron como variables de entrada la altura (H), el caudal (Q), y la precipitación diaria (PP). Mediante los estadísticos Percent Bias (PBIAS), Raíz del Error Cuadrático Medio (RMSE), se compararon los resultados obtenidos por el modelo de ANN y la curva de gasto de sedimento (CGS). Del análisis se pudo concluir que ambos modelos generan predicciones precisas de la C_{ss} , pero comparando los resultados entre los modelos, la ANN presentó mejores valores en los dos estadísticos mencionados.

Por otra parte, se analizaron 8 combinaciones de entradas a la ANN considerando las variables (H), (Q) y (PP) registradas el día que se midió la C_{ss} (tiempo "t") y el día previo a ésta medición (tiempo "t-1"). Según los resultados obtenidos en el presente estudio, se determinó que las variables que generan predicciones más acertadas de la C_{ss} en el modelo de ANN, son $h(t)$ $Q(t)$ $PP(t)$, considerando los estadísticos RMSE y PBIAS.

ABSTRACT

This research project has as main purpose to estimate suitably the daily sediment concentration in suspension (C_{ss}) at the Cautín River located in the Araucanía region, Chile throughout the implementation of neuronal networks (ANN).

On the one hand, it is proposed architecture to formulate the ANN that was used to estimate the sediment concentration in suspension. In addition, the main variables implemented in this research project were the following: height (H), water flow (Q), and daily rainfall (PP). By means of Percent Bias (PBIAS) statistics, Error root mean square (RMSE), the obtained results were compared to the ANN model as well as the sediment expense curve (CGS). Subsequently, from the analysis results it can be concluded that both models generate precise and accurate predictions of C_{ss} ; notwithstanding, comparing the results among models, the ANN showed better values in two statistics previously mentioned above.

On the other hand, 8 ANN entries combinations were analyzed considering (H), (Q), and (PP) variables recorded the day that (time “t”) was measured and the previous day before this measurement (time “t-1”). According to the obtained results in this research project it can said that variables that promote more accurate C_{ss} predictions in the ANN model are $h(t)$ $Q(t)$ $PP(t)$ considering the RMSE and PBIAS statistics.

A mi madre Lilian Del Carmen Barra Esparza.

AGRADECIMIENTOS

En primer lugar, quiero agradecer a todo el cuerpo académico y administrativos de la UCSC, en especial a los profesores de carrera por el esmero en transmitir sus conocimientos y valores durante mi estadía como estudiante.

Al profesor Diego Caamaño por el apoyo, consejo y ayuda que me entregado durante el desarrollo de este trabajo.

A mis amigos, con los cuales compartí momentos difíciles en el transcurso de la carrera, así como también excelentes momentos que se quedaron en mi memoria. Gracias por todas las aventuras que vivimos, por estar cuando más los necesite, por hacerme reír en todo momento, por reconocer mis logros y aconsejarme en mis errores y por el gran lazo de amistad que hemos formado.

A mis familiares por el ánimo que me entregaron día a día para seguir adelante, apoyándome cada uno como solo ellos lo saben hacer. A mis tíos, tías, primos y primas, por su cariño, por los consejos y la ayuda que me han entregado. A mi abuelito Enrique y mi abuelita Irma, los cuales estuvieron a mi lado gran parte de mi infancia y siempre creyeron en mí. A mi querida hermana Fernanda, agradezco su buena disposición cuando necesité ayuda, por soportarme, por la excelente empatía que ha tenido conmigo y por qué siempre ha estado a mi lado. Gracias a mi polola Bárbara, por el ánimo que me dio cuando no lo tenía, por hacerme más responsable con mis deberes, le doy gracias por el apoyo incondicional que me brindó en esta última etapa de mi carrera y por la preocupación que ha mostrado durante estos tres años y medio.

Finalmente agradezco a mi madre Lilian, quien ha puesto todo su esfuerzo para que mi hermana y yo podamos surgir en esta vida, para que tengamos un buen futuro. Si bien todos mis años de estudio no han sido fáciles, mi madre siempre estuvo ahí para apoyarme, con su cariño y paciencia, creyendo en mí en todo momento y como resultado he llegado hasta este punto. Sin dudar es una mujer ejemplar y admirable la cual me ha entregado excelentes

valores morales y de no ser por ella no estaría aquí. Desde antes que decidiera comenzar esta carrera, mi madre fue la motivación para alcanzar lo que he logrado hasta ahora, es por esto y por otras razones que además de estar eternamente agradecido de lo que ha hecho por mí, le dedico este trabajo.

TABLA DE CONTENIDO

1	Introducción	1
1.1	Motivación.	3
1.2	Objetivos.	4
2	Marco Teórico	5
2.1	Modelo de ANN.	5
2.1.1	Componentes básicos de una ANN.	6
2.1.2	Entradas múltiples en una ANN.	7
2.1.3	Una capa de neuronas con entradas múltiples.	8
2.1.4	Múltiples capas de neuronas.	9
2.1.5	Forma abreviada de la arquitectura de las redes neuronales.	10
2.1.6	Configurar la red neuronal.	11
2.1.7	Entrenamiento de redes neuronales.	11
2.1.7.1	<i>Función de entrenamiento.</i>	13
2.1.7.2	<i>Función de rendimiento de la ANN.</i>	14
2.1.8	Aplicación de ANN en el transporte de sedimento.	16
2.2	Variables del modelo.	18
3	Métodos	19
3.1	Rede neuronal de Multicapas Feedforward (MLP) utilizada.	19
3.2	Zona de estudio.	23
3.2.1	Descripción general del río.	24
3.2.2	Clima.	24
3.2.3	Uso de Suelo.	25
3.3	Datos para el modelo	26
3.3.1	Estaciones y datos para el modelo.	27
3.3.2	Métodos de Muestreos y obtención de datos.	28
3.3.3	Análisis de datos para el modelo.	29

3.3.4	Lectura de datos.	33
3.3.5	Procesamiento de entrada-salida.	35
3.3.6	Subconjunto de datos en la red neuronal.	36
3.4	Curva de gasto de sedimento (CGS)	39
3.5	Código Matlab	39
4	Resultados	41
4.1	Estimación de transporte de sedimento en suspensión mediante el modelo de ANN	41
4.2	Análisis de crecidas según resultados	44
4.3	Comparación del modelo de ANN con la CGS	47
4.4	Formulación matemática a partir de la ANN	53
5	DISCUSIÓN Y CONCLUSIONES	55
6	Referencias	57
7	Anexos	60
7.1	Anexo 1: Tablas de Estadístico de rendimiento de resultados 1.	60
7.2	Anexo 2: Grafica de datos estimados por el modelo de ANN vs los datos reales de Css para cada una de las crecidas	68

ÍNDICE DE FIGURAS

Figura 2.1: Componentes de una neurona simple (modificado de <i>Beale et al., 2015</i>). ____	6
Figura 2.2: ANN simple con R elementos de entrada (modificado de <i>Beale et al., 2015</i>).__	7
Figura 2.3: ANN con R elementos de entrada y S neuronas (modificado de <i>Beale et al., 2015</i>). _____	8
Figura 2.4: ANN con tres capas de neuronas (modificado de <i>Beale et al., 2015</i>). _____	9
Figura 2.5: Abreviatura del esquema de una ANN (modificado de <i>Beale et al., 2015</i>). __	10
Figura 3.1: Arquitectura de la Red Multicapas Feedforward utilizada en este estudio (modificado de <i>Beale et al., 2015</i>)._____	20
Figura 3.2: Área de estudio del proyecto (Red Hidrométrica). _____	23
Figura 3.3: Estación de Sedimento y Meteorológica. _____	27
Figura 3.4: Hidrograma de una tormenta aislada. _____	29
Figura 3.5: Hidrograma de la estación “Río Cautín en Rari-Rura”. _____	30
Figura 4.1: Promedio de PBIAS para cada crecida seleccionada. _____	46
Figura 4.2: Promedio de RMSE para cada crecida seleccionada. _____	47
Figura 4.3: Curva de Gasto de Sedimento. _____	48
Figura 4.4: Estimaciones de la ANN por sobre las de CGS. _____	49
Figura 4.5: RMSE de ANN vs RMSE de CGS. _____	52
Figura 4.6: PBIAS de ANN vs PBIAS de CGS. _____	52
Figura 7.1: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 1. _____	68
Figura 7.2: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 2. _____	68
Figura 7.3: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 3. _____	69
Figura 7.4: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 4. _____	69
Figura 7.5: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 5. _____	70
Figura 7.6: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 6. _____	70
Figura 7.7: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 7. _____	71
Figura 7.8: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 8. _____	71
Figura 7.9: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 9. _____	72
Figura 7.10: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 10. _____	72
Figura 7.11: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 11. _____	73

Figura 7.12: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 12.	73
Figura 7.13: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 13.	74
Figura 7.14: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 14.	74
Figura 7.15: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 15.	75
Figura 7.16: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 16.	75
Figura 7.17: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 17.	76
Figura 7.18: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 18.	76
Figura 7.19: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 19.	77
Figura 7.20: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 20.	77
Figura 7.21: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 21.	78
Figura 7.22: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 22.	78
Figura 7.23: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 23.	79
Figura 7.24: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 24.	79
Figura 7.25: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 25.	80
Figura 7.26: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 26.	80
Figura 7.27: Datos estimados vs datos reales de C _{ss} (mg/l) de la crecida 27.	81
Figura 7.28: Datos estimados vs datos reales de C _{ss} (mg/l) para los grupos de datos de entrenamiento y validación de la crecida 14.	81

ÍNDICE DE TABLAS

Tabla 3.1: Combinaciones de entradas para la ANN.	31
Tabla 3.2: Cantidad de datos y porcentajes de concentraciones de sedimento.	31
Tabla 3.3: Periodos de crecidas para el modelo de ANN.	32
Tabla 3.4: Porcentajes de subconjuntos de datos entrenamiento y validación para el modelo.	37
Tabla 3.5: Cantidad de datos de concentración de sedimento mg/l según intervalos.	38
Tabla 4.1: Índices de rendimiento del PBAS para estimaciones de sedimento (<i>Moriassi et al., 2007</i>).	41
Tabla 4.2: Estimaciones más acertadas de C _{ss} según el estadístico PBIAS para cada combinación de variables de entrada.	43
Tabla 4.3: Cantidad de pruebas que tienen indicadores de rendimiento por debajo de lo normal.	45
Tabla 4.4: Estadísticos PBIAS y RMSE de las estimaciones de C _{ss} de los modelos de CGS y ANN	50
Tabla 7.1: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB1.	60
Tabla 7.2: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB2.	61
Tabla 7.3: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB3.	62
Tabla 7.4: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB4.	63
Tabla 7.5: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB5.	64
Tabla 7.6: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB6.	65
Tabla 7.7: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB7.	66

Tabla 7.8: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB8. _____ 67

1 INTRODUCCIÓN

La estimación del sedimento en suspensión debe considerarse en una amplia gama de proyectos en recursos hídricos e hidráulicos tales como, el diseño y mantenimiento de canales, presas y embalses, así como también es un índice importante de la erosión de cuencas y ríos. Por estas razones es importante estimar con exactitud la concentración de sedimento en suspensión (*Rajae et al., 2009*).

Una de las formas más confiables para estimar el sedimento son los registros observados (i.e. medidos), sin embargo, los muestreos son difíciles de realizar y el proceso es lento y costoso (*Melesse et al., 2011*). Por estas razones ha surgido la necesidad de encontrar fórmulas y métodos para predecir el sedimento en suspensión de manera indirecta. Autores como Graf (1971), Einstein (1950), Van Rijn (1984), entre otros, han desarrollado fórmulas para determinar la carga de sedimento en suspensión (*Chanson, 2004*). Sin embargo, estos métodos requieren información detallada, incluyendo características hidrológicas, hidráulicas, geológicas y del sedimento. Por otro lado, al ser métodos complejos requieren de algunos supuestos para simplificar las formulas. Comúnmente las simplificaciones son aplicables al sitio de estudio, lo que dificulta la transferibilidad de los resultados. Esto trae como consecuencia que los métodos y fórmulas encontradas por los diferentes autores se apliquen a una zona reducida y específica (*Heng y Suetsugi, 2013*).

Varios estudios se han centrado en la predicción y la estimación del transporte de sedimento en suspensión utilizando enfoques calibrados según la información disponible de los registros de datos. Este es el caso de las redes neuronales artificiales (ANN, Artificial Neural Networks). La capacidad de las ANN para establecer vínculos no lineales entre las variables de entradas y las variables que se desean estimar la convierte en una herramienta útil para modelar fenómenos hidráulicos e hidrológicos complejos, como lo es la estimación acertada del sedimento en suspensión (*ASCE, 2000*).

La presente tesis aborda la estimación del transporte del sedimento en suspensión por medio del modelo de ANN utilizando una implementación desarrollada en MATLAB en conjunto

con una selección y pre procesamiento de varias variables que se consideran relevantes para la cuantificación de la cantidad de sedimento en suspensión.

1.1 Motivación.

Las redes neuronales artificiales es uno de los métodos que ha tenido en las últimas décadas buenos resultados en las áreas de ingeniería, creciendo la aplicación de este método en estudios de hidrología e hidráulica (*Rajae et al., 2009*).

El transporte de sedimento en suspensión es un proceso complejo por la cantidad de variables que influyen en éste y por las difíciles relaciones que se deben realizar para su correcta determinación.

Las ANN establecen relaciones lineales y no lineales entre variables de entradas y aquella que se desee estimar. Poseen una estructura simple, siendo fácil definir y modificar su arquitectura, la cantidad de variables en la entrada, por lo que no se limita a una zona en específica de estudio.

El presente proyecto busca implementar una red neuronal para estimar de forma acertada la concentración de sedimento en suspensión. Además, identificar la importancia de las variables de entrada en el modelo utilizando. Finalmente se comparará el rendimiento de la red con la curva de gasto de sedimento obtenida sobre la base de mediciones en terreno.

1.2 Objetivos.

Objetivo general:

- Identificar una arquitectura y los algoritmos respectivos para clasificar una ANN de forma tal que ésta sea capaz de estimar concentraciones de sedimentos en suspensión de manera acertada.

Objetivos específicos:

- Reconocer una arquitectura apropiada de una ANN para predecir de manera precisa las concentraciones de sedimento en suspensión en un curso fluvial.
- Definir las variables de entrada a la red que mejor representan el fenómeno.
- Aplicar un indicador estadístico para comparar resultados de estimación de sedimento sobre la base de diferentes eventos de tormenta.
- Obtener desde la arquitectura de la red una formulación matemática simple que tenga interpretación física.

2 MARCO TEÓRICO

2.1 Modelo de ANN.

Las ANN son un modelo “simplificado” de un sistema nervioso biológico. Una ANN consiste en un número de elementos de procesamiento de datos llamadas neuronas o nodos que se agrupan en capas. Las neuronas de las capas de entradas reciben datos de entrada (inputs) o información y transmiten los valores a la siguiente capa de elementos de procesamiento a través de conexiones. Este proceso continúa hasta llegar a la capa de salida, entregado como resultado lo que se desea estimar. Este tipo de red en el que los flujos de datos son sólo hacia adelante se conoce como feedforward (*Beale et al., 2015*). En la práctica, la arquitectura de las ANN se compone de dos elementos principales: capas intermedias (ocultas) y una capa de salida. Estas serán descritas posteriormente.

Las principales diferencias entre los tipos de ANN son la disposición de la arquitectura de la red, las formas de determinar los pesos, las distintas funciones de transferencias que existen y son implementadas en las redes neuronales, y el entrenamiento de la red junto su función de entrenamiento (*Beale et al., 2015*).

2.1.1 Componentes básicos de una ANN.

El bloque de construcción básico para las redes neuronales es la neurona de entrada única, como se muestra en la Figura 2.1.

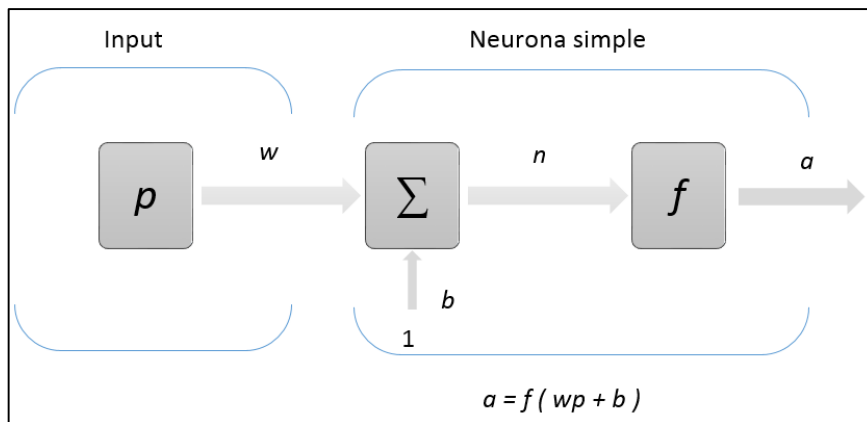


Figura 2.1: Componentes de una neurona simple (modificado de *Beale et al., 2015*).

En primer lugar, la entrada escalar p se multiplica por el peso escalar w para formar el producto wp , nuevo escalar. En segundo lugar, se suma entrada ponderada wp con el sesgo escalar b para formar la entrada neta n ($n=wp+b$). Por último, la entrada neta n pasa a través de la función de transferencia f , que produce la salida escalar a . Los nombres dados a estos tres procesos son: función de peso; función de entrada de red y función de transferencia.

Para muchos tipos de ANN, la función de peso es el producto de un peso multiplicado por la entrada, pero se puede utilizar cualquier tipo de función o relación matemática. La función de entrada neta de la red usualmente es la suma entre el peso ponderado wp con el sesgo b , también podría utilizar cualquier tipo de función. Finalmente, la función de transferencia es la que convierte la entrada neta n en la salida a , como se ilustra en la Figura 2.1.

Se debe tener en cuenta que w y b son ambos parámetros escalares ajustables de cada neurona. La idea central de las redes neuronales es que tales parámetros puedan ser ajustados de modo que la red presente el comportamiento deseado. Por lo tanto, se puede entrenar la red para hacer un trabajo en particular mediante el ajuste de los parámetros de peso y sesgo.

2.1.2 Entradas múltiples en una ANN.

Las neuronas simples pueden ser extendidas para manejar entradas del tipo vector. Una neurona con un único vector de entrada de R-elementos se muestra en la Figura 2.2.

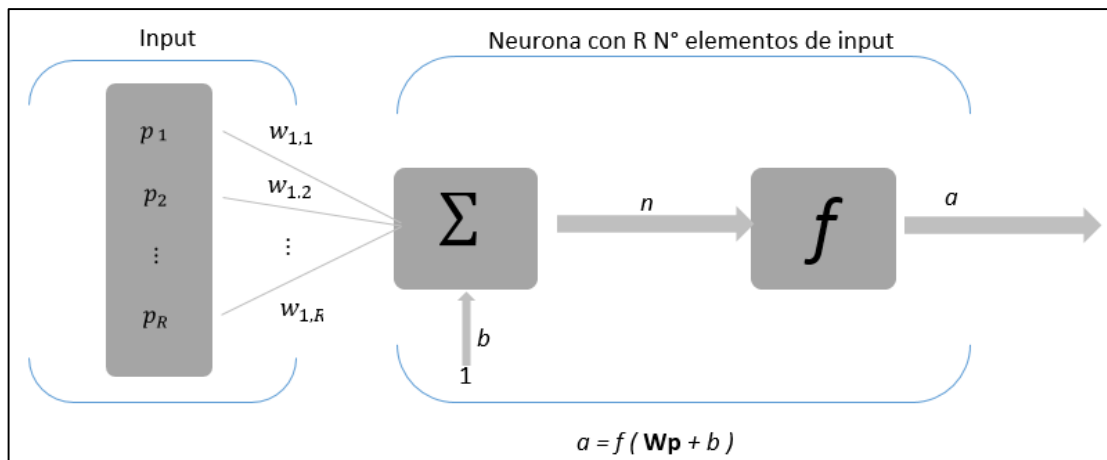


Figura 2.2: ANN simple con R elementos de entrada (modificado de Beale et al., 2015).

Los elementos de entrada del vector p se multiplican por la matriz de pesos W , de manera que se realiza un producto vectorial que se puede expresar como Wp . Por lo tanto, p es un vector de elementos de R entradas. Estas entradas se multiplican posteriormente con una única fila y con R columnas de la matriz W . La neurona mostrada anteriormente tiene un sesgo b , que se suma con las entradas ponderadas para formar la entrada neta n . Finalmente pasa por la función de transferencia para generar la salida a . Una capa de una red incluye los pesos, la multiplicación y la operación suma (producto vectorial Wp), el sesgo b y la función de transferencia. Las variables de entrada y el vector p , no se incluyen en la llamada capa.

2.1.3 Una capa de neuronas con entradas múltiples.

Una red neuronal con una única capa de R elementos de entrada y S neuronas se muestra en la Figura 2.3.

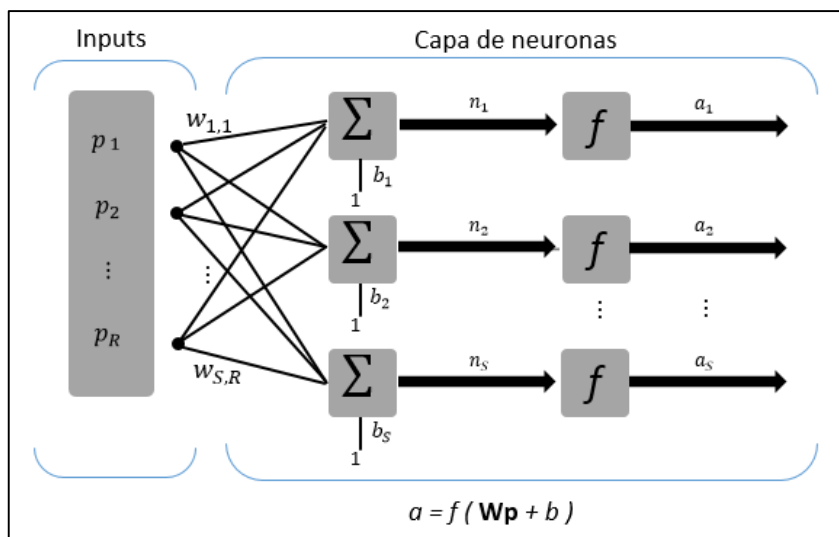


Figura 2.3: ANN con R elementos de entrada y S neuronas (modificado de *Beale et al.*, 2015).

Con R número de elementos del vector de entrada y S número de neuronas en la capa.

En esta red, cada elemento del vector de entrada p está conectado a cada entrada de la neurona a través de la matriz de ponderación W . La neurona i -ésima tiene un estilo que reúne sus entradas y sesgos ponderados para formar su propia salida n_i . Los diversos n_i forman en su conjunto S -elementos del vector de entrada neta n . Finalmente, las salidas de la neurona forman un vector columna a .

Es común que el número de entradas de una capa sea diferente al número de neuronas, es decir, R no es necesariamente igual a S . Los índices de correspondientes a la fila en los elementos de la matriz W indican la neurona de destino, y los índices de la columna indican la fuente de entrada.

2.1.4 Múltiples capas de neuronas.

Una red puede tener varias capas. Cada capa tiene una matriz W de pesos, un vector sesgo b , y un vector de salida a . Para distinguir estos elementos de cada una de las capas, se les añade a estas variables un número como superíndice correspondiente a la capa que pertenece. A continuación en la Figura 2.4, se muestra una red con tres capas de neuronas para ejemplificar lo dicho anteriormente.

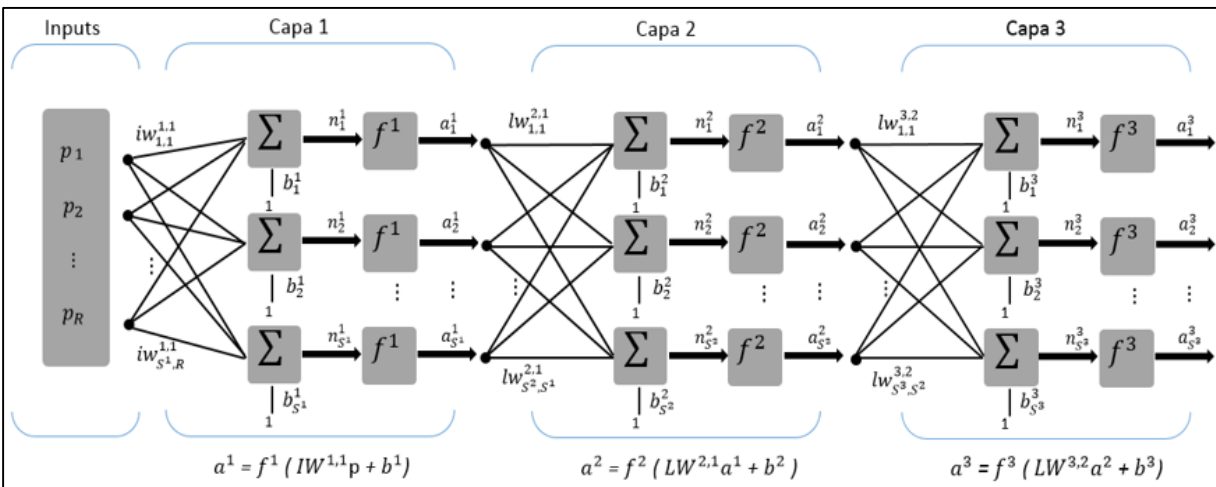


Figura 2.4: ANN con tres capas de neuronas (modificado de Beale et al.,2015).

La red de la Figura 2.4 contiene R entradas, s^1 neuronas en la primera capa, s^2 neuronas en la segunda capa, etc. Las salidas de cada capa intermedia son las entradas de la capa siguiente. Entonces la entrada de la capa dos es a^1 y la salida es a^2 , y así sucesivamente hasta llegar a la última capa.

A la última capa de la red (la que genera la salida de la red neuronal) se le llama capa de salida (layer output). A las demás capas de la red se les denominan capas ocultas (layer hidden). Consecuentemente, la red de la Figura 2.4 posee una capa de salida (i.e. la tercera capa) y dos capas ocultas.

Otra distinción de las redes multicapas son la clasificación de dos matrices de pesos. Se llaman a las matrices de pesos conectadas a la entrada, pesos de entrada (input Weights) y a

las matrices de pesos que conectan capas, pesos de capas (layer Weights). Dado su nombre en inglés, es que su abreviatura es IW para la matriz de pesos de entrada y LW para la matriz de pesos de capas. Además, se añaden superíndices para indicar la fuente (segundo índice) y el destino (primer índice) como se muestra en la red anterior.

En la parte inferior de cada capa de la Figura 2.4, se presentan las ecuaciones de las salidas de dichas capas. Estas contienen la función de transferencia, la matriz de peso, las entradas y el sesgo correspondiente a cada una de las capas. Finalmente, la red se puede descomponer en la siguiente expresión (Beale et al., 2015).

$$a^3 = f^3(LW^{3,2} f^2 (LW^{2,1} f^1 (IW^{1,1}p + b^1) + b^2) + b^3) \quad 2.1$$

Esta expresión se debe extraer una vez inicializado los pesos y sesgos de la red, encontrado las funciones de transferencias y arquitectura, e incorporando las variables procesadas, todo esto para el fin que se está diseñado la red.

2.1.5 Forma abreviada de la arquitectura de las redes neuronales.

Dado la estructura compleja y además la cantidad de capas que tienen las redes neuronales es conveniente mostrar la arquitectura de forma abreviada. En la Figura 2.5 se muestra un tipo de abreviatura, considerando el ejemplo de red de 3 capas mostrado en la Figura 2.4.

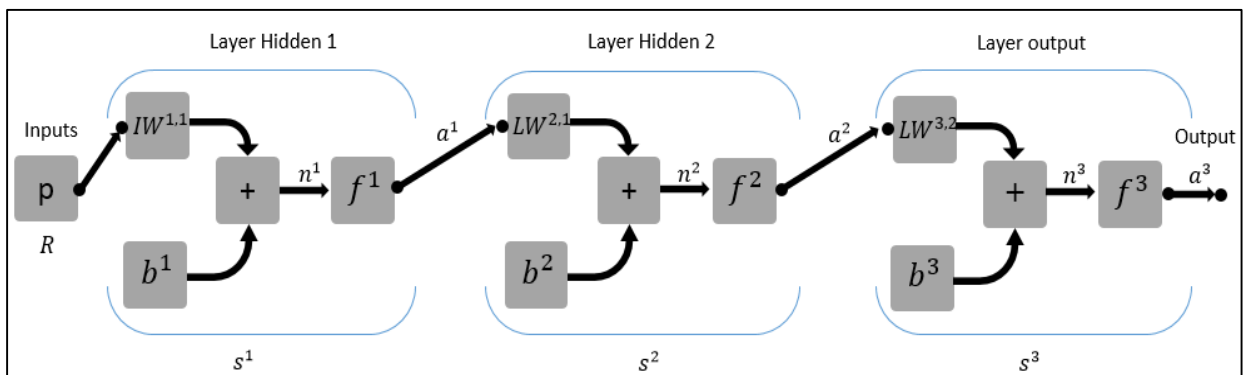


Figura 2.5: Abreviatura del esquema de una ANN (modificado de Beale et al., 2015).

Esto ayuda simplificar el esquema de la red y entrega toda la información relevante de la arquitectura, como son las dimensiones de los vectores y matrices, las funciones que se utilizan, el número de capas, etc (Beale et al., 2015).

2.1.6 Configurar la red neuronal.

La configuración de la red se realiza de acuerdo a la cantidad de datos de las variables de entrada y de destino. Los datos de entrada a la ANN deben ser influyentes en los datos que se desea estimar. Los datos de destino u objetivo son los que se desea estimar. Los datos de salida, son aquellos calculados por la red neuronal, en base a los datos de entrada y de destino.

La configuración de la red consiste en examinar los datos de entrada (variables que influyen en el transporte de sedimento en suspensión) y de destino (concentración sedimento en suspensión observado), establecer los tamaños de entrada y de salida (concentración sedimento en suspensión predicho por la ANN) de la red para que coincida con los datos, y la elección de los ajustes para el procesamiento de las entradas y destinos, y así permitan un mejor rendimiento de la red. La configuración se puede hacer manualmente, pero existen códigos en MATLAB que permiten entrenar y configurar la red al mismo tiempo y que son frecuentemente utilizados.

La configuración entonces, es el proceso de establecer los tamaños de entrada y salida de la red, ajustes de procesamiento de entrada y de destino y la configuración de la inicialización de pesos para que coincida con los datos de entrada y de destino (Beale et al., 2015).

2.1.7 Entrenamiento de redes neuronales.

Las redes neuronales se pueden entrenar de manera incremental (incremental training) y por lotes (batch training). En el entrenamiento incremental los pesos y sesgos se actualizan cada vez que se presenta una entrada a la red. En el entrenamiento por lotes los pesos y sesgos se actualizan sólo después de que se presenten todos los datos de entrada. Según el manual Neuronal Network Toolbox de MATLAB el entrenamiento por lotes suele ser más eficiente, debido a que sus funciones de entrenamiento son más eficientes.

El entrenamiento de redes neuronales, generalmente se divide primero los datos en subconjuntos. El primer subconjunto es el grupo de datos para entrenar, el cual se utiliza para calcular gradientes y las actualizaciones de los pesos y sesgos de la red. El segundo subconjunto es el de validación en el cual se controla el error durante el proceso de formación. El proceso de entrenamiento de la red consiste en afinar los pesos y sesgos para optimizar el rendimiento de la red, según la definición de función de rendimiento.

Cualquier algoritmo de optimización numérica estándar se puede utilizar para optimizar la función de rendimiento. Estos algoritmos son las funciones de entrenamiento las cuales optimizan ya sea el gradiente del rendimiento de la red con respecto a los pesos o el jacobiano de los errores de la red con respecto a los pesos.

El gradiente y el jacobiano se calculan utilizando una técnica llamada el algoritmo backpropagation, que consiste en la realización de cálculos hacia atrás a través de la red.

El objetivo del entrenamiento es encontrar pesos óptimos, esto para que los valores de la salida (concentración de sedimento calculada por la red) se asemejen lo más posible a los datos de destino (datos de concentración de sedimento). Los pesos óptimos se encuentran minimizando el error calculado por la función de rendimiento de la red.

Por ejemplo, si se considera el algoritmo de optimización de descenso de gradiente se actualizan los pesos y sesgos en la dirección en que la función de rendimiento disminuye más rápidamente. Una iteración de este algoritmo se escribe como:

$$w_{k+1} = x_k - a_k * g_k \quad 2.2$$

Donde x_k es el vector de pesos y sesgos actuales, g_k es el gradiente actual y a_k es la tasa de aprendizaje. Esta ecuación se itera hasta que la red converja.

El comando *configure* configura los objetos de la red y también inicializa los pesos y sesgos de la red. Terminado este proceso la red está lista para ser entrenada. Con el comando *train* configura automáticamente la red, inicializa los pesos y sesgos, y además entrena la red.

Si se entrena con el comando *train*, independiente del formato que tenga la entrada, se transforma en vectores concurrentes (columnas de una matriz).

Cada vez que una red neuronal es entrenada, puede resultar en una solución diferente debido a los distintos valores de los pesos y de polarización inicial y por la división de los datos de entrada en conjuntos de entrenamiento, validación o prueba. Como resultado, una red neuronal entrenada en un mismo problema puede dar diferentes salidas para una misma entrada. Es por esto que se debe ejecutar varias veces para encontrar una buena precisión (Beale et al., 2015).

2.1.7.1 Función de entrenamiento.

El algoritmo de Levenberg Marquardt (LM) es el algoritmo de optimización más utilizado. Funciona mejor que el descenso de gradiente simple, y otros métodos de gradiente conjugado, en una amplia variedad de problemas. El método Levenberg Marquardt funciona como el método de descenso rápido (Steepest Decent) cuando los parámetros que está midiendo están lejos de los valores óptimos, y tiende a funcionar como el método de Gauss-Newton cuando los parámetros están cerca de los valores óptimos (Hise et al., 2012). El vector de gradientes se representa como

$$g = \begin{pmatrix} \frac{\partial E}{\partial W_1} \\ \frac{\partial E}{\partial W_2} \\ \vdots \\ \frac{\partial E}{\partial W_R} \end{pmatrix} \quad 2.3$$

Donde E es el error de la red para la estimación que se está realizando y W se refiere a los pesos de la red. El jacobiano es esencialmente cada gradiente para cada una de las estimaciones de entrenamiento y salida de la red. El jacobiano se muestra a continuación.

$$J = \begin{pmatrix} \frac{\partial E_1}{\partial W_1} & \dots & \frac{\partial E_1}{\partial W_R} \\ \vdots & \ddots & \vdots \\ \frac{\partial E_m}{\partial W_1} & \dots & \frac{\partial E_m}{\partial W_R} \end{pmatrix} \quad 2.4$$

Donde R es el número de pesos y m es el número de estimaciones. Una vez se calcula el jacobiano, el algoritmo de Levenberg Marquardt (LM) se puede representar por la siguiente ecuación.

$$W_{k+1} = W_k - (J_k^T J_k + \mu I)^{-1} J_k^T E \quad 2.5$$

Donde E es el error total para todas las estimaciones, I es la matriz identidad y μ es un parámetro de aprendizaje. Entonces el parámetro μ se ajusta varias veces en cada iteración y el resultado con mayor reducción del error se selecciona. Cuando error es muy grande este método se comporta como el método de gradiente rápido y cuando μ es igual a cero se comporta como el método de Newton, el cual es más rápido y exacto cuando el error es cercano a cero. De esta manera el resultado de la función de rendimiento se reducirá siempre con cada iteración.

2.1.7.2 Función de rendimiento de la ANN.

Para medir el rendimiento de la red se comparan los datos de salidas (concentración de sedimento en suspensión estimada) con los datos de destino (concentración de sedimento en suspensión reales) en el proceso de entrenamiento y validación. En estos procesos de actualización de pesos y sesgos es común utilizar en redes neuronales el error cuadrático medio (mse) u otro parámetro estadístico para medir el error que se genera al comparar los datos de salida y de destino de la red. El proceso de entrenamiento finaliza cuando se

encuentran los pesos y sesgos que generan el mínimo error de la función de rendimiento (mse). Luego de esto se almacenan dichos pesos y sesgos, los cuales son utilizados para realizar pruebas con datos diferentes a los que se utilizaron en el entrenamiento y validación. Las salidas de los datos de prueba se comparan con sus datos de destinos utilizando estadísticos que no necesariamente deben ser el mismo de la función de rendimiento. Esto se realiza para analizar la precisión de las predicciones de la red producto del proceso de entrenamiento y validación. Algunos parámetros estadísticos utilizados en las ANN pues estos procesos son:

Error cuadrado medio (MSE): Este estadístico es utilizado en el proceso de entrenamiento y validación de la red, en donde calcula el error de los datos de salida y de destino, para que por medio de la función de entrenamiento (LM) modifique los pesos y sesgos, de forma tal que ésta reduzca el error de los datos de salida y de destino. A continuación, se expone la fórmula del error cuadrático medio.

$$MSE = \frac{\sum_{i=1}^n (P_i - O_i)^2}{N} \quad 2.6$$

Donde P es el parámetro predicho, O es el parámetro observado, \bar{P} es la media del parámetro predicho, \bar{O} es la media del parámetro observado y N el número total de datos.

Percent Bias (Pbias): Este estadístico es utilizado en tres procesos. Primero se utiliza en un ciclo creado en el programa MATLAB para encontrar una de las tres pruebas de cada combinación de entradas en la ANN. Una vez encontrada las tres soluciones para cada combinación de variables, el PBIAS junto al RMSE son los que determinan cual prueba es la que genera estimaciones más acertados de C_{ss}. Finalmente, estos dos estadísticos se utilizan para comparar la prueba seleccionado de la ANN con las estimaciones de la CGS. A continuación, se expone la fórmula del estadístico percent bias.

$$Pbias = \frac{\sum_{i=1}^n (O_i - P_i) * 100}{\sum_{i=1}^n (O_i)} \quad 2.7$$

Raíz del error cuadrático medio (RSME): Es utilizado para identificar cual de todas las pruebas realizadas con las combinaciones de entrada de la ANN es la más acertada en la estimación de C_{ss} . Del mismo modo este estadístico sirve para discernir cuál de los dos modelos utilizados en este estudio para la determinación de la C_{ss} es el que entrega estimaciones más acertadas.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{N}} \quad 2.8$$

2.1.8 Aplicación de ANN en el transporte de sedimento.

Experimentos recientes han informado que la técnica de ANN es una alternativa para la estimación de sedimento suspendido (*Abrahart y White, 2001; Jain, 2001; Tayfur, 2002; Cigizoglu, 2004; Kisi, 2004; Cigizoglu y Kisi, 2005; Cobaner et al, 2009; Melesse et al., 2011; Kia et al., 2013; Barzegari et al, 2015*). Abrahart y White (2001) aplicaron las ANN para predecir el flujo de sedimentos que ocurren en diferentes tipos de agricultura y uso de suelo. Predijeron cargas de sedimento en suspensión a partir de datos de precipitaciones, datos de escorrentía, coeficiente de escorrentía y los parámetros que representan el estado del uso del suelo. Los resultados indicaron que las ANN proporcionan un mejor ajuste a los datos que el método de regresión lineal múltiple. Jain (2001) utilizó una ANN simple para establecer la relación sedimento-caudal y encontró que el modelo ANN genera resultados más acertados que la curva de gastos de sedimentos. Tayfur (2002) utilizó ANN para simular los flujos de sedimentos observados experimentalmente desde diferentes pendientes de laderas en diversas intensidades de lluvia. El estudio indicó que el rendimiento de la ANN sólo con la información pendiente y precipitaciones como entradas podría ser tan buena como la de un modelo de base física con mucho más variables, tales como la velocidad del flujo, tasa de infiltración, esfuerzo de corte y energía de la corriente. Cigizoglu (2004) investigó la exactitud de una ANN simple en la estimación y predicción de los datos de sedimentos en suspensión diaria. Su estudio concluyó que las ANN con arquitectura MLP pueden representar el comportamiento no lineal del sedimento relativamente mejor que los modelos convencionales. Kisi (2004) utilizó diferentes técnicas ANN para predecir y estimar la concentración de sedimentos en suspensión diario e indicó que los modelos multicapa se

desempeñaron mejor que las redes neuronales de regresión generalizada y redes de funciones de base radial. Cigizoglu y Kisi (2005) desarrollaron métodos para mejorar el rendimiento ANN en la estimación de los sedimentos en suspensión. Cobaner et al. (2009) modela la concentración de sedimentos en suspensión a partir de datos hidrometeorológicos. Melesse et al. (2011) estima que la carga de sedimentos suspendidos por tres ríos principales (Mississippi, Missouri y Río Grande) en EE.UU. utilizando enfoque de modelado perceptón de multicapas de redes neuronales artificiales (MLP). Kia et al. (2013) utilizó diferentes redes neuronales para estimar el sedimento en suspensión en el río Babolroud. Con los resultados determinó que las ANN estructuradas MLP mostraron mejores rendimientos que los otros modelos de redes utilizados y que el método de la curva de gasto de sedimento en suspensión. Barzegari et al. (2015) estimó el sedimento en suspensión por el modelo ANN, DT (decisión trees) con el algoritmo CART y diferentes tipos de CGS, en diez estaciones de la provincia de Lorestan de Irán. Los resultados mostraron que la exactitud de la ANN con arquitecturas MLP tienen mejor rendimiento que los otros dos métodos utilizados.

Los autores Maier y Dandy (2000) describen que, en los 43 artículos analizados de la predicción y pronóstico de recursos hídricos con el método de las ANN, la mayoría utiliza el procesamiento backprogration con el flujo de información del tipo feedforward (*Maier y Dandy, 2000*). Estos procesos son descritos en el desarrollo de este documento, y fueron la base para la elaboración de la red neuronal artificial que se utilizó.

En consecuencia, una de las estructuras de red más utilizada y que genera una estimación acertada en la carga o concentración de sedimento en suspensión, es aquella que utiliza una capa oculta y otra de salida, donde sus funciones de transferencias por lo general son del tipo sigmoidea y lineal. El flujo de información en esta red neuronal es en hacia adelante (redes feedforward). El proceso de entrenamiento se realiza por medio de algoritmos de optimización de la función de rendimiento, con métodos de retro-programación (backprogration). Para esto se debe definir la función de rendimiento de la red, la cual calcula el error de las salidas de la ANN (sedimento en suspensión predicha por la red) con los datos observados (sedimento en suspensión provenientes del río). Mediante una función o algoritmo se modifican los pesos con el objetivo de minimizar el error calculado por la

función de rendimiento. Esto se realiza hasta que el error converja a un mínimo y los pesos sean los óptimos.

Toda programación de la red neuronal artificial que se desarrolló en esta tesis se realizó en el software MATLAB. Los datos considerados para el análisis fueron pre-procesados y post-procesados en EXCEL de manera de facilitar la organización y definición de los formatos respectivos.

2.2 Variables del modelo.

En este estudio se utilizaron combinaciones de las variables de altura media diaria, caudal medio diario y precipitación media diaria, para determinar con el modelo de ANN la C_{ss}. Esto debido a que se sabe de investigaciones realizadas por otros autores que el sedimento en suspensión tiene una estrecha relación con dichas variables que se miden en el mismo día y en el anterior al que se toma la muestra de sedimento en suspensión (*Cobaner et al, 2009*).

El modelo de ANN, como ya se ha mencionado, identifica patrones que existen entre las variables de entrada y de salida. En la realidad existen diferentes valores de concentraciones de sedimento en suspensión en periodos donde el caudal es constante y no precipite, lo cual genera patrones erróneos entre las variables de entrada y destino de la red. También se conoce que la correlación entre las variables seleccionadas en este estudio y la C_{ss} es mayor durante eventos de crecidas. Por lo tanto, hace sentido separar los caudales en eventos de crecidas y analizar cada una de ellas de manera independiente, de esta forma se selecciona la combinación de parámetros que mejor representen el conjunto de eventos.

3 MÉTODOS

3.1 Rede neuronal de Multicapas Feedforward (MLP) utilizada.

Este tipo de red es una de las más utilizadas en la estimación de sedimento en suspensión como también de fondo. Es por esto que en el presente estudio se tomará como base esta estructura para estimar de forma acertada el sedimento en suspensión.

Estas redes se caracterizan porque las neuronas de cada nivel solo están conectadas con las neuronas de los niveles posteriores, por lo que la información solo se propaga hacia adelante.

Las redes multicapas Feedforward están compuestas por dos tipos de capas diferentes de neuronas, las capas ocultas y la capa de salida. Se contempla utilizar una capa oculta y una de salida.

La capa oculta es aquella que está en contacto directo con la información de entrada (input). Las neuronas de esta capa están interconectadas con los datos de entrada por medio de pesos produciendo así entradas ponderadas que se suman con los sesgos para formar las entradas netas. Estas entradas netas pasan por la función de transferencia del tipo sigmoide para producir la salida de esta capa. La función de transferencia puede ser modificada por cualquier tipo de función (*Beale et al., 2015*).

A continuación, se muestra la expresión de la función de transferencia del tipo sigmoide utilizada.

$$f(x) = \frac{2}{1 + e^{-2*x}} - 1 \quad 3.1$$

La capa de salida está totalmente interconectada con la capa oculta, donde hay un peso asociado a cada una de las conexiones. Estos pesos junto con las salidas de la capa oculta producen las entradas ponderadas, las cuales se les suman el sesgo al igual que en la capa oculta. Finalmente, las entradas netas de esta capa pasan por la función de transferencia, la cual es del tipo lineal, generando así la salida de la red. En la Figura 3.1 se esquematiza la estructura de la red utilizada.

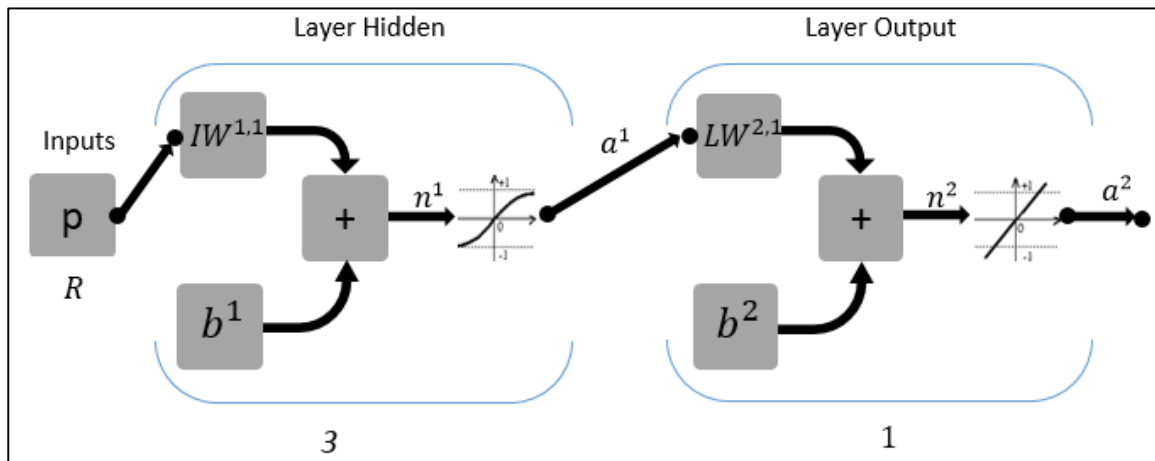


Figura 3.1: Arquitectura de la Red Multicapas Feedforward utilizada en este estudio (modificado de *Beale et al., 2015*).

Los números tres y uno que se muestran en la Figura 3.1, debajo de las capas oculta y de salida respectivamente, representan la cantidad de neuronas en dichas capas. Este número de neuronas se determinó por ensayo y error al igual que el número de capas en la red.

Se describirá más adelante el número de entradas y que variables son las que se tomarán para la estimación acertada la concentración de sedimento en suspensión.

El proceso de aprendizaje consiste en minimizar (respecto a los pesos) la función de rendimiento para todos los patrones del conjunto de entrenamiento, los cuales se van presentando en la red en iteraciones sucesivas hasta lograr la convergencia de la función de error.

Para cada iteración los pesos de la red se irán modificando y modificarán las entradas y salidas de las neuronas y el error de la siguiente iteración.

La función de rendimiento que utiliza esta red para entrenar es la MSE la cual es representada por la ecuación 2.6 mostrada anteriormente y la función o algoritmo de entrenamiento es la LM mostrada en la ecuación 2.5. El procedimiento que realiza la red es el siguiente:

- 1) Inicializa los pesos y sesgos de la red.
- 2) Normaliza los datos de entrada y de destino de la red
- 3) Calcula la salida mediante la arquitectura de la red.
- 4) Calcula el error de la salida y del destino de la red mediante la función de rendimiento.
- 5) Actualiza los pesos de la red según la función de entrenamiento, la cual toma como parámetro el error producido por la salida y el destino en función de los pesos.
- 6) Repite el paso 3, luego el 4 y el 5 en la siguiente iteración. Estas iteraciones continúan hasta que la función de rendimiento (MSE) converja hacia un mínimo, es decir, el error calculado en el 4) alcance un mínimo.
- 7) Aplica el proceso inverso a la normalización de los datos de salida para que las unidades coincidan con las de destino.
- 8) Probar la red neuronal encontrada en el punto 6) con conjuntos de datos diferentes a los utilizados en el entrenamiento y validación de la red.
- 9) Medir la precisión de la predicción de la red neuronal.

Si los datos son correctos, las variables de entradas seleccionadas son representativas del fenómeno que se está estudiando y la estructura de la red utilizada es adecuada, se lograra obtener resultados satisfactorios, por lo tanto, la red entregara salidas de datos calculados que serán similares a los datos observados de sedimento.

Al finalizar el proceso de iteraciones se evaluará el rendimiento de la red por medio de parámetros estadísticos, con los cuales se hará una comparación entre los datos de salida final de la red y los datos observados de sedimento en suspensión, para ver que tanto se asemejan.

Con las modificaciones necesarias a las entradas (input) y a la arquitectura de la red neuronal se logra una estimación adecuada de la concentración de sedimento en suspensión. En este punto es donde se debe descomponer la arquitectura de la red neuronal para obtener una formulación matemática que estime de forma correcta la concentración de sedimento en suspensión y además esta tenga una interpretación física debido a que las variables ingresadas en el modelo estarán en la formulación final.

3.2 Zona de estudio.

La zona de estudio para este proyecto se encuentra en la región de la Araucanía, específicamente se analizó el río Cautín al sur de la localidad de Rari-Ruca, ya que este río presenta una amplia gama de concentraciones de sedimento en suspensión que varía entre 0 a 50 mg/l.

En la Figura 3.2 se muestra la ubicación de donde se analizó el río mencionado anteriormente, tomando en cuenta que las estaciones de sedimento y meteorológica que se utilizaron para obtener los datos del modelo se encuentran en los alrededores del punto indicado.

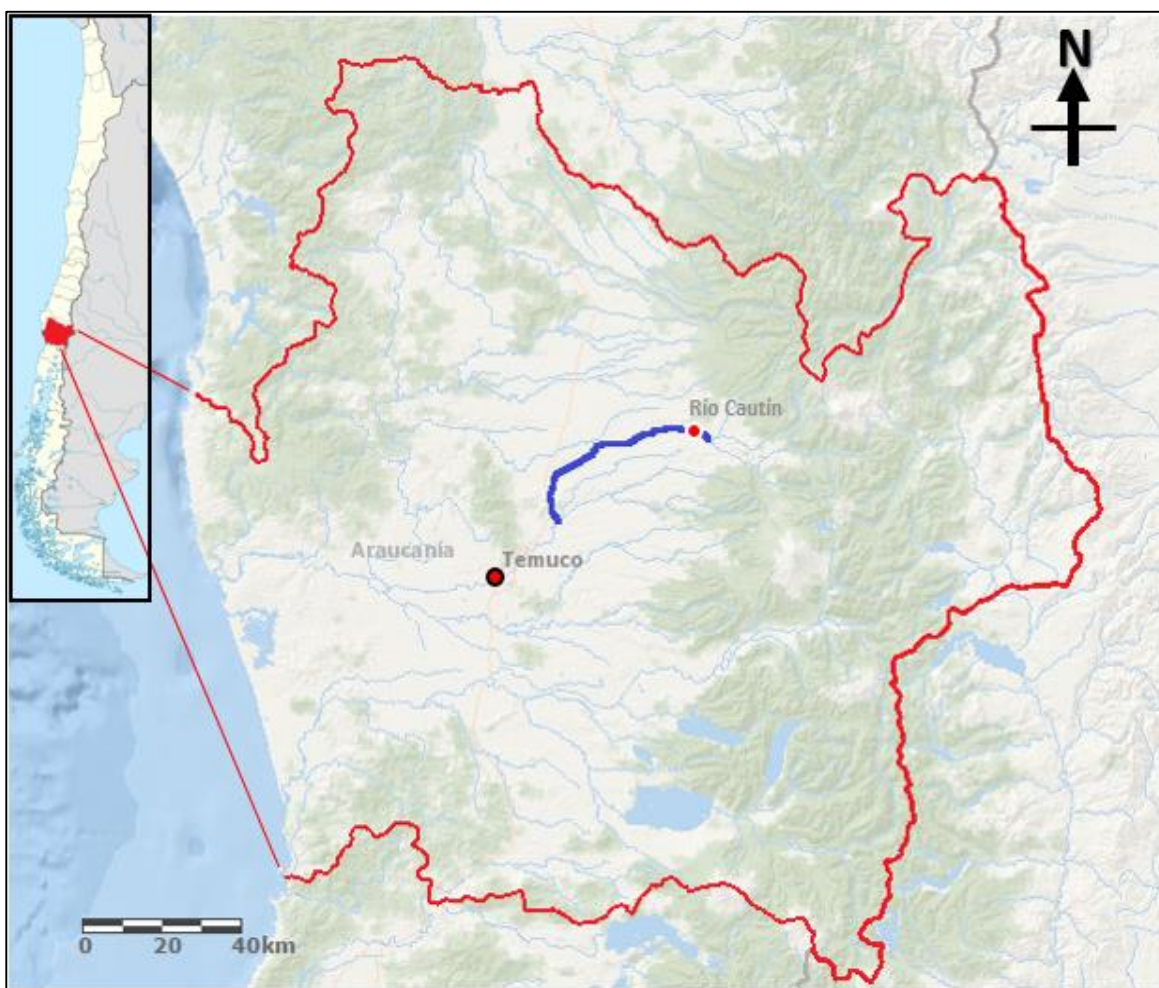


Figura 3.2: Área de estudio del proyecto (Red Hidrométrica).

3.2.1 Descripción general del río.

Río Cautín: Es el principal afluente del río Imperial en atención a su caudal, tiene su nacimiento en la falda occidental de la cordillera de Las Raíces y atraviesa todo el llano con una longitud de 174 km, que sumado al río Imperial completan 230 km. La superficie drenada por el Cautín asciende a 3.100 km².

El más importante tributario del Cautín es el río Quepe, que nace en la laguna homónima en la falda occidental del volcán Llaima y se junta al Cautín cerca de la localidad de Almagro. Este tributario tiene una orientación sensiblemente al W y longitud de 112 km a través de la depresión intermedia y lo integran una gran cantidad de esteros, principalmente por su ribera sur.

El régimen hidrológico de ríos de la Región de la Araucanía es principalmente del tipo pluvial. Esta característica está asociada a que la cordillera de los Andes es normalmente baja en altura en esta región, por lo que no se acumula una cantidad significativa de nieve (*MOP-DGA, 2004*).

3.2.2 Clima.

La cuenca del río Imperial, presenta dos tipos climáticos, el Templado cálido lluvioso con influencia mediterránea (en el sector centro y bajo de la cuenca) y el Clima templado frío lluvioso con influencia mediterránea (sector precordillerano de la cuenca).

Clima templado cálido lluvioso con influencia mediterránea: Este tipo climático se encuentra en la IX y X regiones, desde la cuenca del río Cautín hasta el norte de la ciudad de Puerto Montt. Se caracteriza por presentar precipitaciones a lo largo de todo el año, aunque los meses de verano presentan menor pluviosidad que los meses invernales, en Temuco se registran anualmente 1.345mm. Las temperaturas no sufren una gran variación por latitud, siendo la

unidad térmica y lo poco significativo de las oscilaciones, una notable característica de este clima.

Clima templado frío lluvioso con influencia mediterránea: este tipo climático se presenta en la zona cordillerana de las regiones VIII, IX y sector norte de la X región. (*MOP-DGA, 2004*).

3.2.3 Uso de Suelo.

Río Cautín: La cuenca definida aguas arriba de la estación fluviométrica “Río Cautín en Rari-Ruca” está situada al norte de la IX Región a pocos kilómetros al este de la ciudad de Vilcún. La cuenca se caracteriza por su ubicación a los pies de la Cordillera de los Andes Posee un 51% de bosques, un 7 % de su área esta desprovista de vegetación y el 6 % aproximadamente contiene glaciares y nieve. En la cuenca hay 11 bocatomas, 2 embalses, y 56,8 km de canales. Las ciudades de Curacautín y Malalcahuello se ubican en esta cuenca. Las rocas de esta zona son mayoritariamente mixtas con una composición de 77% proveniente de la clase de secuencias volcánicas y un 17% proveniente de la clase de secuencias volcano-sedimentarias. un 7% de secuencias sedimentarias (*MOP-DGA, 2011*).

3.3 Datos para el modelo

Para producir predicciones precisas de la concentración de sedimento en suspensión el modelo de redes neuronales artificiales necesita un set de datos de entrada como materia prima. Este set de datos debe contener las variables que más influyen o representan el proceso en estudio, de lo contrario la estimación realizada por la red no será la correcta. Además de esto es necesario que dichas variables sean pre-procesadas para que la red pueda identificar un patrón entre las entradas de la red y datos observados (objetivos), teniendo como resultado del entrenamiento y ajuste de los pesos, una salida similar a los datos observados.

En la estimación de sedimento diversos autores han realizado pruebas con parámetros adimensionales como también variables medibles in situ para la estimación correcta de sedimento en suspensión.

Zhu et al (2007) por ejemplo, identificó la precipitación como una de las más influyentes en el transporte de sedimento. Además de esta variable el caudal total del curso fluvial es ingresado como entrada, ya que está estrechamente relacionado con el transporte de sedimento en un río, tanto de fondo como en suspensión. Autores como Kakaei et al (2013) utilizan como variables de entrada para la estimación de sedimento sólo la precipitación y el caudal, obteniendo resultados satisfactorios.

En el presente estudio se utiliza como variables de entrada a la red neuronal combinaciones de la precipitación diaria, el caudal total y la altura del curso fluvial. Como datos objetivos de la red se considera la concentración de sedimento en suspensión observada en el río. Las combinaciones de las variables de entrada del modelo son consideradas en el tiempo “t” y el tiempo “t-1”, teniendo en cuenta que el tiempo “t” corresponde al día que se midió la concentración de sedimento y el tiempo “t-1” al día anterior.

3.3.1 Estaciones y datos para el modelo.

En el presente trabajo se tomaron los reportes de sedimento y meteorológicos de la Dirección General de Aguas (DGA) como la base de datos para la predicción del sedimento en suspensión del río Cautín, utilizando el modelo de redes neuronales artificiales. La estación de sedimento y meteorológica seleccionada para el río Cautín están a una distancia aproximada de 50 m. Las estaciones de sedimento y meteorológica son “RIO CAUTIN EN RARI-RUCA” (Código BNA: 09123002-k; Altitud: 440 msnm; Latitud S:38°25’30”; Longitud W:72°00’38”) y l “RIO CAUTIN EN RARI-RUCA” (Código BNA: 09123001-1; Altitud: 425 msnm; Latitud S:38°25’49”; Longitud W:72°00’38”) respectivamente.

En la Figura 4.6 se muestran las ubicaciones de las estaciones de sedimento y meteorológicas descritas anteriormente.

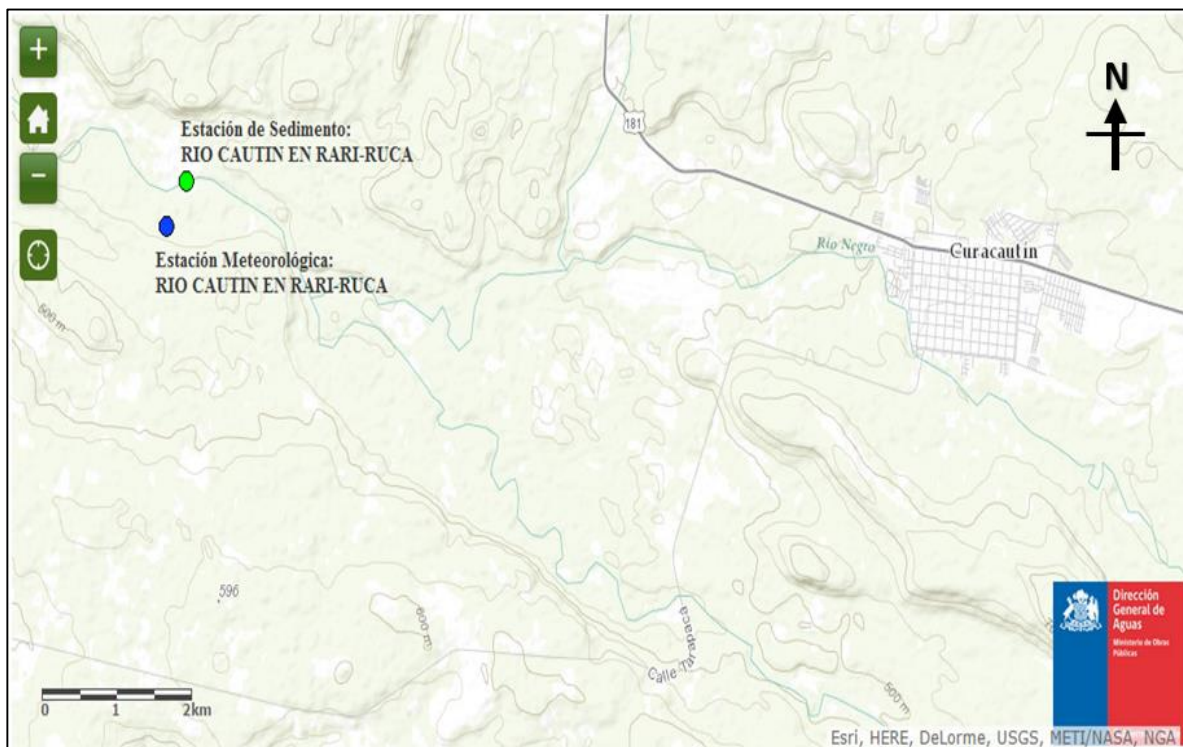


Figura 3.3: Estación de Sedimento y Meteorológica.

Se descargaron los registros de datos de estas estaciones desde el año 2001 hasta el año 2014. Las variables de interés para el modelo de ANN fueron la altura y el caudal medio diario del curso fluvial, la precipitación diaria y la concentración de sedimento en suspensión.

3.3.2 Métodos de Muestreos y obtención de datos.

Los métodos utilizados por la DGA para obtener los datos de las variables altura y caudal medio diario, precipitación diaria y concentración de sedimento en suspensión rutinaria (diaria), disponibles en el sitio web y con los cuales se trabajó en el presente estudio son los siguientes:

- Caudal y altura. Transductores de presión: Son sensores electrónicos que convierten la presión que registran (que es equivalente a una cierta altura de agua) en una señal eléctrica que es guardada en un Data Logger (DL). Estos equipos son programados antes de su instalación de manera que realizan el cálculo directamente para convertir la señal de eléctrica en un dato de altura. Con la altura de escurrimiento de la sección de control, se calcula el caudal utilizando la curva de descarga, característica de la estación.
- Precipitación. Registro por plataforma Satelital (estación T12): La plataforma registra los datos de forma continua obtenidos por un pluviómetro y permite la transmisión de datos a una central entregando información prácticamente en tiempo real. Los tipos más comunes de pluviómetros son el de pesa y el basculante. Los pluviómetros de pesada operan capturando el agua en un sistema de recolección y pesando el peso de esa agua. La precipitación, expresada como altura de agua, se mide en forma continua en pluviógrafos o nivógrafos y en forma esporádica en pluviómetros. Los instrumentos son básicamente recipientes estandarizados que recogen la precipitación y la miden por unidad de área en términos de altura. Aquellos que entregan un registro continuo están equipados con un sistema que va indicando en forma mecánica o electrónica la precipitación acumulada en función del tiempo.
- Concentración de sedimento en suspensión. Muestreo en profundidad. Muestro de llenado continuo de la botella en toda la vertical, en su eje de descenso y ascenso. Se muestrea en verticales elegidas de tal forma que entre ellas no escurra más del 10% del caudal total. En cada vertical se medirá la velocidad media del flujo y la profundidad. En cada vertical se tomará 1 muestra. Las curvas de tarado del

muestreador deben dar, para un cierto punto de la sección y para una determinada velocidad y nivel del río, el tiempo de llenado de la botella. Este tiempo es además función del tipo de boquilla que se emplea. El tipo de boquilla debe elegirse de modo que el ascenso o descenso del muestreador se haga entre los siguientes límites: 1 metro de descenso o ascenso cada 5 segundos a 1 metro de descenso o ascenso cada 10 segundos.

3.3.3 Análisis de datos para el modelo.

Debido a que el río analizado para este trabajo tiene un régimen hidrológico del tipo pluvial, se analizó el hidrograma del río y se seleccionaron los sets de datos de las variables de entradas y objetivos donde el caudal crece. Estas crecidas de caudal se pueden representar en un hidrograma tomando una tormenta aislada, como la mostrada en la Figura 3.4.

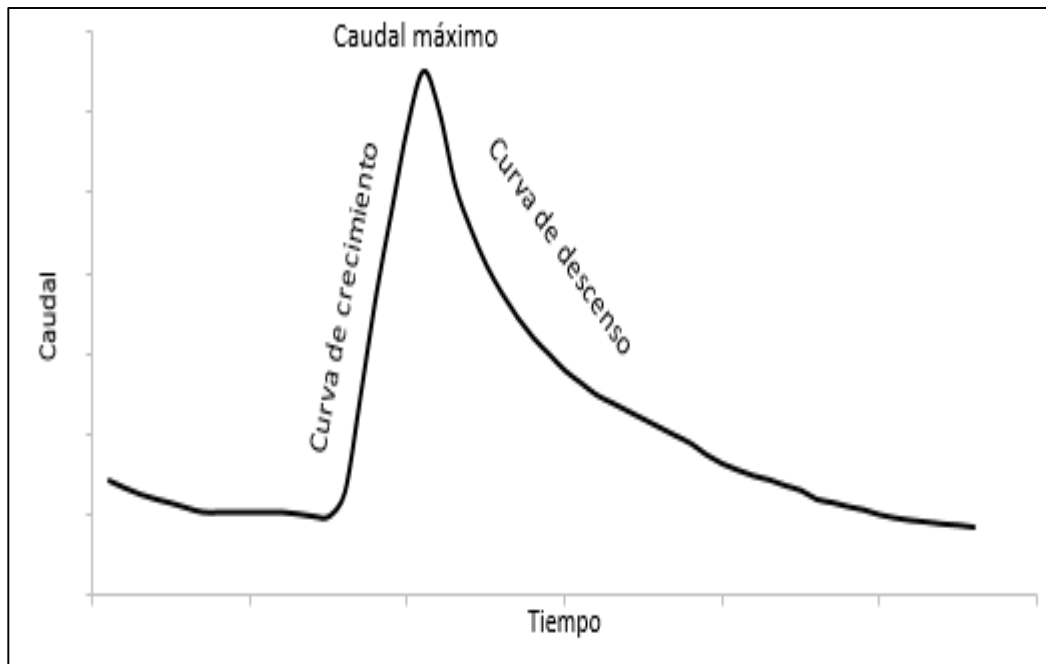


Figura 3.4: Hidrograma de una tormenta aislada.

Estos datos fueron utilizados en el modelo de redes neuronales para la predicción de sedimento en suspensión.

Los datos de la curva de crecimiento, caudal máximo y/o curva de descenso fueron entonces los insumos para que la red neuronal estimara la concentración diaria de sedimento en suspensión. En la Figura 3.5 se muestra el hidrograma de la estación “RIO CAUTIN EN RARI-RUCA” en el periodo que se analizaron los datos, y de forma achurada se destacan los periodos donde se encuentran los datos seleccionados (crecidas) para ser procesados por el modelo de redes neuronales.

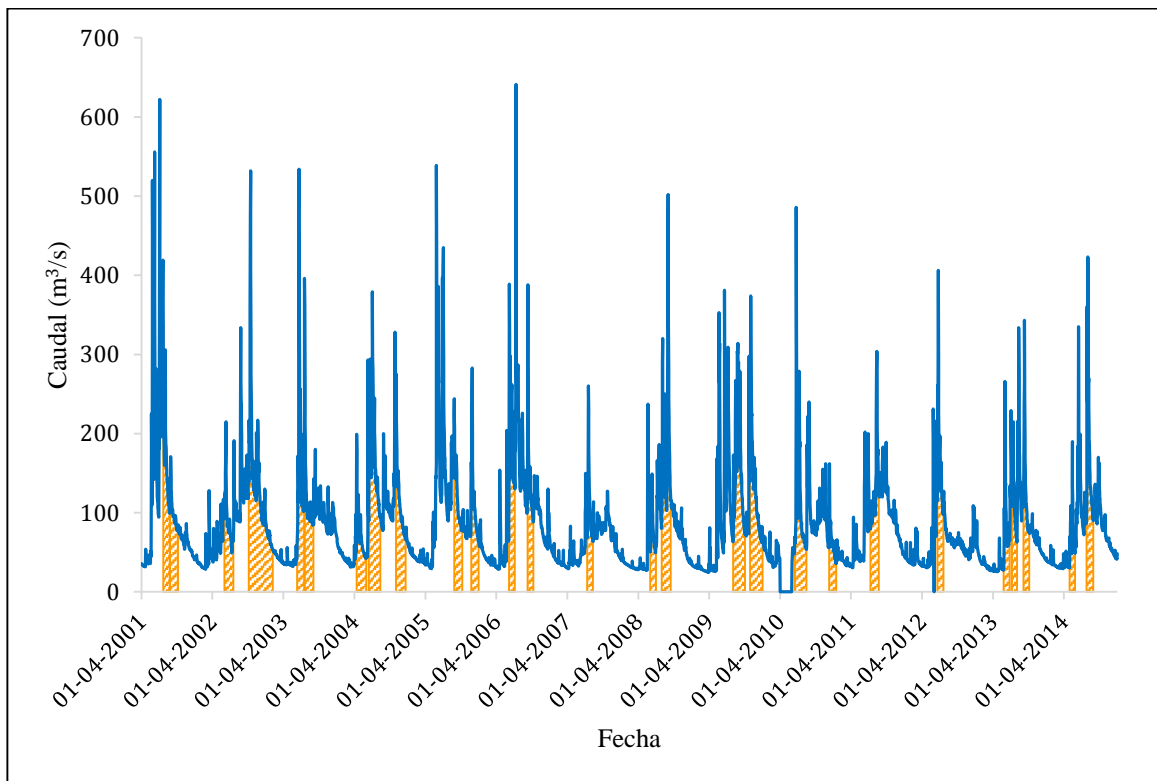


Figura 3.5: Hidrograma de la estación “Río Cautín en Rari-Ruca”.

Los datos deben proporcionar información necesaria para que la ANN genere patrones para la determinación correcta de la C_{ss} . Debido a que existen periodos en donde la C_{ss} varía, sin un cambio en el caudal y altura y teniendo valores de precipitación igual a cero, se eliminaron estos datos, ya que dicha variación de C_{ss} puede deberse a otra variable que no se consideró en este trabajo y lo cual generaría patrones erróneos entre las variables de entrada y objetivo de la ANN. También se tomó en cuenta la cantidad de días que duran las crecidas, seleccionando en la mayoría de los casos crecidas que tienen más de 30 días de duración.

Se analizó una cantidad de 27 crecidas, donde cada una de éstas define las variables de entrada al modelo de redes neuronales, las cuales son la altura, el caudal y la precipitación en el tiempo “t” y “t-1”, y además contienen los datos objetivos, correspondientes a la concentración de sedimento en suspensión en el tiempo “t”. en la Tabla 3.1, se presenta las combinaciones de las variables de entrada consideradas en el modelo de ANN.

Tabla 3.1: Combinaciones de entradas para la ANN.

Combinación	Entradas	Objetivo	Salida ANN
COMB1	H(t) Q(t) P(t)	Css(t)	Css'(t)
COMB2	H(t) Q(t) P(t) H(t-1)	Css(t)	Css'(t)
COMB3	H(t) Q(t) P(t) Q(t-1)	Css(t)	Css'(t)
COMB4	H(t) Q(t) P(t) P(t-1)	Css(t)	Css'(t)
COMB5	H(t) Q(t) P(t) H(t-1) Q(t-1)	Css(t)	Css'(t)
COMB6	H(t) Q(t) P(t) H(t-1) P(t-1)	Css(t)	Css'(t)
COMB7	H(t) Q(t) P(t) Q(t-1) P(t-1)	Css(t)	Css'(t)
COMB8	H(t) Q(t) P(t) H(t-1) Q(t-1) P(t-1)	Css(t)	Css'(t)

Se observó que, para el río seleccionado, aproximadamente un 90% de los datos de concentración de sedimento en suspensión son menores a 50 mg/l. El 10% restante de los datos se considera como datos outlier, producto de eventos puntuales que generan un aumento excesivo de concentración de sedimento. Por lo tanto, para este trabajo solo se consideró los grupos de datos que contenían concentraciones de sedimento menores a 50 mg/l. En la Tabla 3.2 se muestran la cantidad de datos según rangos de concentración de sedimento en mg/l y además el porcentaje con respecto a la cantidad total de datos de sedimento (del año 2001 hasta el año 2014).

Tabla 3.2: Cantidad de datos y porcentajes de concentraciones de sedimento.

Nombre del Río	Concentraciones de sedimentos C_{ss} (mg/l)					
	Numero de datos / Porcentaje					
	C _s < 50	50<C _s <100	100<C _s <150	150<C _s <200	200<C _s <250	C _s >250
Cautín	4536 / 87.9%	405 / 7.8%	65 / 1.2%	33 / 0.6%	32 / 0.6%	85 / 1.6%

Debido a que se están estimando concentraciones de sedimento en suspensión menores a 50 mg/l y la ANN no tiene la capacidad de extrapolar con exactitud datos que no estén dentro de los rangos que se utilizaron para entrenarla, se podrá extraer una expresión matemática que estime de manera acertada la concentración de sedimento en suspensión menor a 50 mg/l. En la Tabla 3.4 se muestran los periodos de las variables utilizadas para las 27 crecidas seleccionadas según los criterios mencionados anteriormente.

Tabla 3.3: Periodos de crecidas para el modelo de ANN.

Crecida	Fecha		Altura (m)		Caudal (m ³ /s)		Precipitación (mm)		Concentración (mg/l)	
	Inicio	Termino	Prom	Max	Prom	Max	Prom	Max	Prom	Max
1	24-08-2001	05-10-2001	1.39	1.82	100.28	171.00	1.41	17.00	14.72	61.84
2	02-06-2002	17-07-2002	1.29	2.05	90.27	215.00	4.25	33.00	18.59	256.12
3	04-10-2002	12-11-2002	1.88	3.18	193.28	532.00	8.97	48.00	52.00	378.90
4	12-11-2002	24-12-2002	1.54	2.00	124.93	217.00	4.46	45.00	15.50	143.16
5	24-12-2002	05-02-2003	1.17	1.59	74.58	130.00	1.35	25.00	7.01	36.50
6	10-06-2003	16-07-2003	1.76	3.05	174.61	534.00	10.80	62.00	53.30	477.78
7	17-07-2003	03-09-2003	1.52	2.64	124.69	396.00	3.51	23.50	15.73	171.89
8	23-11-2005	31-12-2005	1.50	2.31	115.64	283.00	8.23	118.00	16.21	103.52
9	06-09-2006	09-10-2006	1.63	2.65	138.29	388.00	12.96	82.50	23.24	295.71
10	06-07-2006	02-08-2006	2.00	2.80	221.92	425.00	18.12	78.50	43.90	130.94
11	12-07-2007	11-08-2007	1.49	2.15	105.88	260.00	17.56	63.50	24.43	210.18
12	01-06-2008	01-07-2008	1.33	1.74	81.18	149.00	17.11	98.00	24.32	71.27
13	28-07-2008	25-08-2008	1.77	2.40	160.54	320.00	21.12	66.50	37.39	187.35
14	26-08-2008	29-10-2008	1.59	3.07	138.45	502.00	9.63	108.50	26.23	409.07
15	19-05-2009	13-06-2009	1.49	2.45	125.67	353.00	18.85	115.00	25.33	88.52
16	26-06-2009	29-07-2009	1.62	2.29	145.46	309.00	10.66	44.50	37.63	213.80
17	04-09-2009	01-10-2009	1.57	2.17	132.62	279.00	6.63	71.50	15.84	68.41
18	26-10-2009	01-12-2009	1.75	2.54	174.03	374.00	12.07	83.00	20.09	114.52
19	06-07-2010	14-08-2010	1.55	2.38	96.40	279.00	8.91	71.00	12.99	146.80
20	30-05-2010	14-07-2010	1.57	2.38	98.87	279.00	9.56	71.00	17.43	146.80
21	10-07-2011	22-08-2011	1.81	2.52	135.71	304.00	16.02	58.00	9.99	46.43
22	19-06-2012	19-07-2012	2.01	2.90	167.06	406.00	18.23	68.50	36.62	212.97
23	26-05-2013	26-06-2013	1.59	2.41	90.73	266.00	20.05	97.50	20.75	191.30
24	08-07-2013	02-08-2013	1.74	2.22	112.31	215.00	12.94	98.50	15.00	79.95
25	04-09-2013	04-10-2013	1.82	2.69	132.02	343.00	8.58	44.00	18.41	215.81
26	29-04-2014	27-05-2014	1.42	2.12	71.56	190.00	13.66	94.00	15.94	111.50
27	26-07-2014	30-08-2014	2.05	3.01	183.86	423.00	15.71	108.50	41.29	391.67

3.3.4 Lectura de datos.

Los datos son incorporados al programa MATLAB desde archivos block de notas, donde la entrada y objetivo de la red son ordenados de manera que todas las variables correspondientes a un día estén dispuestas en una columna, por lo tanto, el número de columnas corresponde a la cantidad de set de datos o días que se están analizando. El número de filas de la entrada del modelo corresponde a la cantidad de variables que se está utilizando para determinar la concentración de sedimento en suspensión, lo cual para este trabajo y considerando las combinaciones de la Tabla 3.1 serán tres filas si se considera la COMB1; cuatro filas si se consideran las COMB2, COMB3 y COMB4; cinco filas si se consideran las COMB5, COMB6 y COMB7; seis filas si se considera la COM8.

Teniendo en cuenta la arquitectura de la red neuronal de la Figura 3.1 y considerando las variables de entradas de la COMB1, se ejemplifica el proceso dentro de la ANN para obtener el resultado. Las entradas se expresan de la siguiente forma:

$$P = \begin{pmatrix} H_{1,1} & \cdots & H_{1,i} \\ Q_{2,1} & \cdots & Q_{2,i} \\ P_{3,1} & \cdots & P_{3,i} \end{pmatrix} \quad 3.2$$

Donde i es la cantidad de días que se está analizando, H es la altura (m) y Q el caudal (m^3/s) medio diario y P la precipitación diaria (mm).

Los pesos correspondientes a la primera capa se pueden expresar como:

$$IW^{1,1} = \begin{pmatrix} w_{1,1} & w_{1,2} & w_{1,3} \\ w_{2,1} & w_{2,2} & w_{2,3} \\ w_{3,1} & w_{3,2} & w_{3,3} \end{pmatrix} \quad 3.3$$

Donde el número de filas corresponde a la cantidad de neuronas que tiene la capa y el número de columnas a la cantidad de variables de entrada.

Los sesgos de la capa oculta se pueden expresar de la siguiente manera:

$$b^1 = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad 3.4$$

Por lo tanto, la entrada neta de la primera capa quedaría de la siguiente forma:

$$n^1 = \begin{pmatrix} w_{1,1} & w_{1,2} & w_{1,3} \\ w_{2,1} & w_{2,2} & w_{2,3} \\ w_{3,1} & w_{3,2} & w_{3,3} \end{pmatrix} * \begin{pmatrix} H_{1,1} & \cdots & H_{1,i} \\ Q_{2,1} & \cdots & Q_{2,i} \\ P_{3,1} & \cdots & P_{3,i} \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad 3.5$$

En la ecuación **3.5** se puede apreciar que las dimensiones resultantes del producto $WI*P$ serán $3 \times i$, y las del sesgo son 3×1 , por lo que a priori estos dos elementos no se pueden sumar, pero la suma de estos dos elementos se realiza, por ejemplo, cuando la segunda fila de la matriz peso WI se multiplica por alguna columna de la matriz de entrada P y a este resultado se le adiciona el sesgo correspondiente a la fila que se utilizó en la matriz IW para la multiplicación. El resultado entonces de la entrada neta para la capa oculta será la que se muestra en la expresión **3.6**.

$$n^1 = \begin{pmatrix} w_{1,1} * H_{1,1} + w_{1,2} * Q_{2,1} + w_{1,3} * P_{3,1} + b_1 & \cdots & w_{1,1} * H_{1,i} + w_{1,2} * Q_{2,i} + w_{1,3} * P_{3,i} + b_1 \\ \vdots & \vdots & \vdots \\ w_{3,1} * H_{1,1} + w_{3,2} * Q_{2,1} + w_{3,3} * P_{3,1} + b_3 & \cdots & w_{3,1} * H_{1,i} + w_{3,2} * Q_{2,i} + w_{3,3} * P_{3,i} + b_3 \end{pmatrix} \quad 3.6$$

La salida de la capa oculta de la red se obtendrá aplicando la función de transferencia según indicado por la fórmula **3.1** a cada elemento de la matriz de entrada neta, la cual se representa en la siguiente matriz:

$$a^1 = \begin{pmatrix} \frac{2}{1 + e^{-2*n_{1,1}^1}} - 1 & \cdots & \frac{2}{1 + e^{-2*n_{1,i}^1}} - 1 \\ \vdots & \vdots & \vdots \\ \frac{2}{1 + e^{-2*n_{3,1}^1}} - 1 & \cdots & \frac{2}{1 + e^{-2*n_{3,i}^1}} - 1 \end{pmatrix} \quad 3.7$$

La matriz de salida de la capa oculta se convierte ahora en la entrada para la siguiente capa. Esta matriz se multiplica por el vector de pesos de la segunda capa. Debido a que la entrada

a^1 de la capa de salida tiene dimensiones 3 x i y solo hay una neurona en esta capa, entonces la dimensión de la matriz peso generada LW serán 1 x 3. Considerando las dimensiones de los elementos de la capa de salida, la función de transferencia (del tipo lineal) de esta capa y realizando las mismas operaciones matemáticas que en la capa oculta, se obtiene la salida de esta capa o bien la salida de la red neuronal la cual tendrá la siguiente notación:

$$a^2 = LW * a^1 + b^2 \quad 3.8$$

$$a^2 = \left(w_{1,1}^2 \left(\frac{2}{1 + e^{-2*n_{1,1}^1}} - 1 \right) + w_{1,2}^2 \left(\frac{2}{1 + e^{-2*n_{2,1}^1}} - 1 \right) + w_{1,3}^2 \left(\frac{2}{1 + e^{-2*n_{3,1}^1}} - 1 \right) + b^2 \dots \right) \quad 3.9$$

Donde la suma de los cuatro elementos mostrados en la expresión **3.9** corresponden al primer elemento de la salida de la red. Así la salida tendrá tantas columnas o datos de salida como la entrada P o el vector objetivo, para finalmente comparar los datos objetivo con los de la salida de la red neuronal.

Para que la formula pueda ser utilizada los datos de entrada deben ser normalizados y debido a que los resultados obtenidos por la expresión **3.9** provienen de datos normalizados, se debe aplicar el proceso inverso para obtener los valores de concentración de sedimento en las unidades originales.

3.3.5 Procesamiento de entrada-salida.

Una vez incluidos los datos de entrada y objetivos, estos son normalizados para que la red neuronal tenga un mejor funcionamiento y rendimiento. Generalmente la etapa de normalización se aplica tanto a los datos de entrada como a los datos objetivo. De esta manera la salida de la red siempre cae en un rango normalizado. Los datos de salida entonces se pueden transformar de manera inversa en las unidades de los datos objetivos originales.

Es habitual que se apliquen funciones para transformar los datos de entrada y de destino que se proporcionan a la red en los valores que más se adecuen a la formación de la red.

Uno de los comandos que se aplica para normalizar los datos en redes neuronales utilizando el programa MATLAB es el “*mapminmax*”, el cual normaliza los datos en un rango de [-1, 1], tomando como datos máximo y mínimo los extremos del intervalo, por lo tanto, también los datos de salida de la red estarán entre este intervalo. Una vez terminado el proceso de entrenamiento y validación de la red, los datos de salida son convertidos de manera inversa a la normalización.

3.3.6 Subconjunto de datos en la red neuronal.

Es importante que los rangos de los valores de entrada y objetivo sean representativos de cada variable. El motivo de esto es porque las redes de múltiples capas pueden ser entrenadas para realizar una estimación buena dentro de la gama de entradas para las que han sido entrenadas. Sin embargo, no tienen la capacidad para extrapolar con precisión más allá de esta gama, por lo que es importante que los datos de entrenamiento abarcan toda la gama del espacio.

Para este trabajo se ordenaron los datos de concentración de sedimento de menor a mayor en el programa MATLAB. Luego se separaron los datos para los subconjuntos de entrenamiento y validación, obteniendo en cada uno de estos dos los subconjuntos la gama completa de la variable que se quiere estimar.

Los porcentajes para los subconjuntos de datos de entrenamiento y validación fueron 70% y 30% respectivamente. Estos porcentajes se utilizaron tomando en cuenta los que designaban otros autores a su base de datos para el entrenamiento y validación de la red neuronal (Tabla 3.4).

Tabla 3.4: Porcentajes de subconjuntos de datos entrenamiento y validación para el modelo.

Referencia	Porcentaje de datos	
	Entrenamiento	Validación
Agarwal et al. (2009)	66.7	33.3
Ahmat et al. (2008)	81.9	18.1
Barzegari et al (2015)	70	30
Besaw et al. (2010)	72.7	27.3
Bhattacharya et al. (2005)	67.2	32.8
Chutachindakate y Sumi (2008)	76.9	23.1
Cigizoglu (2004)	87.1	12.9
heng y Suetsugi (2013)	75	25
Kia et al (2013)	80	20
Kumar et al (2015)	70	30
Mustafa et al. (2011)	71.4	28.6
Nourani (2009)	75	25
Rajaei (2010)	83.3	16.7
Rajaei et al. (2009)	75	25
Sarangi et al. (2005)	66.7	33.3
Singh y Panda (2011)	66.7	33.3
Singh et al. (2011)	80	20
Yang et al. (2009)	76.7	23.3
Zhu et al. (2007)	76.2	23.8
Zounemat-Kerman et al (2016)	80	20

Tomando en cuenta los métodos descritos en el punto 3.3.3 del presente manuscrito, se analizaron las crecidas para determinar cuál es la más representativa de las 27 seleccionadas, con el fin de definir la que se utilizaría en el proceso de entrenamiento y validación de la ANN. Para esto se contabilizó la cantidad de datos por rango de concentración de sedimentos en suspensión (cada 5 mg/l), se calculó el promedio de datos para cada uno de estos rangos y la cantidad de datos totales de cada crecida.

Tabla 3.5: Cantidad de datos de concentración de sedimento mg/l según intervalos.

Crecida	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	Total
1	4	20	4	4	4	0	1	1	0	1	39
2	7	21	4	3	1	2	4	0	0	0	42
3	0	3	7	10	4	2	2	2	0	1	31
4	2	21	5	5	4	2	0	1	0	0	40
5	11	25	5	0	0	0	0	1	0	0	42
6	2	10	3	6	2	0	2	0	0	1	26
7	18	17	3	2	2	0	0	0	0	0	42
8	13	15	0	2	3	1	1	0	0	1	36
9	3	11	8	3	4	1	1	1	0	0	32
10	0	2	3	3	2	3	2	2	1	0	18
11	0	16	8	2	2	0	0	0	0	0	28
12	0	2	5	9	4	5	0	1	1	3	30
13	0	2	8	1	3	6	2	1	1	0	24
14	10	22	10	7	2	2	0	2	1	0	56
15	0	6	5	3	2	4	1	0	1	0	22
16	0	1	6	5	6	2	1	1	0	1	23
17	3	11	6	3	0	2	0	0	0	0	25
18	6	9	6	5	1	3	1	2	1	0	34
19	9	14	8	3	0	1	1	0	0	0	36
20	9	14	9	5	0	2	1	0	0	0	40
21	15	19	2	3	2	0	1	0	0	2	44
22	0	10	4	4	2	2	1	1	2	0	26
23	5	12	5	2	4	0	1	1	0	0	30
24	8	8	5	1	0	0	2	0	0	0	24
25	14	9	0	1	1	2	1	0	0	0	28
26	8	12	3	0	0	0	2	2	0	0	27
27	8	11	2	4	1	2	0	1	0	1	30
Prom.	5.7	12.0	5.0	3.6	2.1	1.6	1.0	0.7	0.3	0.4	875

En la Tabla 3.5 se muestran la cantidad de datos de concentración de sedimento en suspensión para cada uno de los rangos, los cuales fluctúan entre 0 a 50 mg/l, de las 27 crecidas. También se muestra en la última fila el promedio de datos que existe en cada uno de los rangos y en la última columna la cantidad total de datos por crecida. Tomando en cuenta esto se seleccionó la crecida 14 para entrenar y validar, ya que esta contiene la mayor cantidad de datos y representa en la mayoría de los rangos al doble del promedio de cada uno de estos.

Terminado el proceso de entrenamiento y validación se utilizan las 27 crecidas para probar la red encontrada, determinando la concentración de sedimento en suspensión y verificando la precisión de la ANN con datos diferentes a los que se utilizaron en el proceso de ajuste de pesos y sesgos.

3.4 Curva de gasto de sedimento (CGS)

Desde la década de 1970, ha habido un creciente interés en el control de la descarga de sedimentos en suspensión. La curva de gastos sedimentos (CGS) representa una relación empírica entre el caudal y la concentración de sedimento en suspensión (*Cigizoglu y Kisi, 2005*). Esta relación se define generalmente como una función de potencia como la que se muestra a continuación:

$$C_{ss} = a * Q^b \quad \mathbf{3.10}$$

En donde, a y b son coeficientes de la ecuación y se obtienen del método error cuadrático o trazando una línea de mejor ajuste. Los datos primarios que se recogen para determinar estos coeficientes son el caudal del río (Q) y la concentración de sedimento en suspensión (C_{ss}). Una vez encontrado los coeficientes se puede utilizar el caudal del río para determinar la concentración de sedimento en suspensión.

3.5 Código Matlab

Con el software MATLAB se creó un código con la red neuronal artificial utilizada en este trabajo. El código, por lo tanto, entrena y valida la ANN con la crecida seleccionada. Terminado este proceso, el código toma los datos de entrada y objetivo de las 27 crecidas y los utiliza para realizar pruebas con las cuales se mide el rendimiento de la red encontrada en el entrenamiento, por medio de los estadísticos PBIAS y RMSE que se muestran en las formulas 2.7 y 2.8, respectivamente.

Es importante destacar que cada vez que se entrena se generan soluciones diferentes, aunque se mantenga la arquitectura, las entradas y los objetivos de la ANN. Debido a esta razón es

que, en este trabajo para poder encontrar la solución más acertada en la estimación de sedimento en suspensión, se diseñó el código con un ciclo que repite 50 veces el proceso de entrenamiento y validación, realiza 50 pruebas a las 27 crecidas, calcula los estadísticos que miden el rendimiento de la ANN, y se guardan las combinaciones de pesos y sesgos que generen mejores resultados con respecto al estadístico PBIAS. Este código se ejecuta 3 veces para una misma combinación, por lo tanto, para cada combinación de entradas a la ANN, se realizarán 150 pruebas en total para cada una de las 27 crecidas, guardando los 3 mejores resultados según el estadístico mencionado anteriormente.

4 RESULTADOS

4.1 Estimación de transporte de sedimento en suspensión mediante el modelo de ANN

Se analizó la precisión del modelo de ANN para estimar el transporte de sedimento en suspensión utilizando la arquitectura y métodos descritos en el capítulo anterior.

Se utilizó el estadístico PBIAS para seleccionar las ANN resultantes de los procesos de entrenamiento y validación, que genera estimaciones más acertadas de C_{ss} para cada una de las combinaciones de variables de entrada (tabla 3.1). Este estadístico evalúa el rendimiento de las estimaciones realizadas por la ANN. Tomando en cuenta que se realizan varias pruebas para una misma entrada del modelo y que con esto se genera distintas estimaciones, se selecciona la ANN que obtenga mayor cantidad de crecidas con un rendimiento del estadístico PBIAS por sobre una categoría esperada. Para esto se expone las clasificaciones del rendimiento del estadístico PBIAS de acuerdo estudios de estimaciones de sedimento (Moriassi et al, 2007).

Tabla 4.1: Índices de rendimiento del PBAS para estimaciones de sedimento (Moriassi et al., 2007).

Clasificación de Rendimiento	PBIAS (%)
Muy Bueno	$PBIAS < \pm 15$
Bueno	$\pm 15 \leq PBIAS < \pm 30$
Satisfactorio	$\pm 30 \leq PBIAS < \pm 55$
Insatisfactorio	$PBIAS \geq \pm 55$

Tal como se dijo anteriormente se realizaron varias pruebas para cada combinación de las variables de entrada. Todas estas pruebas se realizaron ya que el modelo de ANN genera soluciones diferentes cada vez que se entrena y valida la red, incluso cuando se mantiene la arquitectura y las entradas del modelo. Esto se atribuye al método que utiliza la ANN para converger a una solución “óptima” de pesos y sesgos, el cual se describió en la sección 2.1.7.1.

De todas las pruebas realizadas para cada combinación de variables de entrada, se seleccionaron tres. Así se obtienen 3 grupos diferentes de pesos y sesgos para cada una de las 8 combinaciones.

El proceso general de la selección de una de las tres pruebas para cada combinación se realiza mediante un ciclo en el programa MATLAB y el cual se resume a continuación.

1. Se entrena y valida la red con la crecida representativa, encontrando así los pesos y sesgos de la ANN.
2. Se estiman 27 grupos de Css utilizando las variables de entrada correspondiente a las 27 crecidas y aplicando la ANN encontrada en (1).
3. Mediante los grupos estimados y observados de Css se calculan 27 PBIAS.
4. Si la cantidad de PBIAS calculado en (3) con rendimiento “bueno” o superior, es mayor a la del ciclo anterior, se guarda los pesos y sesgos del presente ciclo. Si esto no se cumple, entonces se conservan los pesos y sesgos del ciclo anterior.
5. Se repiten los procesos del 1) al 4) 50 veces.

Con este ciclo se guarda los pesos y sesgos que generan mayor cantidad de crecidas con estimaciones que obtienen un índice de rendimiento “bueno” según la Tabla 4.1. Los resultados finales de cada ciclo para cada combinación de variables de entrada se presentan en el anexo 1.

Según el anexo 1 el modelo puede predecir de manera acertada la concentración del sedimento en suspensión en la mayoría de las crecidas seleccionadas. Esto se ve reflejado en el estadístico PBIAS, en donde la mejor estimación contiene 23/27 crecidas con un PBIAS menor a $\pm 30\%$, lo cual sugiere una estimación “buena” y el mínimo de crecidas con este rendimiento considerando todas las pruebas hechas fue de 18/27.

En la Tabla 4.2 se muestran los resultados obtenidos de la pruebas realizadas, buscando las mejores estimaciones con respecto al estadístico PBIAS.

Tabla 4.2: Estimaciones más acertadas de C_{ss} según el estadístico PBIAS para cada combinación de variables de entrada.

Combinación	Prueba	P_{bias}≤±55%	P_{bias}≤±30%	P_{bias}≤±15%
COMB1	1	26	19	11
COMB1	2	26	19	11
COMB1	3	27	19	11
COMB2	4	24	18	11
COMB2	5	25	22	10
COMB2	6	25	18	8
COMB3	7	25	20	9
COMB3	8	22	19	11
COMB3	9	20	20	8
COMB4	10	27	21	12
COMB4	11	25	19	12
COMB4	12	27	20	11
COMB5	13	24	20	10
COMB5	14	25	20	12
COMB5	15	26	19	8
COMB6	16	27	20	12
COMB6	17	26	23	12
COMB6	18	27	22	14
COMB7	19	23	19	9
COMB7	20	27	21	14
COMB7	21	26	21	11
COMB8	22	26	23	11
COMB8	23	26	21	11
COMB8	24	27	22	10

Se obtienen para todas las pruebas de la Tabla 4.2 como mínimo 67% de crecidas con estimaciones que se consideran buenas según el estadístico PBIAS. De todas las pruebas de esta tabla, en promedio se tiene que el 40% de las crecidas obtienen un PBIAS menor al ±15%, lo que se traduce según su clasificación a una muy buena estimación.

En la Tabla 4.2 es importante indicar que las estimaciones de C_{ss} obtenidas solo con las variables de entrada en el tiempo “ t ” son igual de precisas que las estimaciones que incluyen las variables en el tiempo “ $t-1$ ”. Las combinaciones que contienen variables en el tiempo “ $t-1$ ” obtienen en promedio aproximadamente 25 y 20 crecidas con rendimientos por lo menos satisfactorio y bueno. Para las combinaciones que solo tienen variables en el tiempo “ t ” se obtienen en promedio 26 y 19 crecidas con estimaciones que por lo menos tienen rendimiento satisfactorio y bueno.

4.2 Análisis de crecidas según resultados

Aplicando la arquitectura de la ANN y el código de MATLAB diseñado, se encontraron las soluciones más acertadas de concentración de sedimento según los dos estadísticos descritos anteriormente, obteniendo tres pruebas por crecida para cada una de las combinaciones de entradas utilizadas. De este modo con los datos de entrada y objetivo de una crecida se obtienen nuevas concentraciones de sedimento calculadas por la ANN con diferentes pesos y sesgos, obteniendo un total de 24 set de sedimentos estimados. Cada uno de estos sets se comparan con los datos reales de sedimento (objetivos de la ANN) por medio de los estadísticos PBIAS y RMSE.

Con los resultados de los estadísticos se puede afirmar que la ANN aplicada puede estimar de manera acertada la concentración de sedimento para la condición de un aumento de caudal en el curso fluvial. Pero dentro de los resultados existen crecidas que en general no presentan una estimación satisfactoria con respecto a los dos estadísticos utilizados.

Analizando el PBIAS y el RMSE para cada una de las crecidas que se muestra en los resultados del anexo 1, se destacan 2 crecidas de las 27 seleccionadas por sus bajos rendimientos en la estimación de la C_{ss} . En la siguiente tabla se resume la cantidad de pruebas para cada una de las crecidas con los resultados menos precisos de acuerdo a los estadísticos PBIAS y RMSE, según el comportamiento general de éstos.

Tabla 4.3: Cantidad de pruebas que tienen indicadores de rendimiento por debajo de lo normal.

Crecida	N° de pruebas	
	PBIAS $>\pm 30\%$	RMSE >15 mg/l
1	3	0
2	1	0
3	0	0
4	0	0
5	10	0
6	0	5
7	23	0
8	1	0
9	1	1
10	3	3
11	1	0
12	24	20
13	0	0
14	0	0
15	9	2
16	9	0
17	6	0
18	18	1
19	0	0
20	1	0
21	15	2
22	4	1
23	6	1
24	4	0
25	12	0
26	3	1
27	9	4

En la Tabla 4.3 se identifica que para los resultados según el PBIAS, las crecidas 7 y 12 tienen rendimientos menos acertados que las otras crecidas, obteniendo 23/24 y 24/24 pruebas con rendimientos que están bajo el promedio. Estas crecidas presentan una anomalía con respecto a las otras 25 seleccionadas.

En la Figura 4.6 se muestra el promedio de 24 pruebas realizadas para el estadístico PBIAS de crecidas.

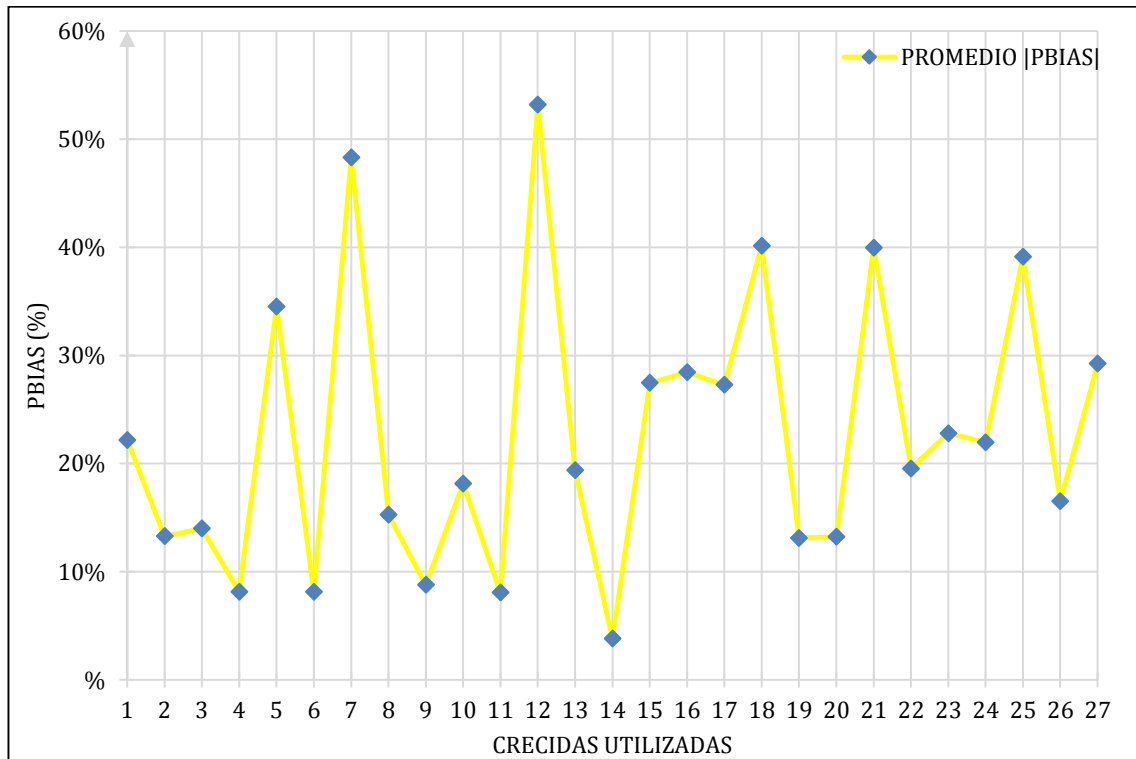


Figura 4.1: Promedio de PBIAS para cada crecida seleccionada.

En la Figura 4.6 se destacan los promedios de PBIAS de las crecidas 7 y 12. El rendimiento poco asertivo de las estimaciones para estas dos crecidas es producto de que los pesos y sesgos encontrados en el proceso de entrenamiento y validación no representan las C_{ss} de estas dos crecidas.

En cuanto al RMSE la crecida 12 es la que genera un mayor error en sus estimaciones, alcanzando un promedio de este estadístico de 16.9 mg/l, lo cual es un 200% del promedio de las otras crecidas. Los promedios de los RMSE para cada una de las crecidas considerando las 24 pruebas realizadas se exponen en la Figura 4.6.

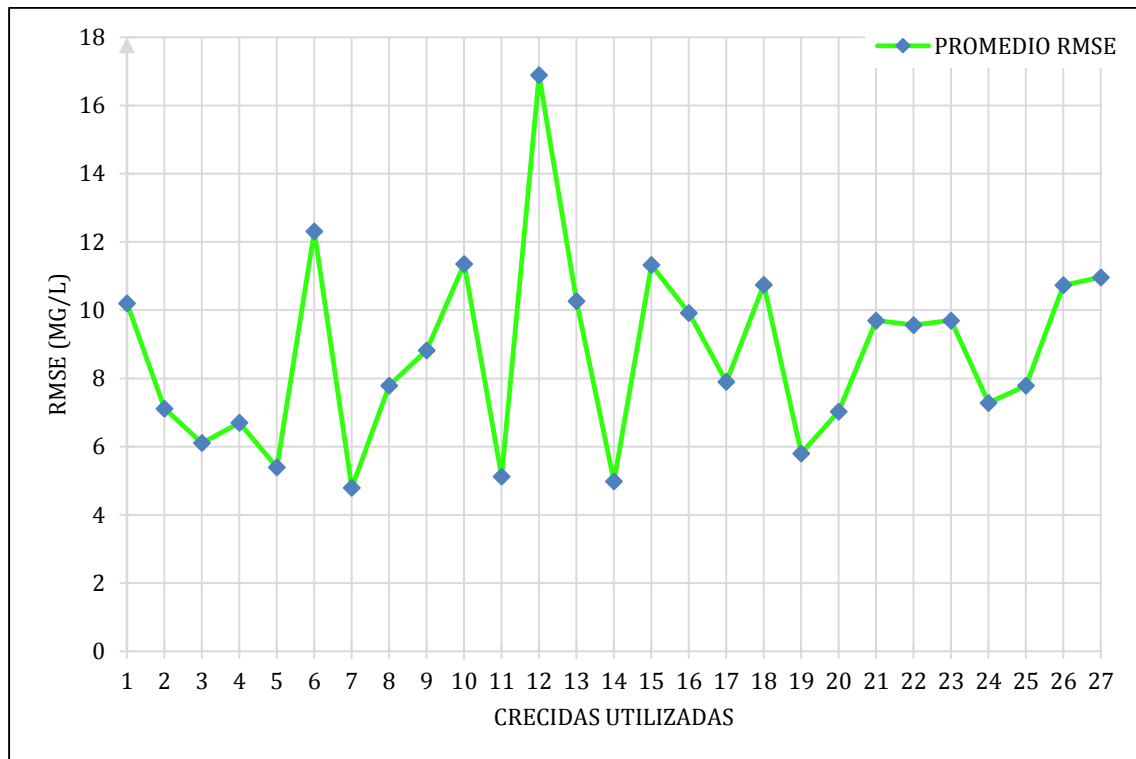


Figura 4.2: Promedio de RMSE para cada crecida seleccionada.

Como en la mayoría de las estimaciones de C_{ss} se obtienen rendimientos satisfactorios, se puede decir que la ANN es un modelo capaz de predecir de forma acertada la C_{ss} . Por esto es probable que la poca exactitud que obtuvieron las estimaciones de la crecida 7 y 12 se atribuyen a que la C_{ss} está ligada a otra variable o fenómeno que no se consideró en este estudio, y se tratan de casos puntuales que no ocurren con frecuencia.

4.3 Comparación del modelo de ANN con la CGS

Además de verificar el rendimiento de las predicciones de ANN con los estadísticos utilizados, se utiliza la curva de gasto para estimar la concentración de sedimento, y así verificar cuál de estos dos métodos proporciona valores más precisos en la estimación de la variable de interés.

Para la curva de gasto de sedimento se utilizó la misma crecida con la cual se entrenó y validó la red, por lo tanto con los caudales y la concentración de sedimento de la crecida 14 se determinó los coeficientes a y b de la ecuación 3.10. Luego con las 27 crecidas seleccionadas,

se compararon los resultados del modelo de ANN con las estimaciones que proporciona el método de CGS.

La fórmula de la curva de gasto de sedimento se obtiene ajuste de una función potencial de los datos de caudal y de concentración de sedimento utilizados para entrenar y validar la ANN. En la Figura 4.3 se muestra el grafico obtenido con el método de CGS.

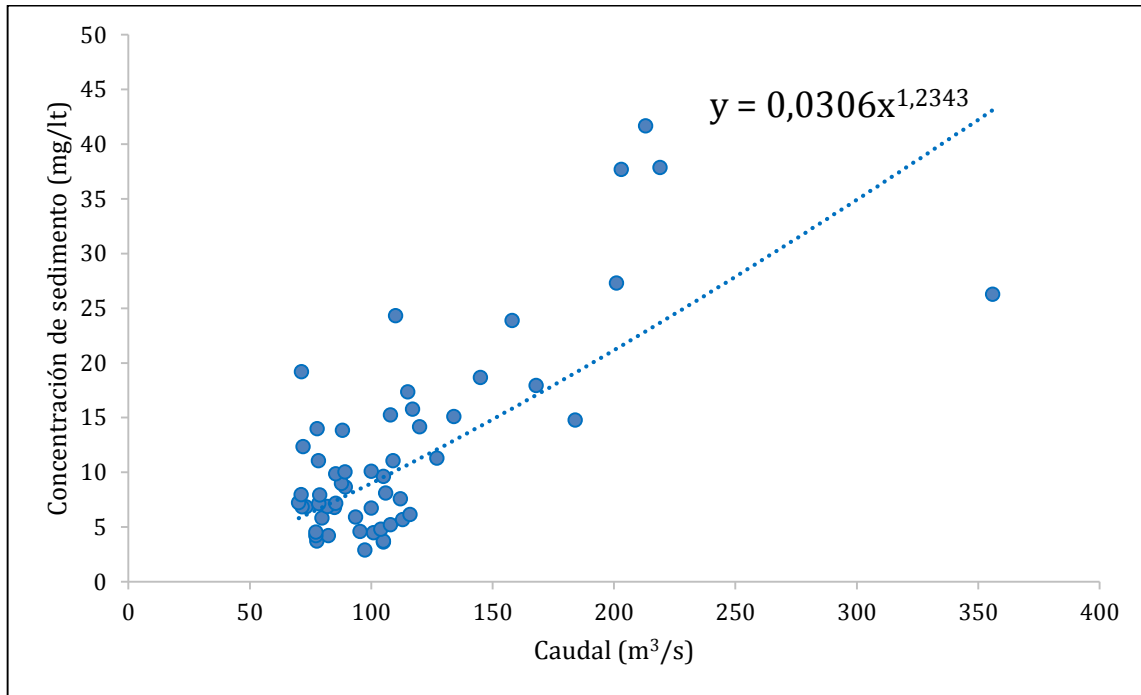


Figura 4.3: Curva de Gasto de Sedimento.

Así como se muestra en la Figura 4.3, la fórmula para determinar la concentración de sedimento, en base al caudal del afluente en m³/s es:

$$C_{ss}^{(2)} = 0.0306 * Q^{1.2343} \quad \mathbf{4.1}$$

Para esto, se estimó mediante la fórmula **4.1** las C_{ss} de las 27 crecidas y luego se calculó los estadísticos PBIAS y RMSE de las observaciones y estimaciones de C_{ss} .

Teniendo 24 combinaciones de pesos y sesgos correspondientes a las 24 pruebas, se seleccionará la que estime de forma más acertada la C_{ss} que el método de CGS.

Los estadísticos PBIAS y RMSE de las 24 pruebas realizadas por la ANN y del método de CGS se comparan y se elige la ANN que obtenga mayor cantidad de crecidas con rendimientos superiores según estos dos estadísticos.

A continuación, se presenta un gráfico en donde se observan la cantidad estadísticos del modelo de ANN que obtienen un rendimiento por sobre las estimaciones del modelo de CGS.

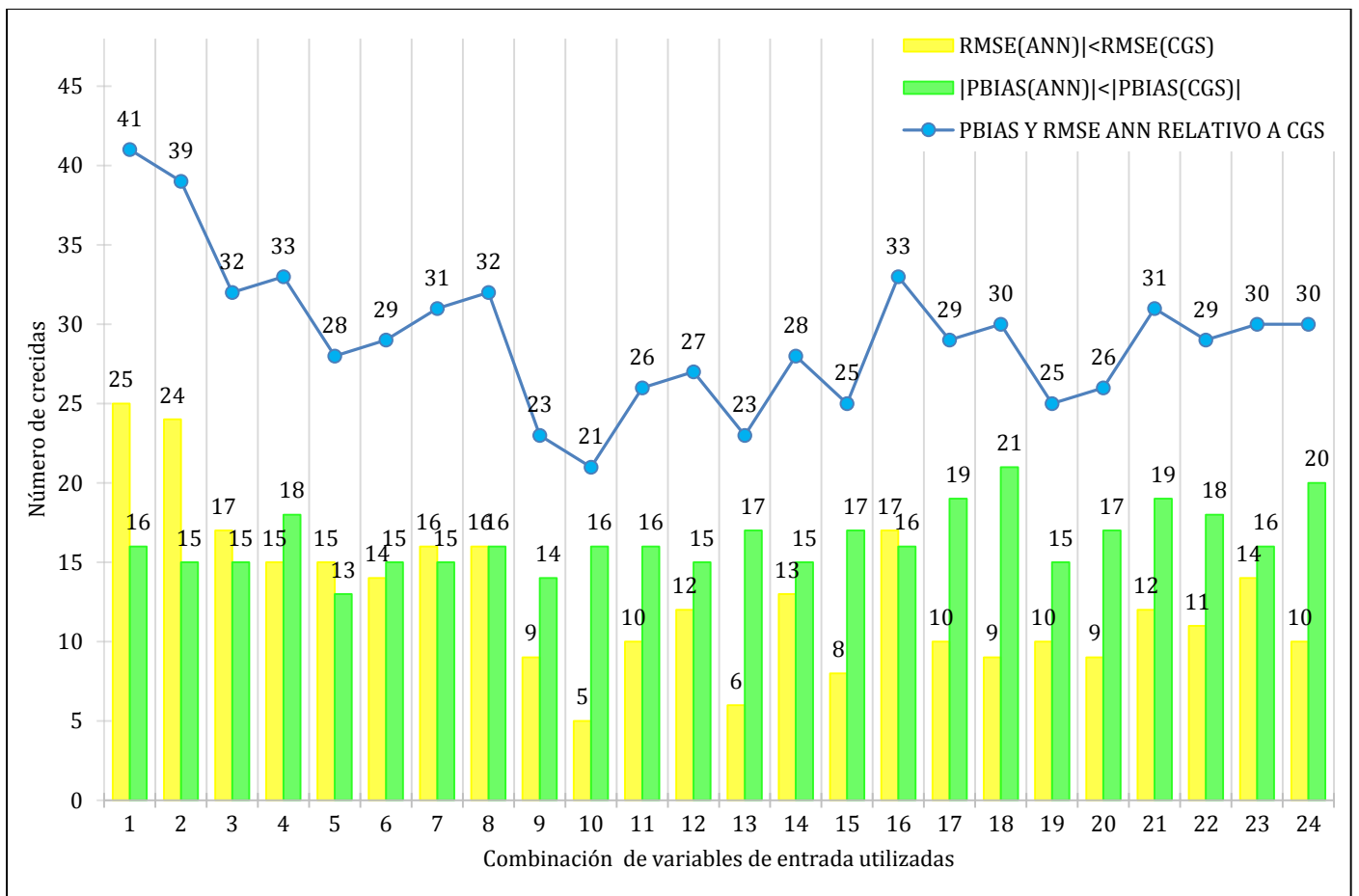


Figura 4.4: Estimaciones de la ANN por sobre las de CGS.

Las barras de color amarillo mostradas en la Figura 4.6 representan la cantidad de estimaciones de la ANN que generan rendimientos mejores que las estimaciones del modelo

de CGS, según el RMSE. Las barras de color verde son las estimaciones generadas con la ANN que están por sobre el modelo de CGS, según el PBIAS. Esto se realiza para cada una de las 24 pruebas de la ANN que se muestran en el anexo 1.

En esta figura se puede observar que de las 24 pruebas solo en 7 las estimaciones de C_{ss} realizadas por el modelo de CGS están por sobre las realizadas por la ANN, esto tomando en cuenta los dos estadísticos utilizados. Por lo tanto, la ANN genera en el 70 % de los casos mejores estimaciones que la CGS, independiente de las variables de entrada que se utilicen.

Para seleccionar la mejor estimación realizada por la ANN se consideran los dos estadísticos utilizados, por lo tanto, en la Figura 4.6 se muestran la suma de las cantidades de PBIAS y RMSE calculados de las estimaciones de la ANN, que se consideren más precios que los PBIAS y RMSE obtenidos de la CGS. De esta forma la prueba 1 es la que obtiene más crecidas, con 25/27 crecidas con mejores estimaciones hechas con la ANN que con la CGS según el RMSE y 16 de 27 crecidas con mejores estimaciones hechas con la ANN que con la CGS, según PBIAS.

En la Tabla 4.4 se muestran los resultados del PBIAS y RMSE calculados según el modelo de ANN y CGS.

Tabla 4.4: Estadísticos PBIAS y RMSE de las estimaciones de C_{ss} de los modelos de CGS y ANN

CRECIDA	CGS		PRUEBA 1 (ANN)	
	RMSE(mg/l)	PBIAS(%)	RMSE(mg/l)	PBIAS(%)
1	10.76	28.22	10.03	24.69
2	7.37	30.03	7.28	13.63
3	8.34	25.25	6.64	26.76
4	5.17	4.79	4.47	-0.80
5	4.66	13.22	3.59	-43.81
6	12.00	19.72	10.11	11.34
7	3.54	-28.54	4.02	-40.89
8	7.82	6.71	7.04	-12.29
9	5.65	7.75	5.08	0.31
10	8.75	7.07	8.67	-9.55

11	5.14	22.76	3.60	6.77
12	18.33	69.19	16.70	62.28
13	11.26	35.55	10.28	29.79
14	6.67	13.35	5.07	3.03
15	11.17	50.97	12.28	43.36
16	11.97	41.95	9.74	34.18
17	5.17	-10.48	4.86	-19.67
18	8.99	-16.20	8.78	-26.72
19	5.46	16.26	3.88	-5.67
20	6.54	22.74	5.00	14.94
21	7.87	-27.84	6.66	-35.72
22	9.28	22.94	8.13	27.46
23	8.88	44.39	8.21	31.75
24	6.55	-7.88	5.90	-1.72
25	7.21	-29.54	6.82	-41.28
26	9.57	47.20	8.50	25.39
27	7.10	-16.38	5.44	-22.33

El promedio del RMSE para el método de CGS es de 8.19 mg/l y para el modelo de ANN es de 7.29, lo cual es un 11 % más bajo. Se observa que el RMSE de las estimaciones es más bajo en el modelo de ANN que en el de CGS, obteniendo de las 27 crecidas, 25 con valores menores correspondiente a un 93% aproximadamente. Con esto se puede decir que la ANN genera menores errores en sus estimaciones que el método de CGS.

Con respecto al estadístico PBIAS en la ANN hay 19 crecidas que están por sobre la calificación buena y para la CGS son 20. Si bien existe 1 crecida más que se estima de forma más precisa con la CGS que en el modelo de ANN, según el estadístico PBIAS, para cada una de las clasificaciones de este estadístico hay: con calificación muy buena, 11 crecidas de la ANN y 8 de la CGS; con calificación buena, 8 crecidas de la ANN y 12 de la CGS; con calificación satisfactoria, 7 crecidas de la ANN y 6 de la CGS; con calificación insatisfactoria, 1 crecidas de la ANN y 1 de la CGS. Ahora comparando el valor obtenido para cada crecida de este estadístico son 16/ 27 crecidas con mejores estimaciones de la ANN que del modelo de CGS. También para el PBIAS los promedios de los valores absolutos son de 24.7% para el modelo de CGS y de 22.8% para el modelo de ANN por lo cual, según lo mencionado el modelo de ANN sería más preciso que el modelo de CGS.

Esto se ve gráficamente en las Figura 4.5 y Figura 4.6 al comparar los valores del RMSE y del PBIAS de los modelos de ANN y CGS.

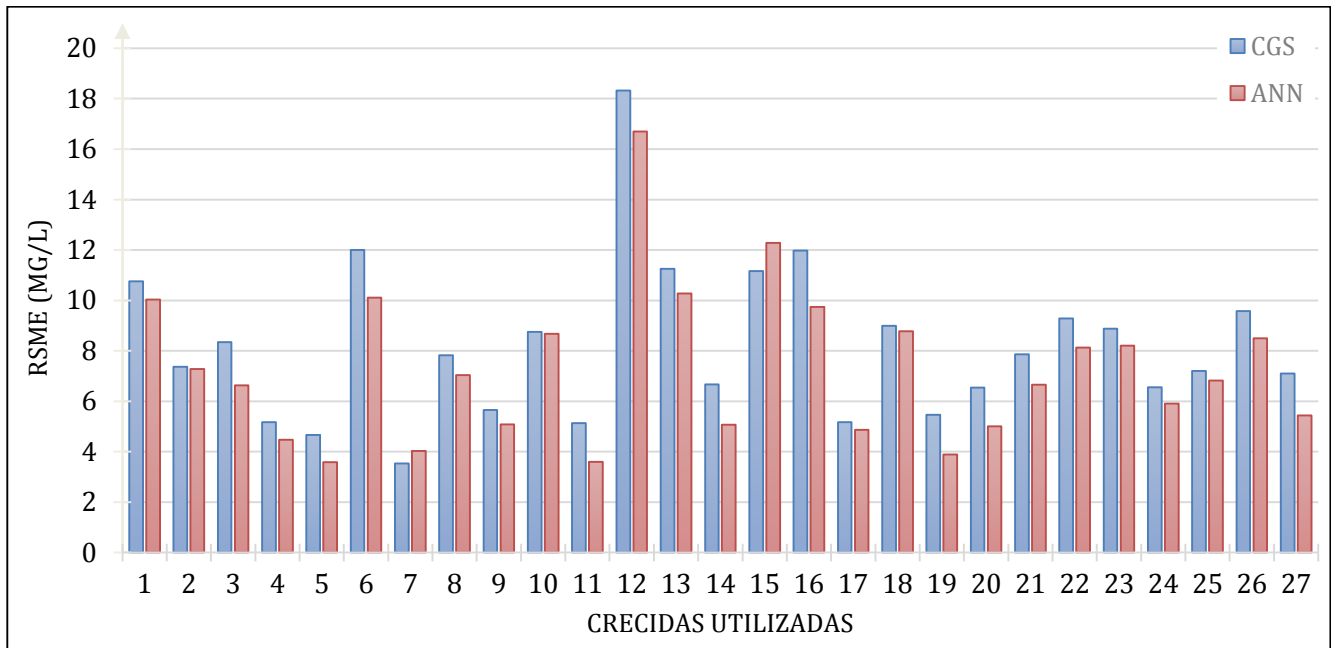


Figura 4.5: RMSE de ANN vs RMSE de CGS.

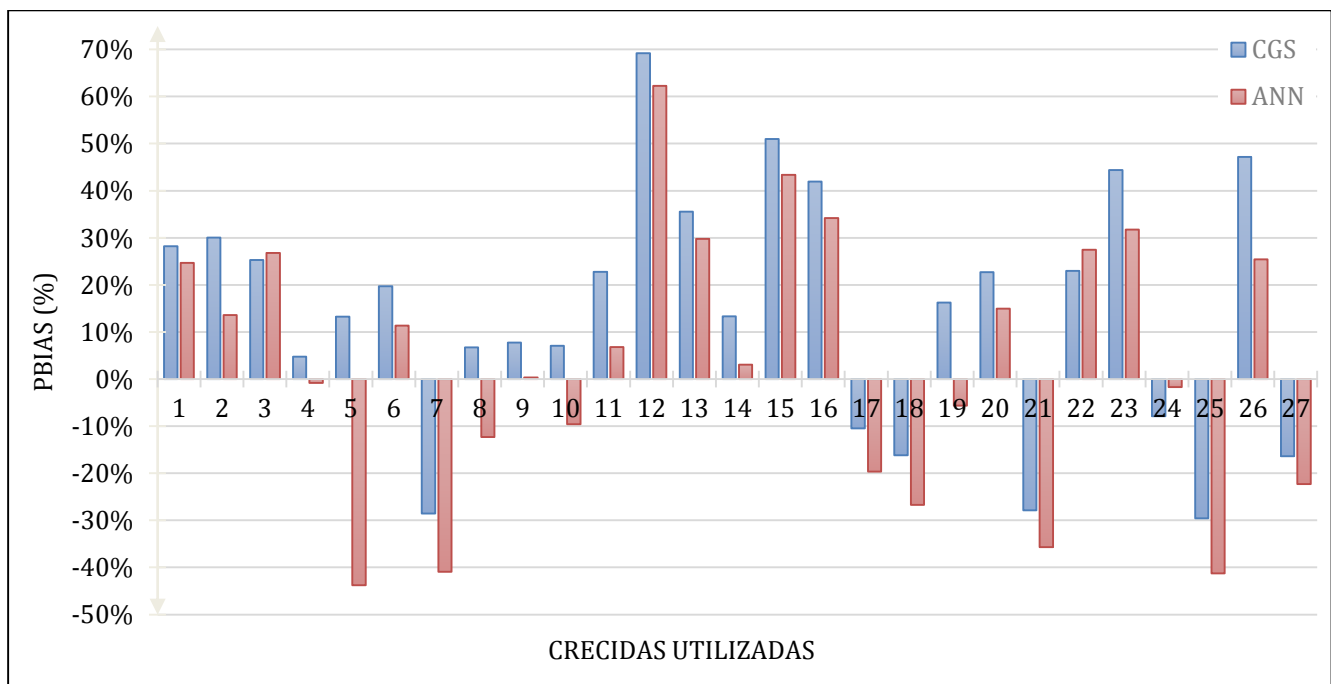


Figura 4.6: PBIAS de ANN vs PBIAS de CGS.

En el anexo 2 se incluyen las gráficas para cada una de las crecidas con los datos de concentración de sedimento en suspensión estimados (output) por la ANN y reales (target) en mg/l, que corresponden a la prueba de 1 de la Tabla 4.4; **Error! No se encuentra el origen de la referencia.** Se incluye la línea de tendencia (línea azul) de los datos junto su ecuación, el Coeficiente de Correlación y una línea en 45° que representa la igualdad de los datos estimados y reales para una visualización más clara de cómo deben ser la distribución de los datos con una estimación perfecta.

4.4 Formulación matemática a partir de la ANN

De la arquitectura de la ANN se puede construir una fórmula que contenga como variables de la estimación de sedimento, las entradas utilizadas en el modelo. De los resultados obtenidos se extraen los pesos y sesgos con los que se definen las constantes de la fórmula, las cuales se ingresan en la descomposición de la arquitectura de la ANN, que para este trabajo sería la Figura 3.1 mostrada anteriormente.

Según lo descrito anteriormente la prueba 1 es la que estima de forma más acertada la C_{ss} utilizando el modelo de ANN.

Por lo tanto, las variables $H(t)$, $Q(t)$, $P(t)$ son las que definirán la concentración de sedimento en suspensión en la fórmula.

Los pesos y sesgos de la prueba 1 son los siguientes.

Los pesos de la primera capa:

$$IW^{1,1} = \begin{pmatrix} 1.99224205606536 & 3.26611299664986 & 0.430991236817096 \\ 1.06877726686549 & 1.4740170174899 & 0.452495357606154 \\ -0.198321010417552 & -0.348326108112707 & 1.99224205606536 \end{pmatrix} \quad 4.2$$

Los pesos de la segunda capa:

$$LW^{2,1} = (0.583105937691059 \quad 0.387101257044076 \quad -0.654444830430267) \quad 4.3$$

Los sesgos de la primera capa:

$$b^1 = \begin{pmatrix} 0.693896411396805 \\ 0.138510898193662 \\ -1.82143141718272 \end{pmatrix} \quad 4.4$$

Los sesgos de la segunda capa:

$$b^2 = -0.4637330104580625 \quad 4.5$$

Por lo tanto, como resultado de las entradas y los pesos y siguiendo el procedimiento utilizado para definir la fórmula 3.9, se tiene que la concentración de sedimento en suspensión es:

$$C_{ss}^1 = \frac{-1.38779}{e^{-3.05651H(t)+6.53223*Q(t)+0.861982*P(t)+1.38779} + 1} + \frac{0.277022}{e^{2.13755*H(t)+2.4803*Q(t)+0.904991*P(t)+0.277022} + 1} + \frac{3.64286}{e^{-0.396642*H(t)-0.696652*Q(t)+3.98448*p(t)-3.64286} + 1} - 1.45276 \quad 4.6$$

Considerando la altura H m, el caudal Q en m³/s, la precipitación P en mm y la concentración de sedimento en suspensión C_{ss} en mg/l.

La fórmula 4.6 entrega las estimaciones de sedimento en suspensión de manera acertada para las condiciones del río seleccionado, tomando en cuenta además que se deben normalizar los datos de entrada. Si bien no es posible darle una explicación física a su estructura, se integran más variables que la utilizada en la curva de gasto de sedimento, la cual solo integra como variable de entrada el caudal del afluente.

5 DISCUSIÓN Y CONCLUSIONES

Con los datos seleccionados se pone a prueba la capacidad de la ANN para predecir la C_{ss}. Con los resultados obtenidos según el estadístico PBIAS la ANN puede generar en promedio estimaciones con clasificación muy buena para aproximadamente un 40% de las crecidas seleccionadas. Además, al comparar las estimaciones de la prueba 1 realizada por la ANN, con las estimaciones de CGS, se obtuvo que en aproximadamente el 60% de las crecidas generan mejores estimaciones con el modelo de ANN, por lo que según estos resultados el modelo de ANN obtendría un 18.5 % de estimaciones de C_{ss} mejores que el modelo de CGS.

Esto es confirmado con el RMSE, el cual obtiene un total de 25 crecidas con mejores resultados en modelo de ANN respecto al modelo de CGS. De este modo se obtendría para estimaciones futuras en 85 % de los casos mejores resultados con el modelo de ANN que con el modelo de CGS. Como es de esperar el error promedio calculado con el RMSE de la ANN es menor que el calculado con la CGS, obteniendo 7.29 y 8.16 mg/l respectivamente.

De todas formas, el método de CGS genera estimaciones acertadas de C_{ss} obteniendo en 26/27 crecidas rendimientos que se consideran por lo menos satisfactorias según el PBIAS.

El criterio de analizar las crecidas seleccionadas de manera separada es afirmado en este estudio, ya que como se demostró que las crecidas 7 y 12 no generan estimaciones acertadas. Si se hubieran analizado todas las crecidas juntas para entrenar, validar y realizar pruebas, estas dos crecidas serían outlier y disminuirían el rendimiento de las estimaciones de la ANN.

De acuerdo combinaciones de entradas utilizadas, no se puede afirmar que al incorporar variables en el tiempo “t-1” aumenta el rendimiento de la ANN para la estimación de C_{ss}. Teniendo 3 pruebas para las variables en “t” se determinó que una de estas pruebas era mejor de las 24 que se analizaron.

Si bien es posible encontrar una fórmula con la cual se pueda determinar de forma acertada la concentración de sedimento en suspensión tomando como base la arquitectura de la ANN y los pesos y sesgos hallados, se debe verificar que procesos realizó la ANN para la correcta

utilización de dicha fórmula. La fórmula **4.6** antes de ser utilizada, se deben estandarizar los valores de las entradas en un rango de $[-1,1]$, tomando como los valores mínimos y máximos de cada variable, los extremos de este rango. Además de esto se debe tener en consideración que la formula realiza estimaciones correctas según las condiciones y variables para la cual fue entrenada la ANN, por lo que no asegura una estimación acertada fuera de estos parámetros.

Un aspecto importante de este trabajo es que se encontraron estimaciones satisfactorias separando los datos de un río por una característica, la cual en este caso es la crecida del caudal del curso fluvial. Dado esto y con las entradas utilizadas se pudo obtener soluciones satisfactorias con el modelo de ANN como también con el método de CGS, lo cual deja abierta la investigación sobre la influencia de las características de los hidrogramas y los datos de entrada, sobre la estimación de sedimento en suspensión.

Si bien la metodología utilizada en el presente trabajo puede ser implementado en cualquier río que se desee estimar de forma acertada la concentración de sedimento en suspensión, se deben tener en consideración la disposición y el tipo de sedimento del río, el clima de la zona de estudio, los valores máximos, mínimos y promedios de las variables de entrada y de destino de la ANN, etc. Esto es importante, ya que por ejemplo la estructura de la ANN utilizada en este trabajo permite estimar de forma acertada concentraciones de sedimento en suspensión que varía entre 0 a 50 mg/lit, pero no asegura que la arquitectura de la ANN sea la adecuada para estimar de forma acertada valores que estén fuera de este rango. También es posible que el suelo de la cuenca del río estudiado no presente zonas desprovistas de vegetación, lo que afectaría a la erosividad de la precipitación al punto tal de dejar de ser ésta una variable importante en el transporte de sedimento en suspensión.

6 REFERENCIAS

Agarwal, A., Rai, R., Upadhyay A. (2009). “Forecasting of Runoff and Sediment Yield Using Artificial Neural Networks”. *Journal of Water Resource and Protection*, Vol.1, No. 5 pp. 368-375.

Ahmat, N., Harun, S., Ariffin J., Abdul-Talib, S., (2008). “Predicción transporte de sedimentos mediante 3-Layer-Feed adelante MLP Networks”. *The 11th Internacional Conferencia de drenaje urbano*, Edimburgo, agosto 31-septiembre 5 -tember.

Barzegari, F., Yosefi, M.y Talebi, A. (2015). Estimating Suspended Sediment by Artificial Neural Network (ANN), Decision Trees (DT) and Sediment Rating Curve (SRC) Models (Case study: Lorestan Province, Iran). *Civil Engineering Infrastructures Journal*, 48(2): 373-380- ISSN: 2322-2093.

Beale, M. Hagan, M., Demuth, H., (2015). “Neural Network Toolbox. User’s Guide. The MathWorks R2015a.

Besaw, L., Rizzo, D., Bierman P., Hackett, W., (2010). “Advances in Ungauged Streamflow Prediction Using Artificial Neural Networks”. *Journal of Hydrology*, Vol. 386, No. 1-4, 2010, pp. 23-37.

Bhattacharya, B., Price R., Solomatine, D., (2005). “Data- Driven Modelling in the Context of Sediment Transport”. *Physics and Chemistry of the Earth*. Vol. 30, No. 4-5, pp. 297-302.

ASCE (2000). Task Committee on the application of ANNs in hydrology. Artificial neural networks in hydrology, II: hydrologic application. *J Hydrol Eng*;5(2):124–37.

Chanson, H. (2004). *The Hidraulics of Open Channel Flow: An Introduction Second Edition*. Elsevier Butterworth-Heinemann.

Chutachindakate C., Sumi, T., (2008). “Sediment Yield and Transportation Analysis: Case Study on Managawa River Basin”. *Annual Journal of Hydraulic Engineering*. Vol. 52, 2008, pp. 157-162

Cigizoglu, H.K. (2004). Estimation and forecasting of daily suspended sediment data by multi layer perceptrons. *Adv. Water Resour.* 27, 185–195.

Cigizoglu, H.K., Kisi, O. (2005) Flow prediction by three back propagation techniques using k-fold partitioning of neural network training data *Nordic Hydrology*, 36 (1), pp. 49–64

Cobaner, M., Unal, B., Kisi, O. 2009. Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data. *J. Hydrol.* 367, 52–61.

Heng, S., Suetsugi, T. (2013). Using Artificial Neural Network to Estimate Sediment Load in Ungauged Catchments of the Tonle Sap River Basin, Cambodia. *Journal of Water Resource and Protection* 5, 111-123.

Hsie, M., Ho, Y.F., Lin, Ch.T., Yeh I.Ch., (2012). Modeling asphalt pavement overlay transverse cracks using the genetic operation tree and Levenberg–Marquardt Method. *Expert Systems with Applications* 39 4874–4881

Jie, L.C., Yu, S.T. (2011). Suspended sediment load estimate using support vector machines in Kaoping River basin. 978-1-61284-459-6/11 IEEE.

Kakaei, E., Moghaddam, A., Ahmadi, A. (2013). Daily suspended sediment load prediction using artificial neural networks and support vector machines. *Journal of Hydrology*, Vol. 478, pp 50-62.

Kia, E., Emadi, A., Fazlola, R., (2013). “Investigation and Evaluation of Artificial Neural Networks in Babolroud River Suspended Load Estimation. *J. Civil Eng. Urban.*,3 (4): 183-190.

Kisi, O. (2004). Daily suspended sediment modelling using a fuzzy differential evolution approach *Hydrological Sciences Journal*, 49 (1) pp. 183–197.

Kumar, D., Pandey, A., Sharma, N., Wolfgang-Albert F. (2015) "Daily suspended sediment simulation using machine learning approach". *Catena* 138 77–90.

Nourani, V., (2009). “Using Artificial Neural Networks (ANNs) for Sediment Load Forecasting of Talkherood River Mouth”. *Journal of Urban and Environmental Engineering*. Vol. 3, No. 1, pp. 1-6.

Maier, H. R., Dandy, G. C., (2000). Neural Networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software*, 15 (1) pp. 101-124.

Melesse, A. M., Ahmad, S., McClain, M. E., Wang, X. y Lim, Y. H. (2011). Suspended Sediment Load Prediction of River Systems: An Artificial Neural Network Approach, *Agricultural Water Management*, Vol. 98, No. 5 pp. 855-866.

MOP-DGA. (2004). Diagnóstico y clasificación de los cursos y cuerpos de agua según objetivos de calidad: Cuenca del río Imperial.

MOP-DGA. (2011). Regionalización de cálculos en escorrentía en cuencas pluviales. Dirección de proyectos y planificación.

Moriasi, D., Arnold, J., Van Liew, M., Bingner, R., Harmel, R., Veith, T. (2007). “Model evaluation guidelines for systematic quantification of accuracy in watershed simulations”. *Transactions of the ASABE*. Vol.50 (3): 885-900.

Mustafa, M., Isa, M., Rezaur, R., (2011). "A Comparison of Artificial Neural Networks for Prediction of Suspended Sediment Discharge in River: A Case Study in Malaysia". World Academy of Science, Engineering and Technology. Vol. 81, 2011, pp. 372-376.

Rajae, T., (2010). "Wavelet and ANN Combination Model for Prediction of Daily Suspended Sediment Load in River". Science of the Total Environment. Vol. 409, No. 15, pp. 2917-2928.

Rajae, T., Mirbagheri, S.A., Zounemat-Kermani, M., Nourani, V. (2009). Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models, Science of the Total Environment, 407: 4916-4927.

Sarangi, A., Madramootoo, C., Enright, P., Prasher S., Patel, R., (2005). "Performance Evaluation of ANN and Geomorphology-Based Models for Runoff and Sediment Yield Prediction for a Canadian Watershed". Current Science. Vol. 89, No. 12 pp. 2022-2033.

Singh, A., Imtiyaz, M., Isaac R., Denis, D., (2011). "Comparison of Soil and Water Assessment Tool (SWAT) and Multilayer Perceptron (MLP) Artificial Neural Network for Management, Vol. 104, 2011, pp. 113-120.

Singh, G., Panda, R., (2011). "Daily Sediment Yield Modeling with Artificial Neural Network Using 10-Fold Cross Validation Method: A Small Agricultural Watershed, Kapgari, India". International Journal of Earth Sciences and Engineering. Vol. 4, No. 6, pp. 443-450.

Tayfur, G. (2002). Artificial neural networks for sheet sediment transport, J Hydrological Sciences, 47 (6): 879-892.

Yang, C., Marsooli, R., Aalami, M., (2009). "Evaluation of Total Load Sediment Transport Formulas Using ANN". International Journal of Sediment Research. Vol. 24, No. 3, pp. 274-286.

Zhu, Y.M., Lu, X.X., Zhou Y.M. (2007). Suspended sediment flux modeling with artificial neural network: an example of the Longchuanjiang River in the Upper Yangtze Catchment, China. Geomorphology, (84): 111–125.

Zounemat M., Kisi, O., Adamowski J., Ramezani, A., (2016). "Evaluation of data driven models for river suspended sediment concentration modeling". Journal of Hydrology. 535 457–472.

7 ANEXOS

7.1 Anexo 1: Tablas de Estadístico de rendimiento de resultados 1.

Tabla 7.1: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB1.

Crecida	COMB1					
	Prueba 1		Prueba 2		Prueba 3	
	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)
1	10.0	24.7	10.2	0.2	10.8	32.8
2	7.3	13.6	6.7	0.1	6.4	20.5
3	6.6	26.8	6.9	0.2	5.0	10.3
4	4.5	-0.8	4.0	-0.1	4.2	-8.8
5	3.6	-43.8	3.1	-0.3	2.5	-11.3
6	10.1	11.3	9.0	0.1	8.8	-1.0
7	4.0	-40.9	4.0	-0.5	3.2	-31.5
8	7.0	-12.3	5.8	-0.2	5.8	-20.4
9	5.1	0.3	5.3	-0.1	6.4	-16.1
10	8.7	-9.6	8.0	0.1	9.5	-15.4
11	3.6	6.8	4.2	-0.1	6.3	-3.4
12	16.7	62.3	14.5	0.5	14.7	53.8
13	10.3	29.8	8.4	0.2	8.1	9.7
14	5.1	3.0	5.3	0.0	4.9	8.0
15	12.3	43.4	8.4	0.3	9.3	30.4
16	9.7	34.2	9.4	0.3	9.7	27.3
17	4.9	-19.7	4.8	-0.2	6.5	-23.6
18	8.8	-26.7	8.5	-0.2	10.5	-45.6
19	3.9	-5.7	4.6	-0.2	6.0	-4.9
20	5.0	14.9	5.2	0.0	7.3	10.4
21	6.7	-35.7	8.4	-0.5	8.9	-54.4
22	8.1	27.5	7.7	0.1	7.9	5.8
23	8.2	31.8	6.2	0.1	8.5	15.7
24	5.9	-1.7	6.9	-0.3	8.0	-11.3
25	6.8	-41.3	6.7	-0.6	5.4	-41.6
26	8.5	25.4	8.5	0.0	9.2	19.4
27	5.4	-22.3	6.9	-0.3	8.4	-35.9

Tabla 7.2: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB2.

Crecida	COMB2					
	Prueba 4		Prueba 5		Prueba 6	
	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)
1	10.8	33.8	10.6	21.9	9.3	5.2
2	6.9	21.3	5.9	6.0	8.2	-22.7
3	5.4	10.4	4.4	7.0	5.7	18.4
4	6.3	-4.1	6.2	-9.8	6.8	-11.9
5	4.7	-6.4	5.5	-28.2	9.1	-97.2
6	10.5	11.4	11.3	7.8	11.0	7.9
7	3.1	-28.3	4.5	-47.4	6.2	-70.0
8	5.4	-6.9	7.4	-12.7	7.2	-23.8
9	7.4	6.5	7.8	-13.8	7.1	-6.1
10	8.4	-9.4	10.6	-16.2	8.0	-17.1
11	5.1	0.7	5.9	-8.2	5.1	-10.8
12	17.0	60.6	16.2	56.2	15.9	51.4
13	10.3	19.6	10.2	18.1	9.6	22.5
14	5.2	10.9	4.7	-1.7	4.9	-10.2
15	9.9	41.7	8.7	29.8	8.1	17.7
16	10.2	33.5	9.1	17.3	8.3	17.0
17	7.3	-16.0	8.2	-29.9	7.8	-36.6
18	9.8	-31.3	11.9	-57.5	11.1	-54.9
19	5.8	-15.7	5.4	23.8	5.2	16.6
20	7.0	-0.1	7.2	29.6	6.8	31.2
21	10.4	-51.7	7.2	-21.2	6.0	0.5
22	8.6	-14.6	8.0	27.8	12.6	51.2
23	7.8	15.9	9.6	51.3	9.9	43.5
24	7.7	-57.6	5.1	17.1	5.3	16.7
25	8.3	-71.8	6.1	5.2	8.0	6.1
26	9.1	21.9	10.5	51.0	10.8	33.8
27	9.8	-43.0	7.7	-7.2	7.3	22.0

Tabla 7.3: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB3.

Crecida	COMB3					
	Prueba 7		Prueba 8		Prueba 9	
	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)
1	10.2	22.6	10.5	25.1	9.9	20.5
2	6.8	14.7	6.4	16.8	6.9	11.5
3	4.7	5.9	4.2	3.4	6.8	5.1
4	6.9	-15.4	6.7	-13.8	9.5	-20.1
5	5.0	-18.3	4.5	-8.9	5.2	-21.7
6	11.5	-3.8	10.3	-2.1	15.5	-16.5
7	4.9	-49.8	4.1	-45.0	9.8	-60.7
8	7.0	-11.5	5.9	-15.8	14.0	-15.7
9	7.8	-1.9	7.2	-7.1	12.8	-13.6
10	11.0	-27.5	10.0	-27.8	19.8	-58.7
11	3.8	8.2	4.3	-6.5	4.2	14.4
12	18.1	62.7	16.8	57.9	18.1	62.3
13	10.7	23.4	9.3	13.3	10.7	27.2
14	4.9	-1.4	4.9	-1.0	4.5	-0.8
15	9.1	34.8	8.6	33.1	36.2	-14.7
16	9.2	28.1	9.0	26.2	11.6	21.7
17	8.0	-28.7	7.7	-27.2	8.7	-29.4
18	10.4	-39.7	10.2	-43.2	13.7	-56.6
19	5.1	-16.8	5.7	-28.1	9.1	-18.3
20	6.0	1.7	6.4	-9.5	10.0	1.8
21	9.5	-49.1	11.0	-69.3	15.1	-78.6
22	6.5	2.4	7.3	-16.8	13.1	-4.6
23	8.3	27.1	7.1	8.3	9.9	24.5
24	6.7	-23.6	7.3	-59.3	5.2	-12.0
25	8.2	-69.6	9.4	-90.3	10.0	-77.2
26	9.3	18.1	9.3	5.3	11.3	15.0
27	9.4	-50.9	10.8	-63.9	24.8	-112.4

Tabla 7.4: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB4.

Crecida	COMB4					
	Prueba 10		Prueba 11		Prueba 12	
	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)
1	10.0	20.9	10.2	22.8	9.8	17.0
2	7.9	10.0	7.4	12.3	7.1	4.4
3	8.8	27.5	6.1	14.5	5.3	11.9
4	8.4	3.3	6.8	-7.7	6.7	-8.1
5	5.9	-39.4	5.4	-27.5	5.9	-43.6
6	15.9	6.6	12.0	10.1	11.8	0.6
7	5.3	-53.3	4.3	-46.1	4.5	-52.5
8	9.3	-18.7	7.8	-8.4	8.3	-17.2
9	11.6	-10.7	7.9	3.6	9.0	-12.8
10	12.8	12.8	11.2	-0.5	9.4	-24.7
11	4.0	11.2	4.6	5.1	4.8	-1.6
12	18.3	52.2	17.3	58.3	17.3	51.1
13	13.0	27.1	11.3	23.7	8.6	5.3
14	4.8	2.3	4.9	-0.2	4.3	-2.0
15	8.6	27.2	9.7	33.7	9.0	26.1
16	12.3	42.7	10.4	36.4	8.8	25.1
17	11.9	-42.3	7.6	-22.7	8.9	-37.5
18	9.5	-13.5	10.4	-23.5	9.9	-48.6
19	6.1	13.5	6.0	-27.3	6.6	7.0
20	7.7	27.2	6.7	-3.3	8.2	14.8
21	10.2	-25.8	10.4	-43.6	9.0	-50.3
22	12.4	38.0	8.6	-1.6	7.9	8.3
23	11.7	24.1	7.7	15.1	11.1	16.4
24	9.1	4.9	7.2	-49.6	11.3	-22.4
25	7.5	13.2	10.3	-72.3	6.0	-37.7
26	12.9	8.1	9.3	9.0	10.9	25.9
27	12.4	17.4	12.3	-49.1	9.7	-18.6

Tabla 7.5: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB5.

Crecida	COMB5					
	Prueba 13		Prueba 14		Prueba 15	
	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)
1	10.1	23.3	10.0	16.1	10.9	30.7
2	7.5	15.5	6.6	-2.3	6.2	4.6
3	7.4	22.6	4.9	13.7	8.1	-2.6
4	6.8	4.9	6.5	-7.7	8.1	-8.9
5	5.1	-25.7	6.0	-51.5	4.9	-20.2
6	18.6	9.6	12.0	5.5	13.6	-17.2
7	4.1	-41.8	5.1	-57.9	3.8	-39.3
8	11.9	-4.6	8.8	-29.6	9.7	-34.0
9	17.7	-5.1	10.4	-9.3	11.6	-34.0
10	24.3	-35.2	11.1	-18.2	17.7	-26.5
11	6.3	-7.8	9.2	-39.7	7.8	-6.8
12	17.1	58.4	14.8	39.7	14.7	42.1
13	14.4	15.9	10.5	6.7	8.5	15.7
14	8.3	-3.8	4.9	-8.5	3.6	1.9
15	23.5	18.8	10.9	11.0	10.8	28.8
16	11.2	27.5	9.1	18.2	11.1	21.3
17	7.2	-18.7	8.2	-33.5	9.8	-33.3
18	18.2	-45.0	12.0	-56.3	12.4	-63.0
19	7.7	-23.1	4.9	-12.0	7.4	-3.4
20	8.5	-0.5	5.9	9.5	9.2	16.9
21	18.4	-81.5	8.0	-38.6	12.3	-41.3
22	17.3	-9.7	9.0	26.7	12.5	16.7
23	8.7	26.4	9.8	15.4	13.4	25.3
24	11.9	-7.3	6.3	-18.1	9.1	-3.3
25	8.3	-65.1	8.2	-27.1	7.2	-23.0
26	8.8	7.7	10.0	3.5	10.7	20.1
27	18.1	-44.9	8.9	-5.0	11.3	5.7

Tabla 7.6: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB6.

Crecida	COMB6					
	Prueba 16		Prueba 17		Prueba 18	
	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)
1	10.2	23.2	10.4	18.8	10.5	21.2
2	6.5	8.2	7.6	17.9	8.2	21.8
3	5.7	17.1	6.5	15.2	7.2	22.5
4	6.0	-3.2	7.5	-6.2	7.4	-0.1
5	5.3	-29.9	4.6	-19.0	4.3	-13.5
6	10.9	13.0	13.7	6.2	15.3	5.2
7	4.2	-43.8	4.7	-53.7	4.2	-48.0
8	6.8	-8.7	7.6	-14.4	7.7	-4.7
9	6.8	-2.2	9.0	-3.8	10.9	-7.7
10	7.4	-8.4	12.1	0.6	9.9	-3.6
11	5.0	-3.0	4.8	12.8	5.0	3.7
12	16.4	54.9	18.7	57.1	20.0	47.3
13	9.0	21.3	12.5	28.3	9.0	10.9
14	4.9	4.7	4.7	-2.4	5.0	-1.5
15	9.0	33.4	10.5	29.4	10.1	20.0
16	9.3	26.0	10.3	36.2	9.5	31.6
17	7.1	-20.7	9.2	-36.9	8.2	-29.4
18	9.6	-42.3	9.8	-28.9	9.8	-31.6
19	4.9	19.7	4.4	-3.5	5.6	4.6
20	6.6	29.5	5.8	14.9	7.1	20.9
21	6.6	-11.4	8.6	-25.1	8.2	-21.8
22	10.0	39.9	9.4	27.5	10.2	35.1
23	9.9	46.2	11.1	10.4	15.9	-1.5
24	4.8	14.0	7.4	-20.5	10.4	-21.7
25	6.2	1.6	9.2	-28.9	8.3	-10.8
26	11.0	36.7	11.1	13.1	15.4	3.8
27	6.7	9.3	15.5	-12.7	13.8	-10.4

Tabla 7.7: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB7.

Crecida	COMB7					
	Prueba 19		Prueba 20		Prueba 21	
	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)
1	10.2	21.4	10.4	28.1	8.8	14.5
2	7.1	15.6	7.7	6.9	9.7	-36.5
3	4.2	7.1	5.4	15.1	8.3	21.1
4	7.7	-15.3	7.4	-5.5	7.0	-2.4
5	4.8	-13.8	6.1	-42.6	11.4	-123.0
6	15.8	-16.4	12.1	8.6	11.3	11.4
7	5.8	-53.4	4.1	-38.9	5.6	-55.0
8	7.9	-23.6	7.0	-19.7	7.9	-7.2
9	10.0	-12.5	9.5	-10.6	6.9	-3.8
10	13.9	-34.9	10.2	-12.7	7.8	5.9
11	4.2	11.8	6.6	-10.4	4.6	2.3
12	19.5	55.2	16.8	40.8	15.7	46.4
13	9.7	22.8	10.7	8.2	9.1	28.8
14	4.4	-1.4	4.2	5.2	5.0	7.0
15	9.7	25.5	10.8	13.9	9.7	18.5
16	9.3	27.6	9.6	31.5	10.1	28.7
17	8.4	-29.8	8.4	-26.0	7.5	-20.9
18	9.9	-50.5	9.7	-44.5	10.3	-39.3
19	7.4	-16.2	5.9	3.2	5.6	13.6
20	8.4	2.3	7.6	16.6	6.6	21.0
21	9.8	-55.5	9.7	-38.1	10.5	-22.0
22	7.8	5.7	8.6	21.5	11.2	26.9
23	10.5	14.1	12.0	12.7	8.7	34.1
24	6.5	-41.8	9.4	-26.5	5.1	-9.0
25	8.4	-64.4	7.6	-2.4	8.2	-26.6
26	13.3	3.8	13.9	2.3	12.1	21.7
27	15.4	-69.5	11.4	-13.8	11.1	2.8

Tabla 7.8: Estadísticos de rendimiento de la estimación de ANN para la combinación de entrada COMB8.

Crecida	COMB8					
	Prueba 22		Prueba 23		Prueba 24	
	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)	RMSE (Mg/l)	PBIAS (%)
1	10.3	15.1	10.2	22.0	10.5	26.9
2	6.2	-5.1	6.8	10.6	6.7	6.6
3	8.7	14.1	5.2	8.6	4.9	13.8
4	5.5	-14.3	6.6	-13.9	7.4	-4.2
5	5.6	-44.7	5.3	-26.9	5.8	-37.0
6	11.1	3.8	10.4	3.3	12.9	8.8
7	6.1	-64.0	4.2	-47.1	5.0	-46.1
8	8.3	-14.4	6.7	-11.6	5.9	-10.0
9	8.4	-14.9	7.3	2.8	7.8	-5.9
10	10.5	18.9	9.2	-15.3	11.4	-25.2
11	4.8	-10.1	3.6	-0.7	5.2	-2.5
12	15.5	46.9	19.0	52.9	16.2	54.1
13	9.7	25.0	10.5	22.3	12.3	15.6
14	5.9	0.9	5.1	-0.7	5.1	7.3
15	9.9	28.2	9.4	35.2	9.6	29.7
16	11.0	31.7	9.5	28.6	10.3	29.8
17	7.0	-18.3	7.7	-28.3	8.6	-25.8
18	11.2	-25.8	10.4	-36.7	10.0	-40.6
19	5.9	0.0	4.5	-16.1	5.6	4.1
20	7.0	15.8	5.3	4.5	7.5	17.5
21	9.6	-26.7	9.1	-37.9	9.3	-26.2
22	10.1	21.2	7.3	10.0	7.4	21.7
23	9.4	15.5	7.8	21.3	9.3	42.9
24	6.4	-28.5	7.0	-25.8	4.7	6.6
25	6.7	-26.2	7.8	-57.4	8.0	-15.4
26	11.4	13.8	10.0	6.1	10.3	28.7
27	9.5	0.5	8.0	-29.8	8.5	-22.4

7.2 Anexo 2: Grafica de datos estimados por el modelo de ANN vs los datos reales de Css para cada una de las crecidas

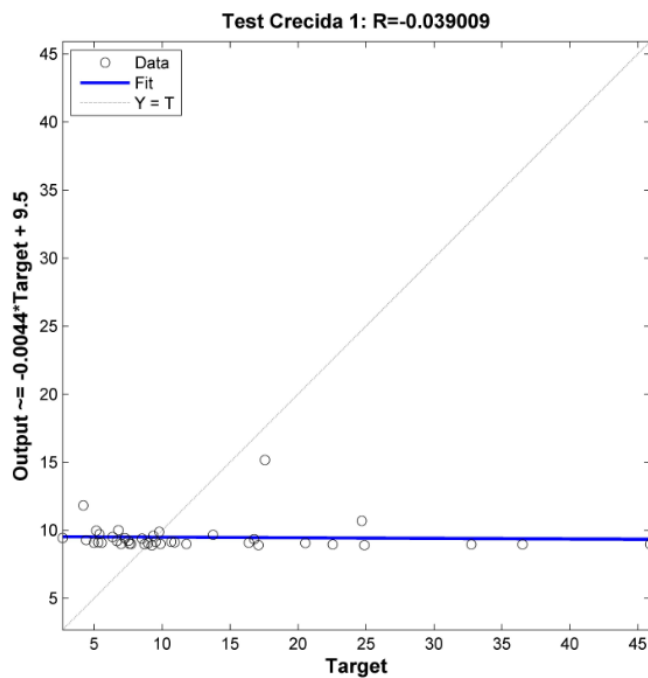


Figura 7.1: Datos estimados vs datos reales de Css (mg/l) de la crecida 1.

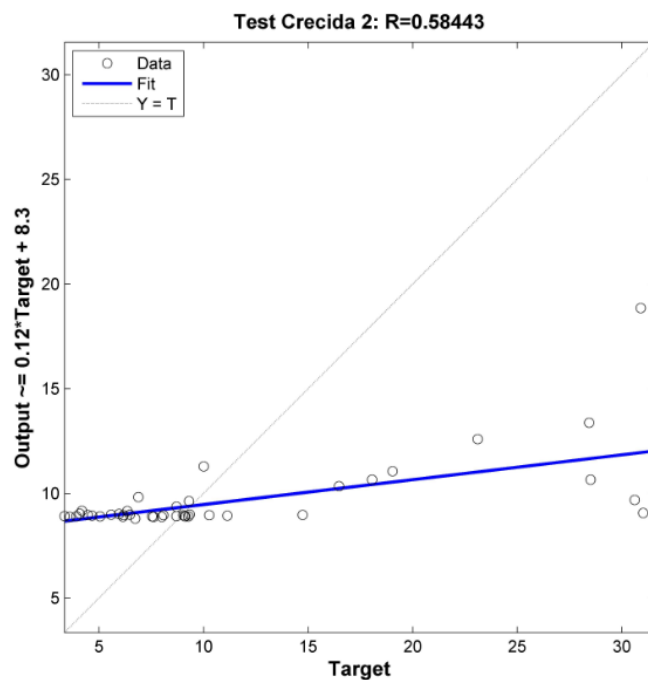


Figura 7.2: Datos estimados vs datos reales de Css (mg/l) de la crecida 2.

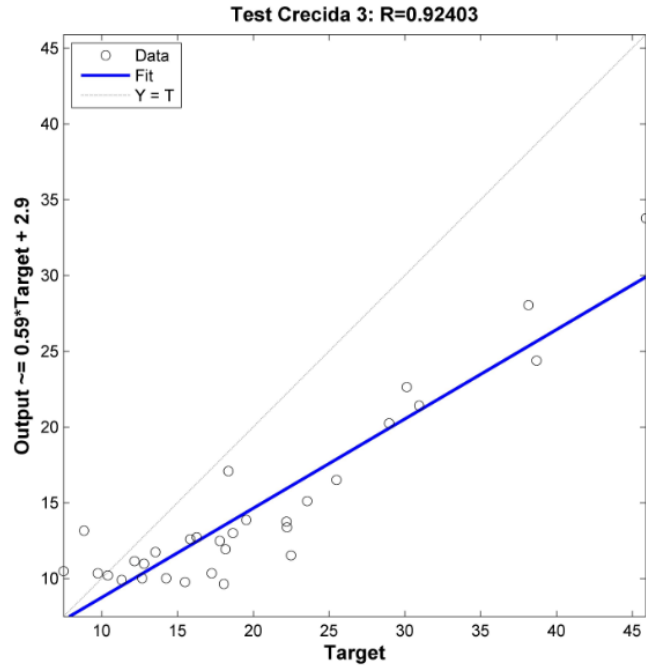


Figura 7.3: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 3.

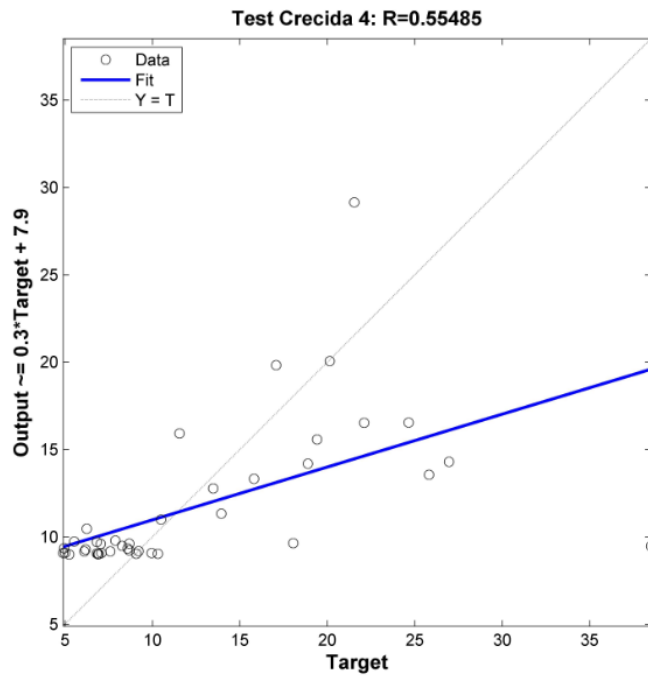


Figura 7.4: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 4.

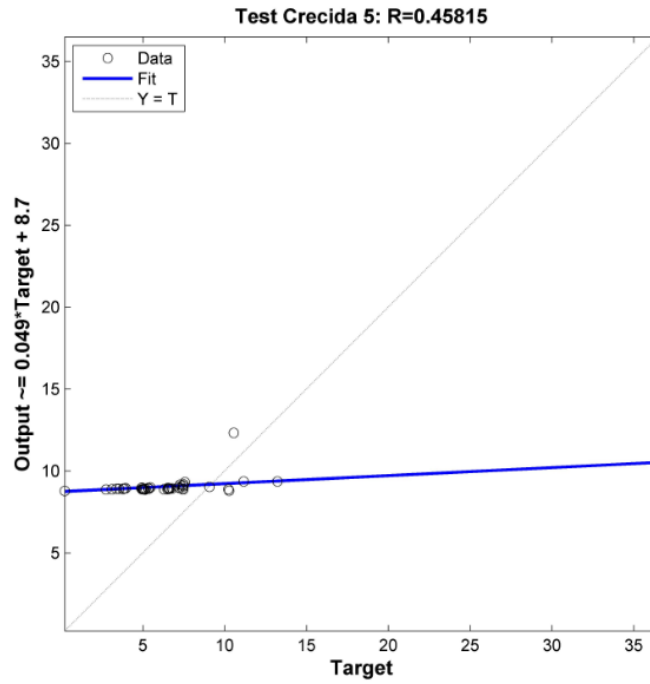


Figura 7.5: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 5.

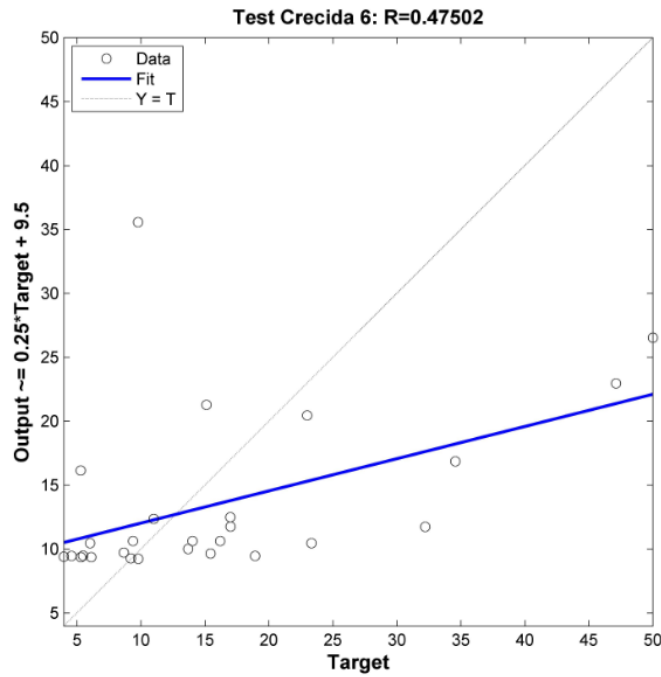


Figura 7.6: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 6.

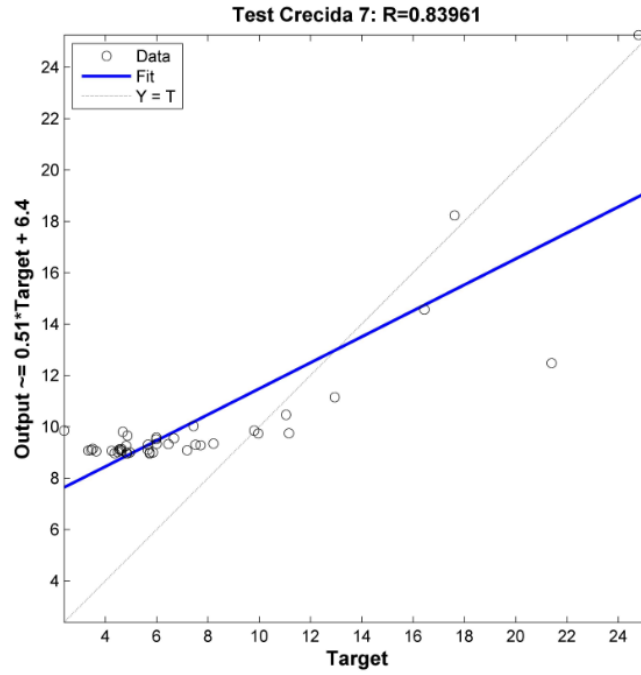


Figura 7.7: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 7.

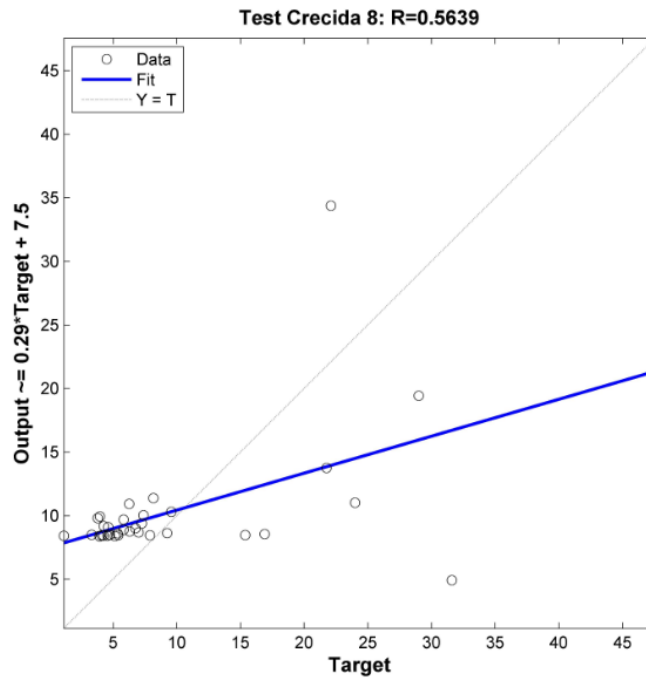


Figura 7.8: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 8.

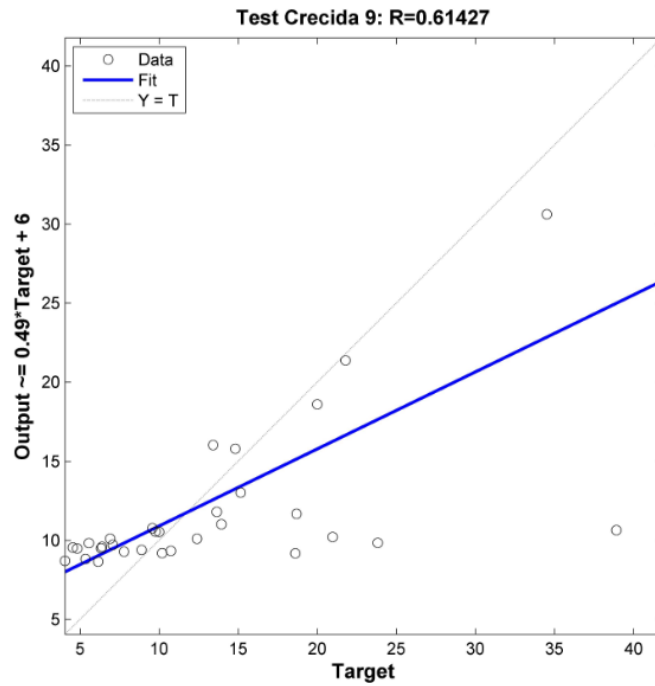


Figura 7.9: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 9.

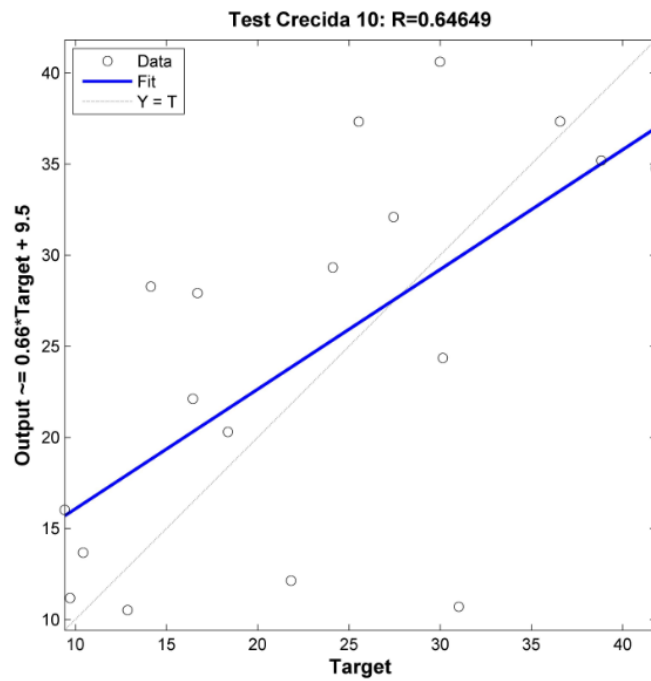


Figura 7.10: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 10.

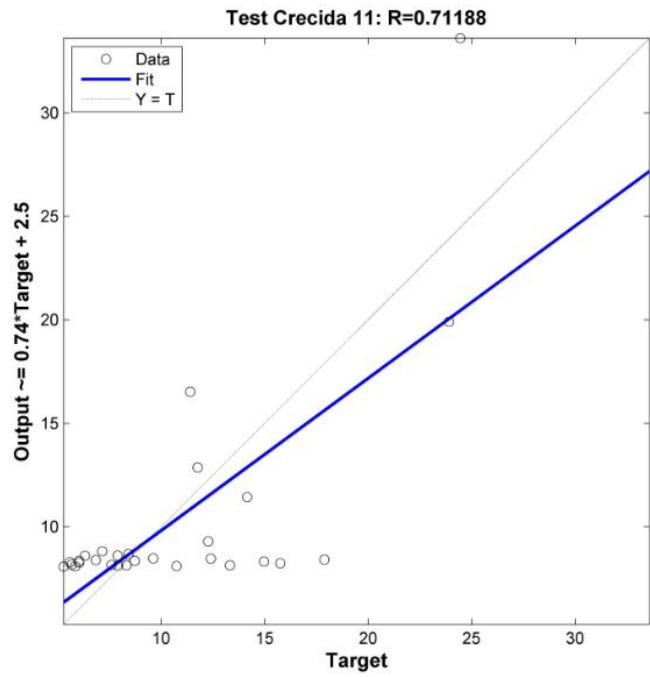


Figura 7.11: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 11.

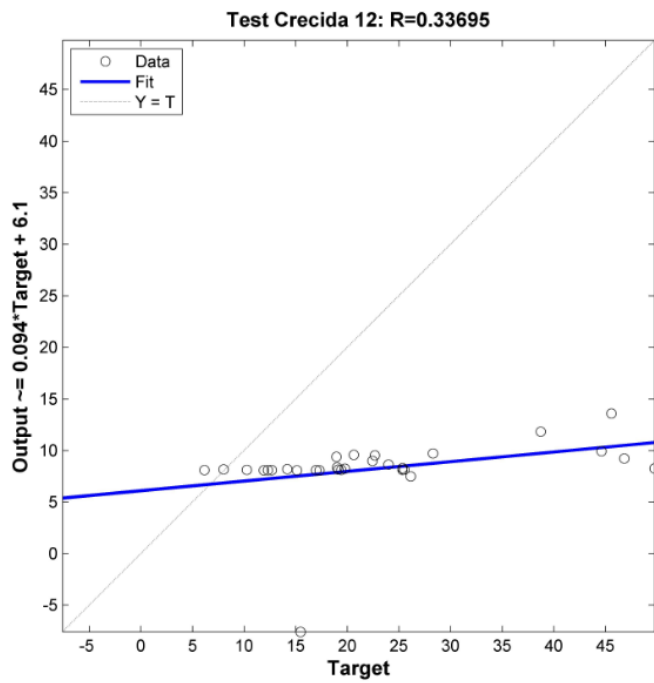


Figura 7.12: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 12.

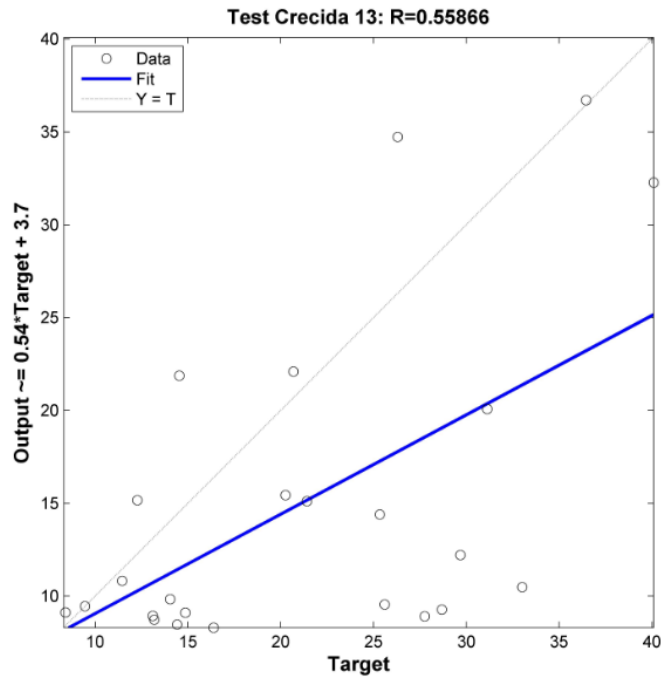


Figura 7.13: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 13.

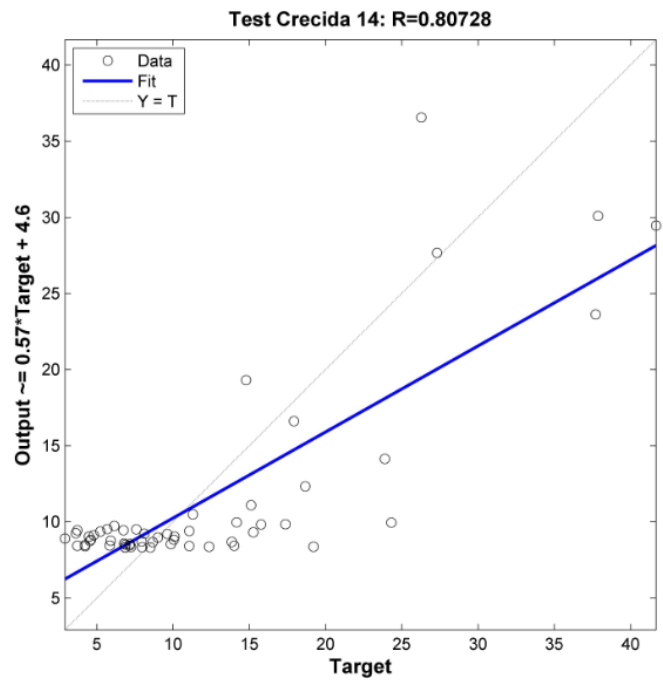


Figura 7.14: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 14.

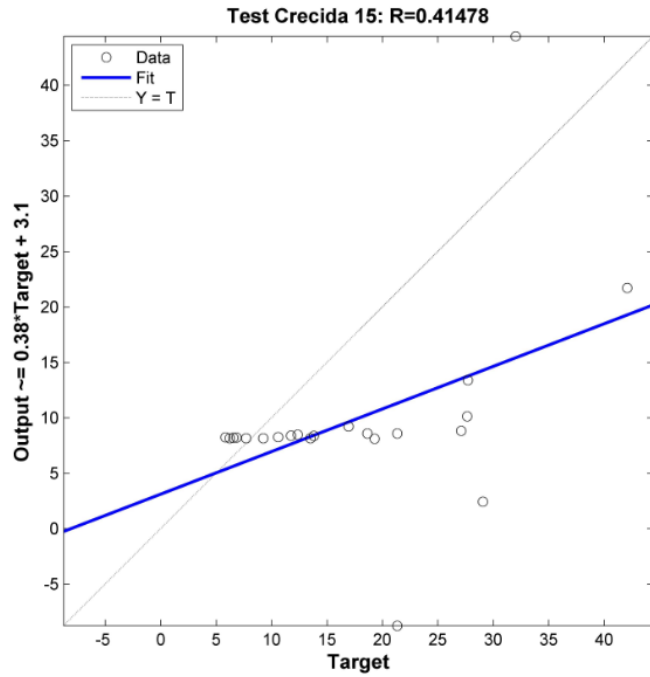


Figura 7.15: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 15.

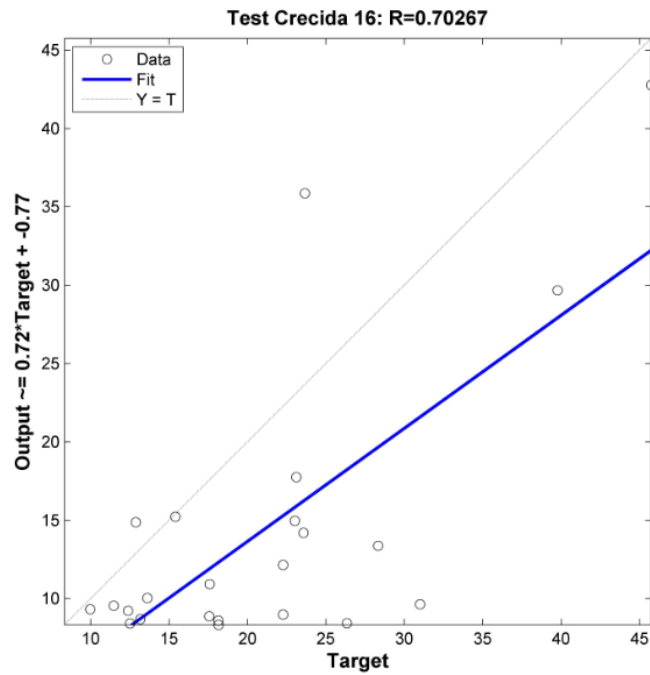


Figura 7.16: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 16.

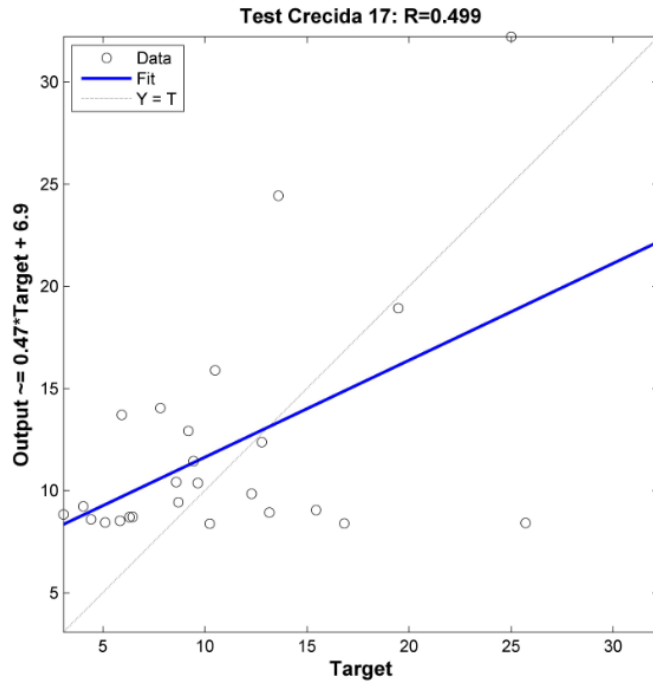


Figura 7.17: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 17.

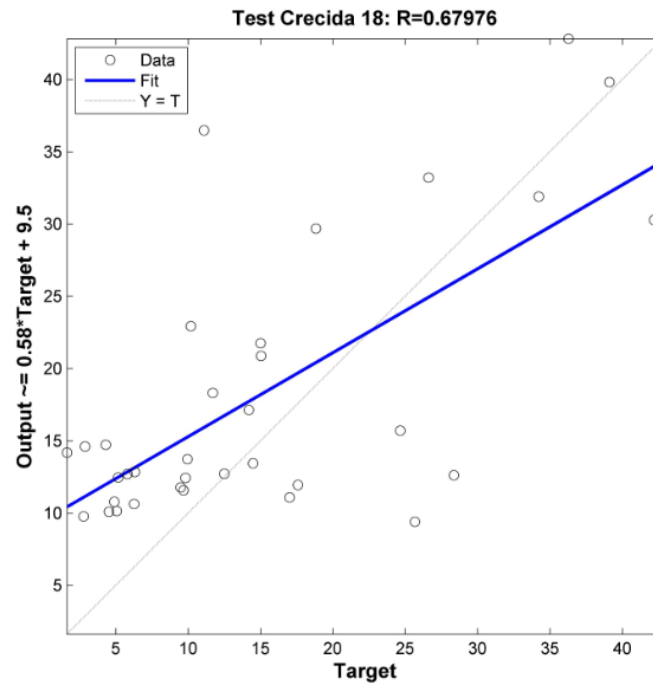


Figura 7.18: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 18.

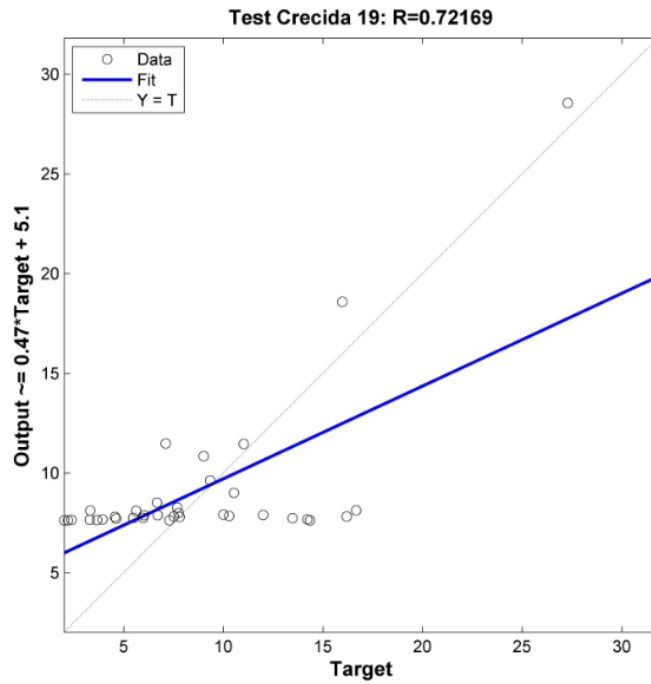


Figura 7.19: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 19.

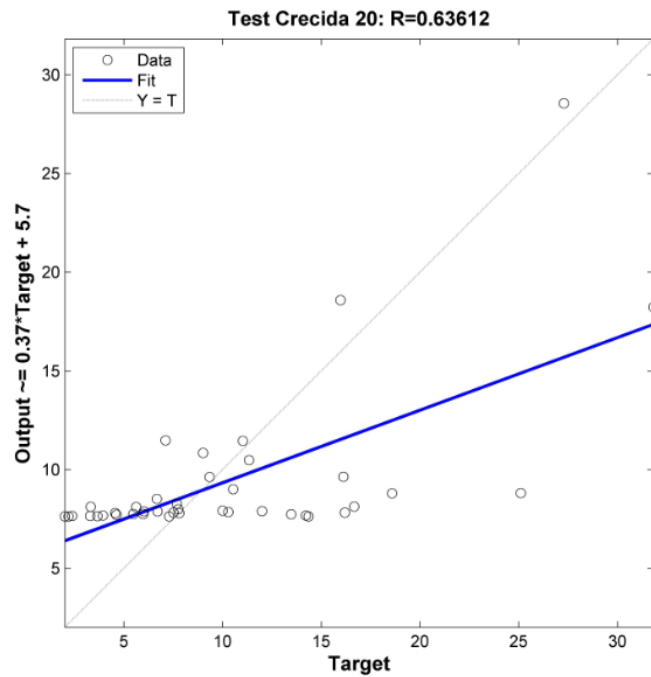


Figura 7.20: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 20.

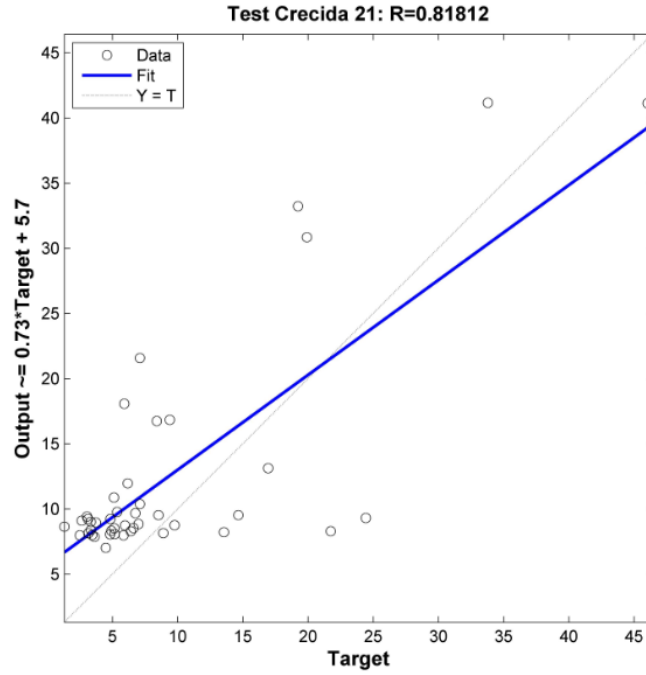


Figura 7.21: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 21.

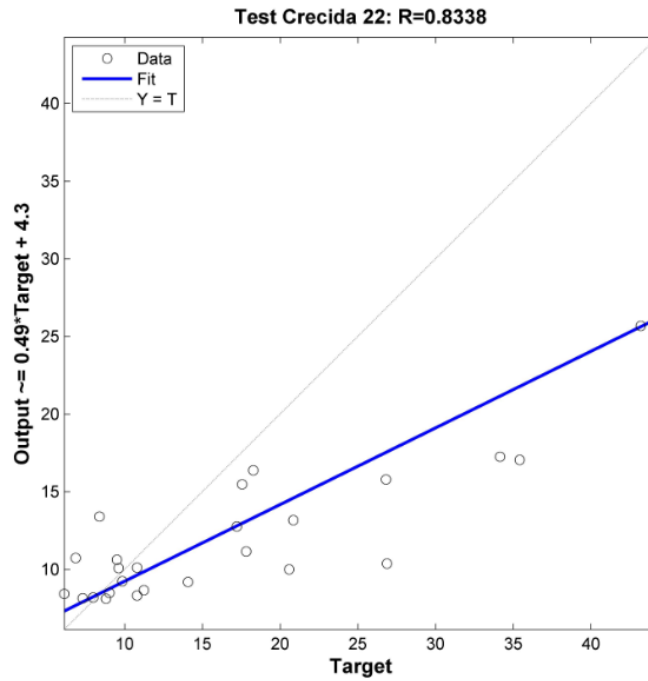


Figura 7.22: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 22.

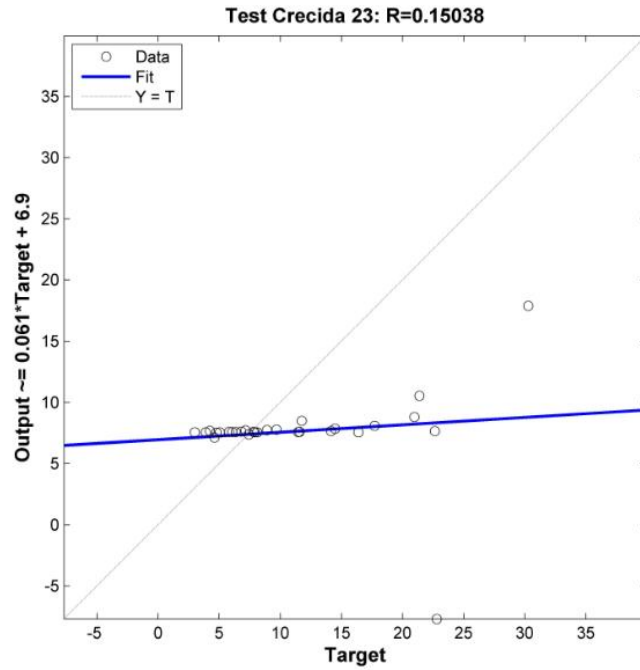


Figura 7.23: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 23.

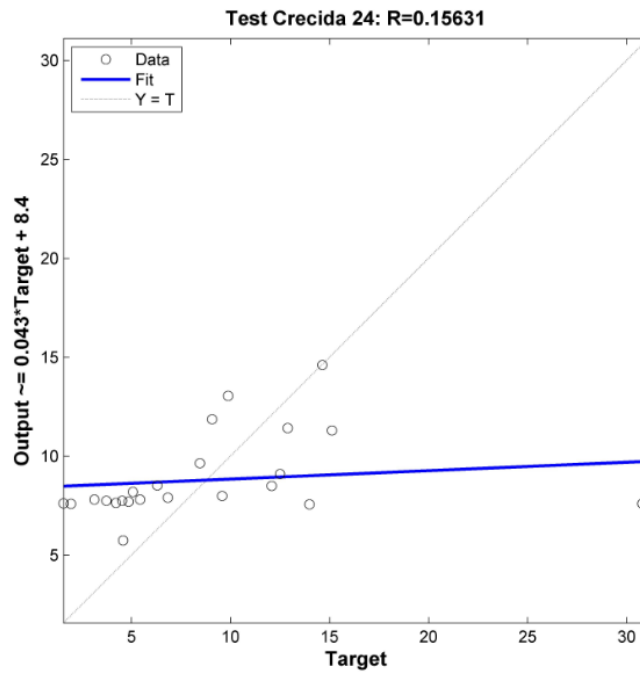


Figura 7.24: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 24.

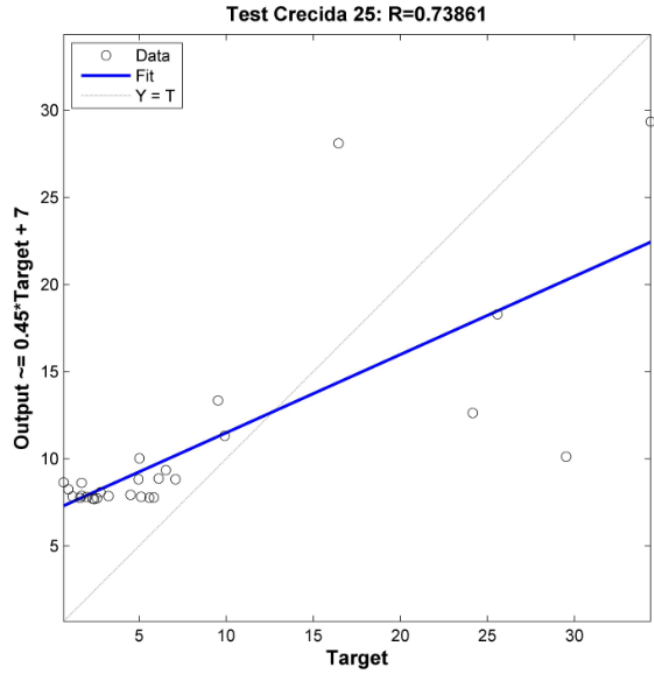


Figura 7.25: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 25.

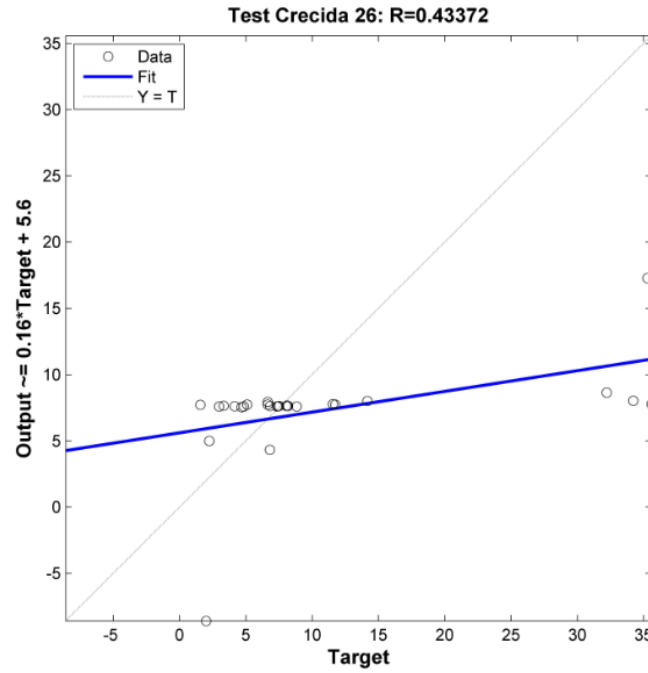


Figura 7.26: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 26.

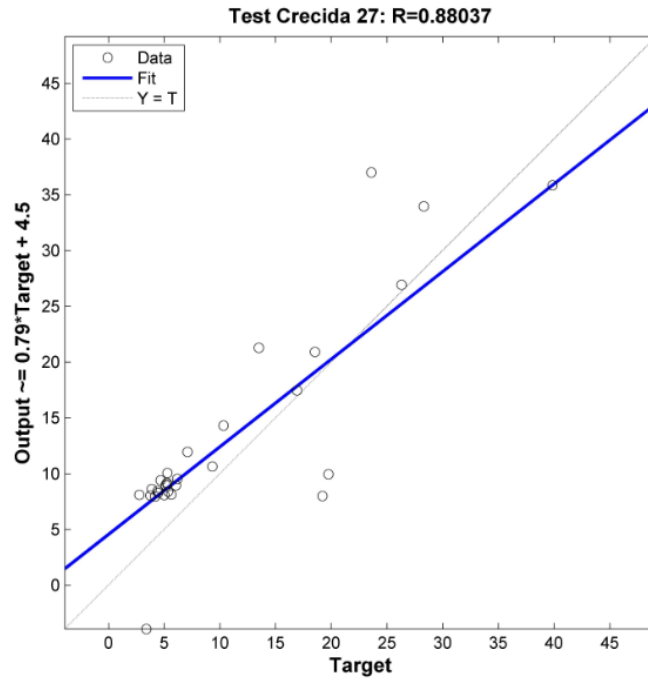


Figura 7.27: Datos estimados vs datos reales de C_{ss} (mg/l) de la crecida 27.

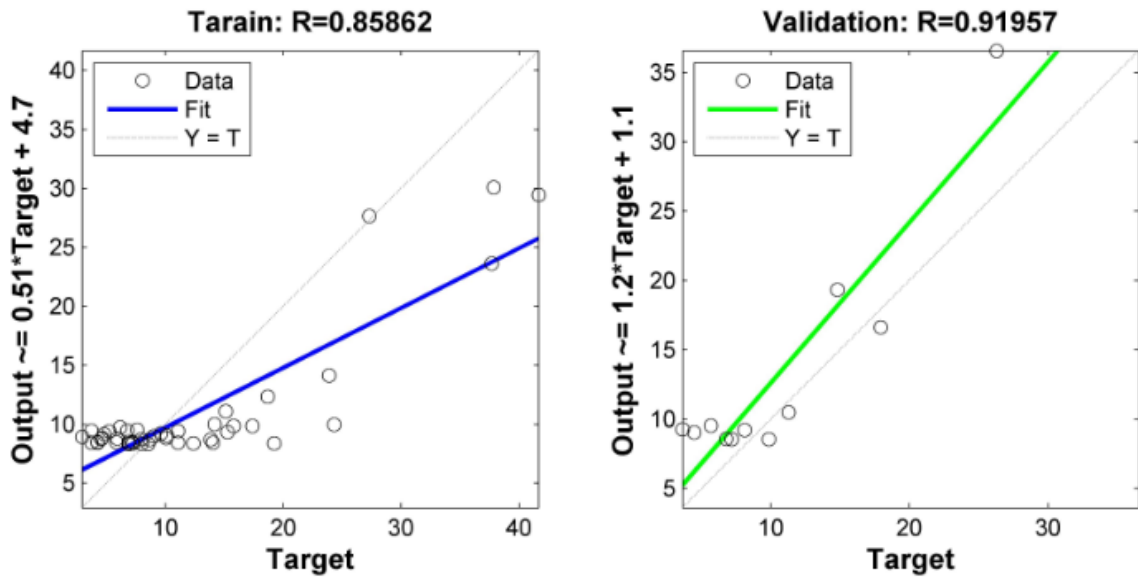


Figura 7.28: Datos estimados vs datos reales de C_{ss} (mg/l) para los grupos de datos de entrenamiento y validación de la crecida 14.