

ANDREA CAROLINA AGUILAR AGUILAR



ANÁLISIS COMPARATIVO PARA LA APLICACIÓN DE  
UN MODELO PREDICTIVO EN LA ESTIMACIÓN DE LA  
CALIDAD DE AGUA

ANEXOS

Universidad del Cauca  
Facultad de Ingeniería Electrónica y Telecomunicaciones  
Maestría en Electrónica y Telecomunicaciones

Popayán  
2021

## Tabla de Contenido

Anexo A .....	6
Análisis de Correlación de Parámetros y Significancia Estadística.....	6
A.1. Correlación de Parámetros .....	6
A.2. Significancia Estadística .....	8
Anexo B .....	10
Entrenamiento de Modelos de Regresión .....	10
B.1. Aplicación Regression Learner de Matlab.....	10
Anexo C .....	14
Estimación de Parámetros .....	14
C.1. Resultados Estimación Muestra AL-2 .....	14
C.2. Resultados Estimación Muestra AL-4 .....	16
C.3. Resultados Estimación Muestra AA-2.....	17
C.4. Resultados Estimación Muestra AA-4.....	19
C.5. Resultados Estimación Muestra AE-1 .....	20
C.6. Resultados Estimación Muestra AE-3.....	22
Anexo D .....	24
Modelo Estimación del ICA y Procedimiento de Validación.....	24
D.1. Características del Modelo para la Estimación del ICA.....	24
D.2. Procedimiento para la Estimación de Nuevos Datos.....	25
D.3. Estimación del ICA sin el parámetro temperatura .....	27

## Lista de Figuras

Figura B.1. Diagrama de flujo herramienta de aprendizaje automático - Matlab...	10
Figura B.2. Interfaz de Matlab – APPS Regression Learner.....	11
Figura B.3. Selección de datos – APPS Regression Learner. ....	11
Figura B.4. Exploración de datos – Interfaz APPS Regression Learner. ....	12
Figura B.5. Selección de modelo – Interfaz APPS Regression Learner. ....	12
Figura B.6. Resultados de simulación – Interfaz APPS Regression Learner.....	13
Figura B.7. Exportar modelo interfaz -Matlab. ....	13
Figura C.1.Estimación pH modelo AR muestra AL-2: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.....	15
Figura C.2.Estimación TDS modelo AR muestra AL-2: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.....	15
Figura C.3.Estimación pH modelo AR muestra AL-4: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.....	16
Figura C.4.Estimación TDS modelo AR muestra AL-4: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.....	17
Figura C.5.Estimación pH modelo AR muestra AA-2: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.....	18
Figura C.6.Estimación TDS modelo AR muestra AA-2: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.....	18
Figura C.7.Estimación pH modelo AR muestra AA-4: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.....	19
Figura C.8.Estimación TDS modelo AR muestra AA-4: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.....	20
Figura C.9.Estimación pH modelo AR muestra AE-1: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.....	21
Figura C.10.Estimación TDS modelo MVS muestra AE-1: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados. ....	21
Figura C.11.Estimación pH modelo AR muestra AE-3: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.....	22
Figura C.12.Estimación TDS modelo MVS muestra AE-3: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados. ....	23
Figura D.1. Arquitectura del modelo (AR-AA1) Estimación ICA.....	24
Figura D.2.Modelo AR-AA1 para validación. ....	25
Figura D.3.Parámetros de entrada modelo AR-AA1.....	25
Figura D.4.Parámetros de entrada modelo AR-AA1.....	26
Figura D.5.Código interpolación de parámetros. ....	26

Figura D.6. Estimación ICA modelo AR- muestra AL-3 .....	27
Figura D.7. Estimación ICA modelo AR- muestra AA-1 .....	28
Figura D.8. Estimación ICA modelo RL- muestra AE-2 .....	29

## Lista de Tablas

Tabla A.1. Coeficientes de correlación – muestras agua lluvia.....	6
Tabla A.2. Coeficientes de correlación – muestras Agua de acueducto .....	7
Tabla A.3. Coeficientes de correlación – muestras de agua envasada. ....	7
Tabla A.4. Significancia estadística valor p – parámetros pH y TDS. ....	8
Tabla C.1. Resultados estimación pH y TDS - muestra AL-2. ....	14
Tabla C.2. Resultados estimación pH y TDS - muestra AL-4. ....	16
Tabla C.3. Resultados estimación pH y TDS - muestra AA-2. ....	17
Tabla C.4. Resultados estimación pH y TDS - muestra AA-4. ....	19
Tabla C.5. Resultados estimación pH y TDS - muestra AE-1. ....	20
Tabla C.6. Resultados estimación pH y TDS – muestra AE-3. ....	22
Tabla D.1. Resultados estimación ICA – muestra AL-3. ....	27
Tabla D.2. Resultados estimación ICA – muestra AA-1. ....	28
Tabla D.3. Resultados estimación ICA – muestra AE-2. ....	28

## Anexo A

### Análisis de Correlación de Parámetros y Significancia Estadística

#### A.1. Correlación de Parámetros

El análisis de correlación se aplicó a las muestras restantes de agua lluvia (AL-1, AL-2 y AL-4) para obtener el coeficiente de correlación entre los parámetros pH, TDS y temperatura. El coeficiente se da en un rango de -1 a +1, donde un valor de 0 indica que no existe asociación entre los parámetros y valores cercanos a +1 y -1 una asociación positiva y negativa respectivamente. Los resultados de correlación se presentan en la Tabla A.1.

Tabla A.1. Coeficientes de correlación – muestras agua lluvia

AL-1			
	pH	TDS	T
pH	1		
TDS	0.9786	1	
T	-0.0323	-0.0437	1

AL-2			
	pH	TDS	T
pH	1		
TDS	0.9681	1	
T	-0.0646	-0.0299	1

AL-4			
	pH	TDS	T
pH	1		
TDS	0.7285	1	
T	-0.0814	-0.0293	1

En los resultados de la Tabla A.1 se observa que existe una fuerte correlación positiva entre los parámetros pH y TDS, con valores positivos más cercanos a 1 en las muestras AL-1 y AL-2 y de 0.73 para AL-4 y una asociación inversa más pequeña de estos parámetros respecto a la temperatura.

Se analizó la correlación de parámetros sobre las muestras de agua de acueducto AA-2, AA-3, AA-4. Los resultados se muestran en la Tabla A.2.

Tabla A.2. Coeficientes de correlación – muestras Agua de acueducto

AA-2			
	<b>pH</b>	<b>TDS</b>	<b>T</b>
pH	1		
TDS	0.9634	1	
T	-0.0932	-0.0322	1
AA-3			
	<b>pH</b>	<b>TDS</b>	<b>T</b>
pH	1		
TDS	0.9775	1	
T	-0.0830	-0.0280	1
AA-4			
	<b>pH</b>	<b>TDS</b>	<b>T</b>
pH	1		
TDS	0.9396	1	
T	-0.0945	-0.0274	1

De acuerdo a los resultados se puede apreciar una fuerte correlación entre el pH y TDS en las muestras con valores superiores a 0.94 y una asociación negativa (inversa) respecto a temperatura también en las tres muestras. El análisis de correlación se aplicó en las muestras restantes de agua envasada, los resultados aparecen en la Tabla A.3.

Tabla A.3. Coeficientes de correlación – muestras de agua envasada.

AE-1			
	<b>pH</b>	<b>TDS</b>	<b>T</b>
pH	1		
TDS	0.9624	1	
T	-0.0836	-0.0408	1

AE-3			
	pH	TDS	T
pH	1		
TDS	0.9589	1	
T	-0.1227	-0.0411	1
AE-4			
	pH	TDS	T
pH	1		
TDS	0.9482	1	
T	-0.0670	-0.0381	1

Las muestras de agua envasada presentan un comportamiento similar al de agua lluvia y de acueducto, con correlaciones positivas cercanas a 1 entre los valores de pH y TDS y una relación negativa inversa de estos parámetros con la temperatura.

## A.2. Significancia Estadística

Para determinar si existe una relación estadísticamente significativa entre los parámetros pH y TDS se utiliza el *valor p* (*p-value*) cuyo valor oscila entre 0 y 1, si el valor *p* cumple con la condición de ser menor que un nivel de significancia, este se considera como un resultado estadísticamente significativo y, por lo tanto, permite rechazar la hipótesis nula. Se estableció entonces un nivel de significancia de 5% (0.05) y la hipótesis nula ( $H_0$ ) y alterna ( $H_1$ ):

$H_0$ : No existe relación estadísticamente significativa entre los parámetros.

$H_1$ : Existe relación significativa entre los parámetros.

La Tabla A.4 contiene los resultados de los *valores p* correspondientes en las muestras restantes de agua lluvia, agua de acueducto y envasada.

Tabla A.4. Significancia estadística valor *p* – parámetros pH y TDS.

Muestra	Valor -p
AL-1	3.82E-69
AL-2	1.08E-60
AL-4	8.75E-18
AA-2	7.90E-58
AA-3	5.12E-68
AA-4	2.06E-47



AE-1	2.74E-57
AE-3	2.14E-55
AE-4	1.28E-50

Como se aprecia en la Tabla A.4, el valor-p obtenido en cada muestra para la correlación de los parámetros pH y TDS es menor al nivel de significancia de 0.05, por lo que se descarta la hipótesis nula  $H_0$ .

## Anexo B

# Entrenamiento de Modelos de Regresión

## B.1. Aplicación Regression Learner de Matlab

La aplicación *Regression Learner* está diseñada para entrenar modelos de regresión mediante técnicas como la regresión lineal, árboles de regresión, modelos de regresión de procesos gaussianos y máquinas de vectores de soporte. Permite además explorar los datos, seleccionar características, especificar esquemas de validación, evaluar resultados almacenar y exportar un modelo al espacio de trabajo para usarlo con nuevos datos. El diagrama de flujo de la Figura B.1 representa el flujo de trabajo que se lleva a cabo en la herramienta.

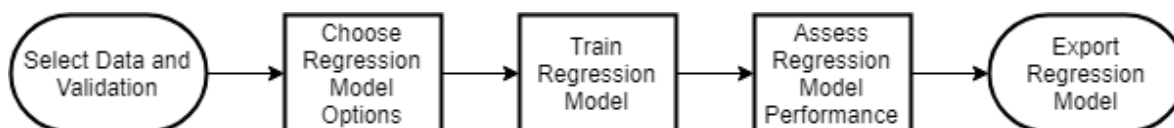


Figura B.1. Diagrama de flujo herramienta de aprendizaje automático - Matlab.

De acuerdo al diagrama, se elige entre varios algoritmos (técnicas) para entrenar y validar los modelos de regresión. Después de entrenar los modelos, se comparan sus errores de validación en paralelo y posteriormente se selecciona el mejor modelo según sus valores de desempeño. Para iniciar el entrenamiento en la interfaz de Matlab, se selecciona en la pestaña *APPS* la opción *Regression Learner* como en Figura B.2.

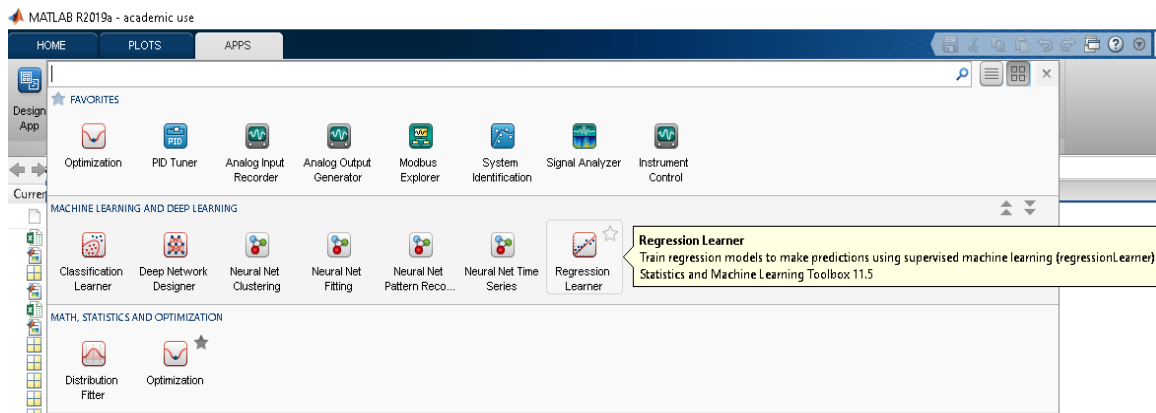


Figura B.2. Interfaz de Matlab – APPS Regression Learner.

En *Regression learner* se selecciona la opción *New Session* (Figura B.3) en cuya ventana se cargan los datos (variable o tabla) del espacio de trabajo. En este caso se utilizará una tabla con los valores de las muestras de agua lluvia, con los datos de pH TDS y temperatura por columna.

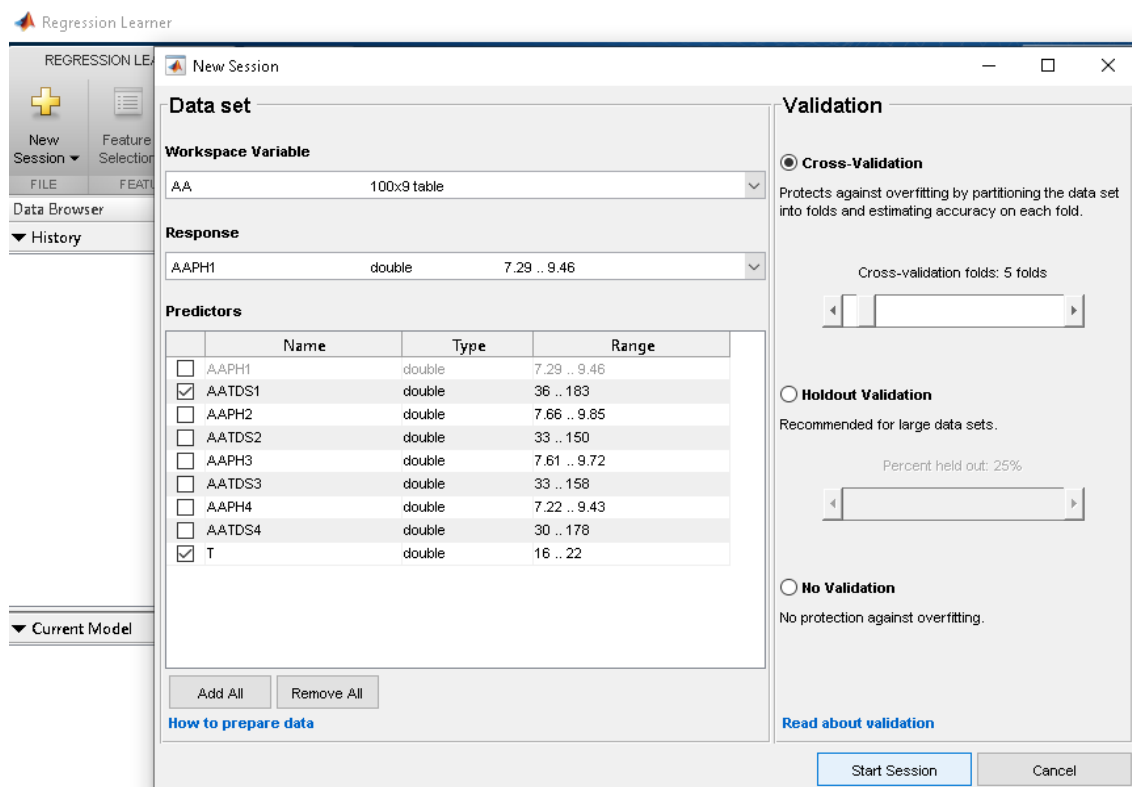


Figura B.3. Selección de datos – APPS Regression Learner.

En la opción *Response* se selecciona la variable que se desea estimar o variable de salida y en *Predictors* las variables que servirán como predictores, también se establece el tipo validación, en este caso una validación cruzada de 5 iteraciones. Una vez se encuentre listo se presiona *Start Session*. La herramienta grafica los

valores reales de la variable de salida que se desea estimar en este caso pH. Figura B.4.

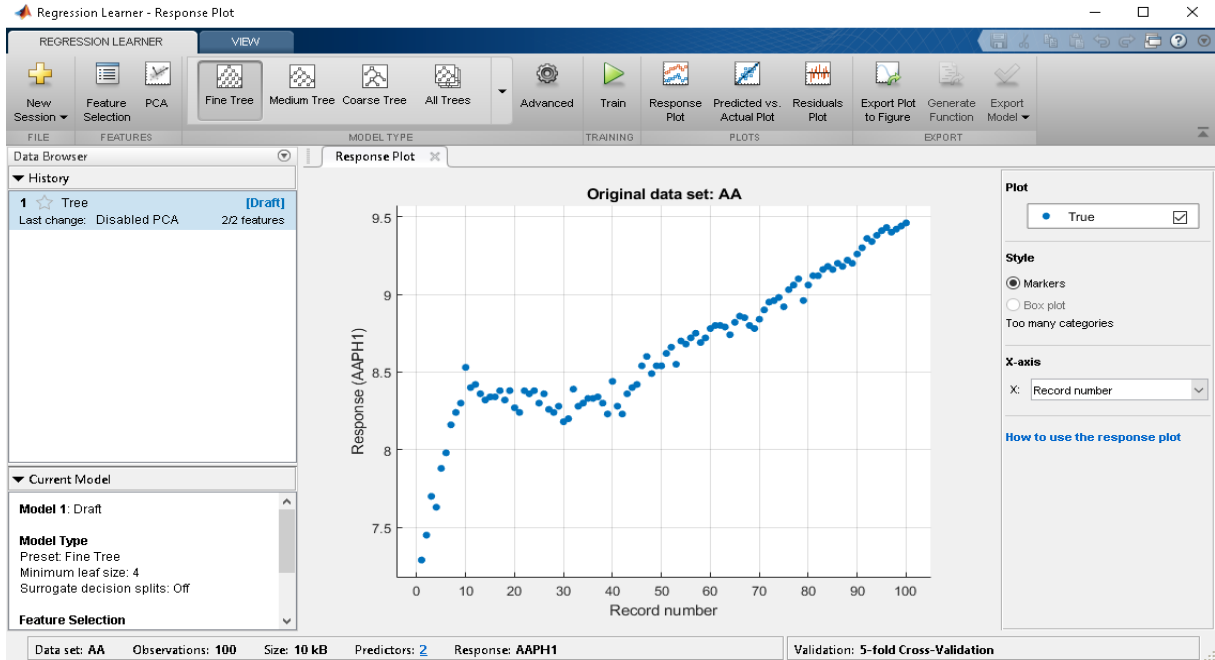


Figura B.4. Exploración de datos – Interfaz APPS Regression Learner.

En la pestaña *Regression Learner*, en la sección *Model Type* haciendo clic en la flecha para expandir, se encuentra la lista de técnicas de regresión para escoger el de preferencia, se selecciona el tipo de técnica a utilizar en este caso un árbol de regresión y luego se pulsa en *Train*. Figura B.5.

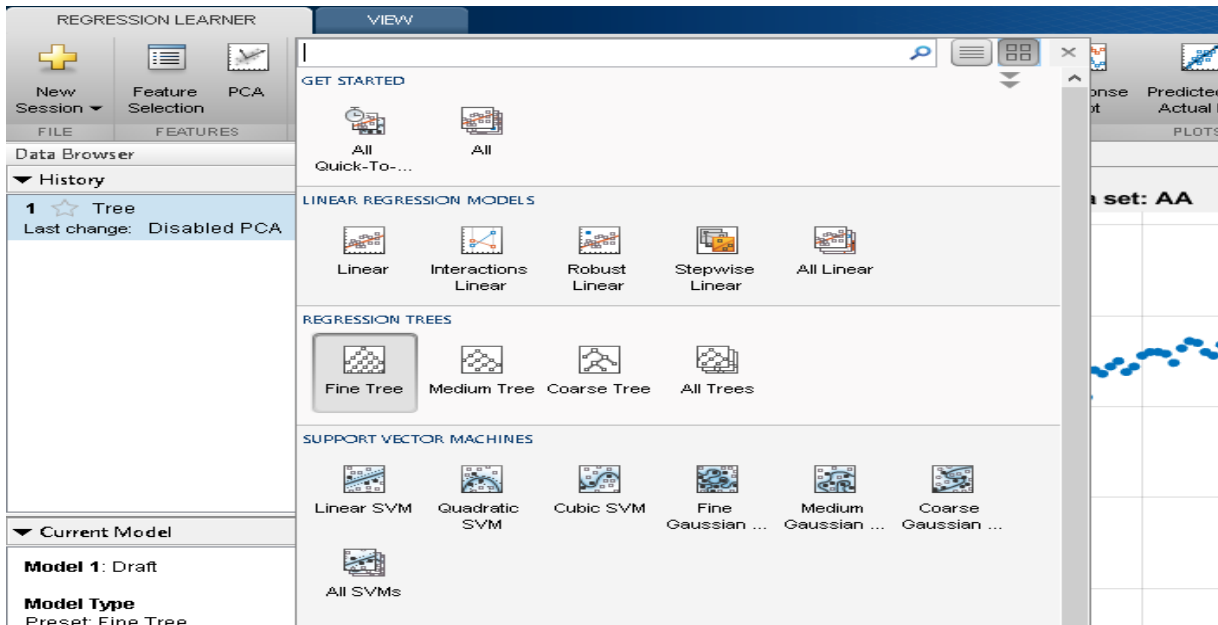


Figura B.5. Selección de modelo – Interfaz APPS Regression Learner.

Una vez pasado el tiempo de simulación, en la pestaña de *Response Plot* se grafica con puntos de color amarillo los valores estimados y los errores con líneas rojas, la aplicación tiene otras opciones graficas y de edicion de figuras. En el recuadro *Current Model* aparecen los resultados de las metrica como RMSE,  $R^2$ , MSE, MAE, y los tiempos de simulación. Es posible simular otros modelos sin iniciar una nueva session utilizando otras tecnicas. En el recuadro *History* aparece el resumen de cada simulación, Figura B. 5. La aplicación resalta el modelo de mejores resultados.

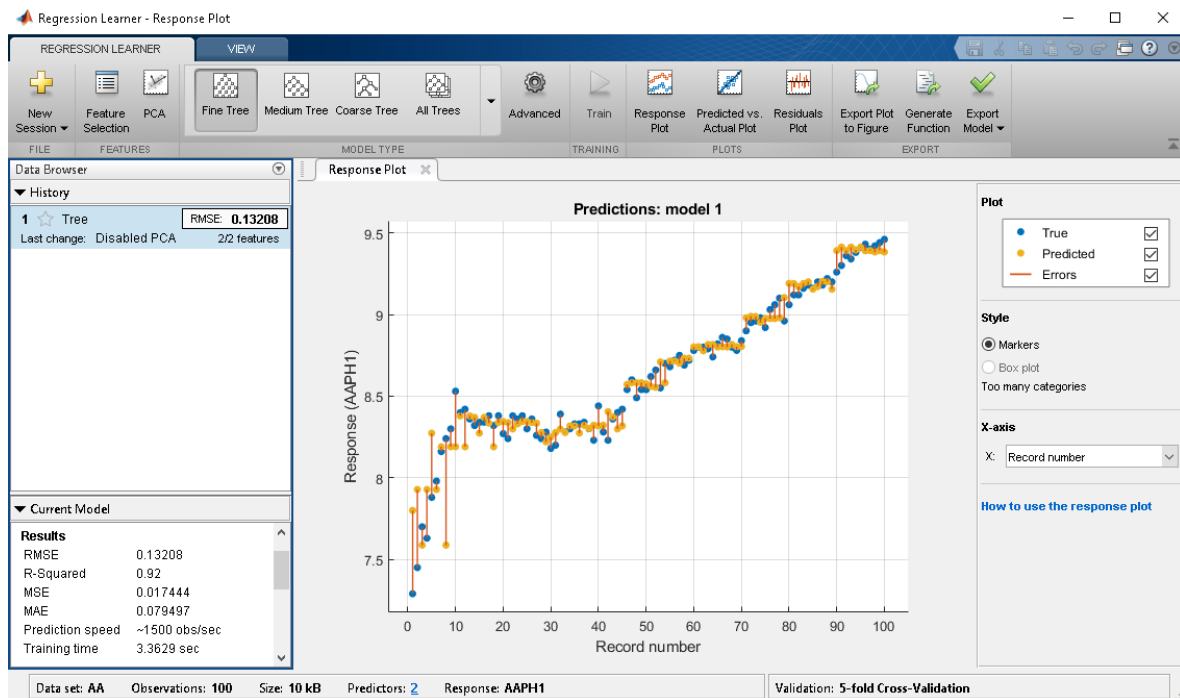


Figura B.6. Resultados de simulación – Interfaz APPS Regression Learner.

Para guardar un modelo en la barra de herramientas de la interfaz, la opción *Export Model* (Figura B.7) permite guardar el modelo que se almacenara en el espacio de trabajo de Matlab y poder utilizarlo así más adelante con nuevos datos.

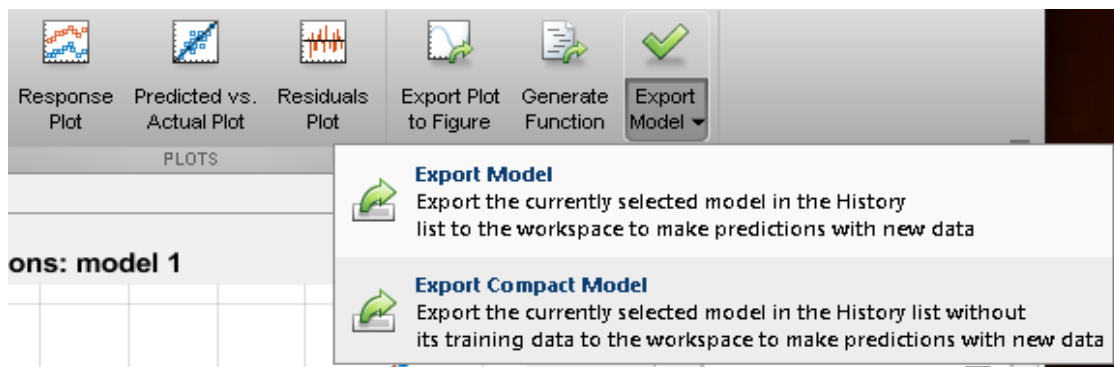


Figura B.7. Exportar modelo interfaz -Matlab.

## Anexo C

### Estimación de Parámetros

Para cada caso de estudio agua lluvia (AL), agua de acueducto (AA) y agua envasada (AE) se analizaron cuatro muestras, las muestras AL-3, AA-1 y AE-2 se seleccionaron al azar para presentar los resultados en la estimación de parámetros de pH, TDS en el documento principal. Las muestras AL-1, AA-3 y AE-4 se destinaron para realizar la validación del modelo de regresión en el capítulo 4, por lo que las muestras restantes de cada clase se utilizaron para replicar el procedimiento de estimación de parámetros de pH y TDS a fin de brindar mayor confiabilidad a los resultados.

#### C.1. Resultados Estimación Muestra AL-2

Mediante la aplicación *Regression Learner* de Matlab se utilizaron las muestras de agua lluvia AL-2 y AL-4 para estimar los parámetros de pH utilizando como entrada los parámetros de TDS y temperatura. Para la estimación del TDS las entradas fueron el pH y temperatura. Se compararon los modelos de regresión lineal, árboles de regresión y máquinas de vectores de soporte. Los resultados obtenidos se presentan en la Tabla C.1.

Tabla C.1. Resultados estimación pH y TDS - muestra AL-2.

Parámetro	pH			TDS		
	RL	AR	MVS	RL	AR	MVS
MAE	0.1612	0.0932	0.1601	19.493	12.538	19.647
RMSE	0.1965	0.1202	0.1981	23.239	17.835	23.567
R <sup>2</sup>	0.94	<b>0.98</b>	0.94	0.94	<b>0.96</b>	0.94

De acuerdo a los valores de la Tabla C.1, se puede observar que en la estimación del pH el modelo de árboles de regresión presenta valores pequeños de error en las métricas MAE y RMSE y un ajuste de los datos estimados a los reales de un 0.98 para R<sup>2</sup>. De manera similar en la estimación del TDS, el mejor modelo fue el AR.

Las representaciones graficas de estos modelos se presentan en las Figuras C.1 y C.2.

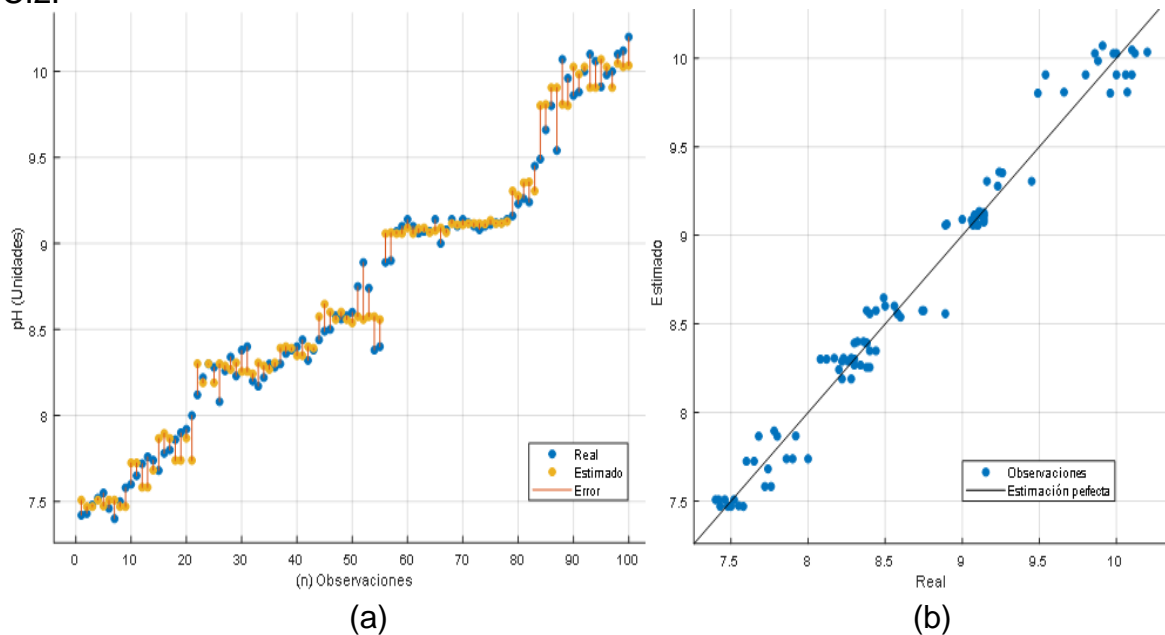


Figura C.1. Estimación pH modelo AR muestra AL-2: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

En la Figura C.1 (a), se representan los valores estimados de color amarillo, de color azul los valores reales y los errores por líneas rojas. La Figura C.1 (b) presenta la distribución de valores estimados cuya línea diagonal equivale a una estimación perfecta. Los puntos se encuentran muy cercanos sobre la línea, lo que evidencia la capacidad del modelo para explicar el 98% de varianza de los datos.

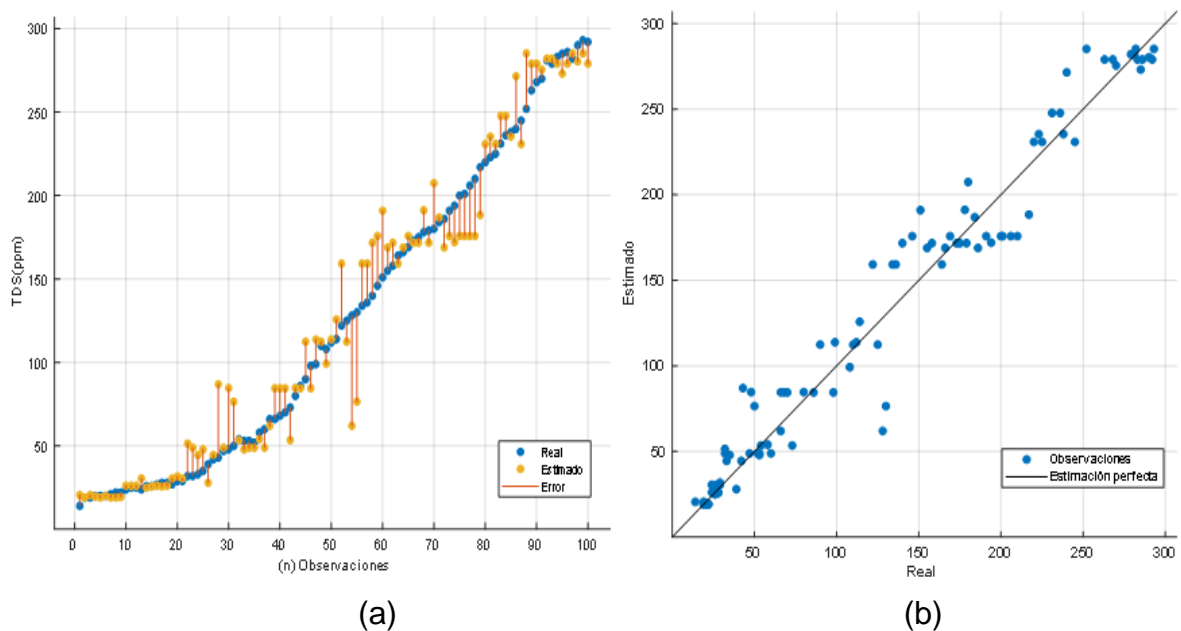


Figura C.2. Estimación TDS modelo AR muestra AL-2: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

En la Figura C.2(a) se muestran los resultados de la estimación de TDS y el diagrama de dispersión de los datos estimados (b), se aprecia también una buena aproximación de los datos predichos con un ajuste del 96% a los datos reales.

## C.2. Resultados Estimación Muestra AL-4

Para la muestra AL-4 se construyeron los modelos para la estimación del pH y TDS, los resultados comparativos se muestran en la Tabla C.2.

Tabla C.2. Resultados estimación pH y TDS - muestra AL-4.

Parámetro	pH			TDS		
	RL	AR	MVS	RL	AR	MVS
<b>MAE</b>	0.2220	0.1323	0.2214	19.594	380.85	558.84
<b>RMSE</b>	0.2784	0.1681	0.2882	22.023	19.515	23.64
<b>R<sup>2</sup></b>	0.50	<b>0.82</b>	0.46	0.49	<b>0.60</b>	0.41

Al comparar los valores de la Tabla C.2, el modelo de AR presenta resultados superiores frente al desempeño de los modelos RL y MVS en las dos estimaciones. Las Figuras C.3 y C.4 contienen los desempeños de los modelos AR.

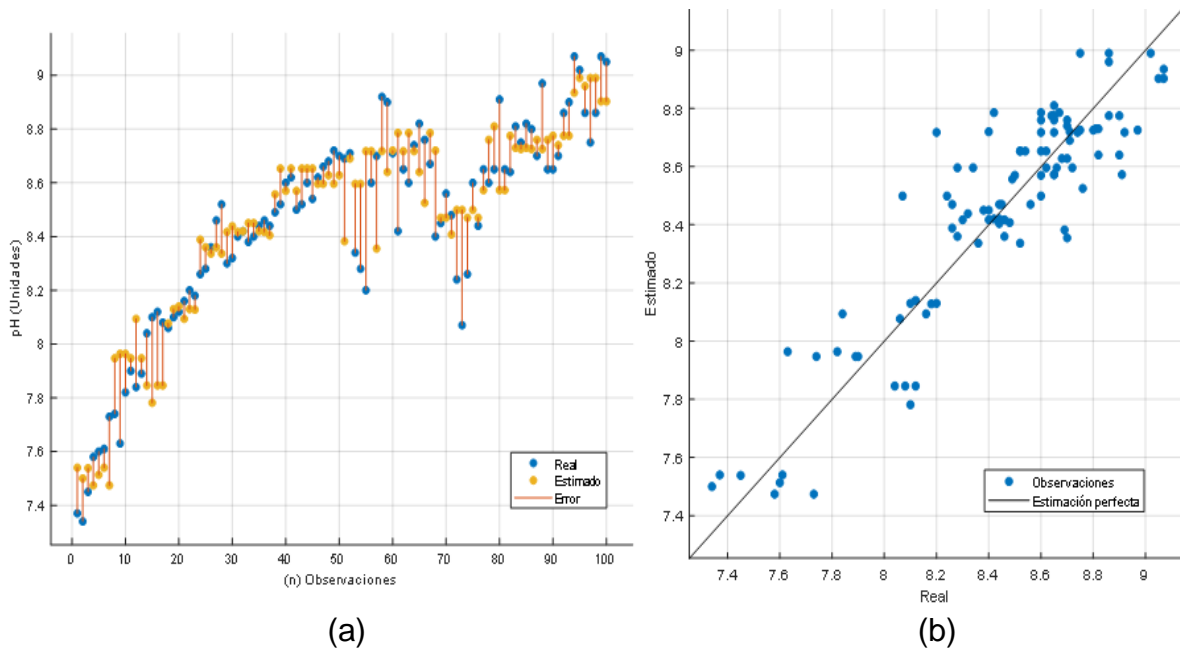


Figura C.3. Estimación pH modelo AR muestra AL-4: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

En la Figura C.3 se aprecia un buen ajuste de los datos estimados y reales con un valor de 82% de ajuste del modelo y los valores menores a uno para las métricas MAE y RMSE



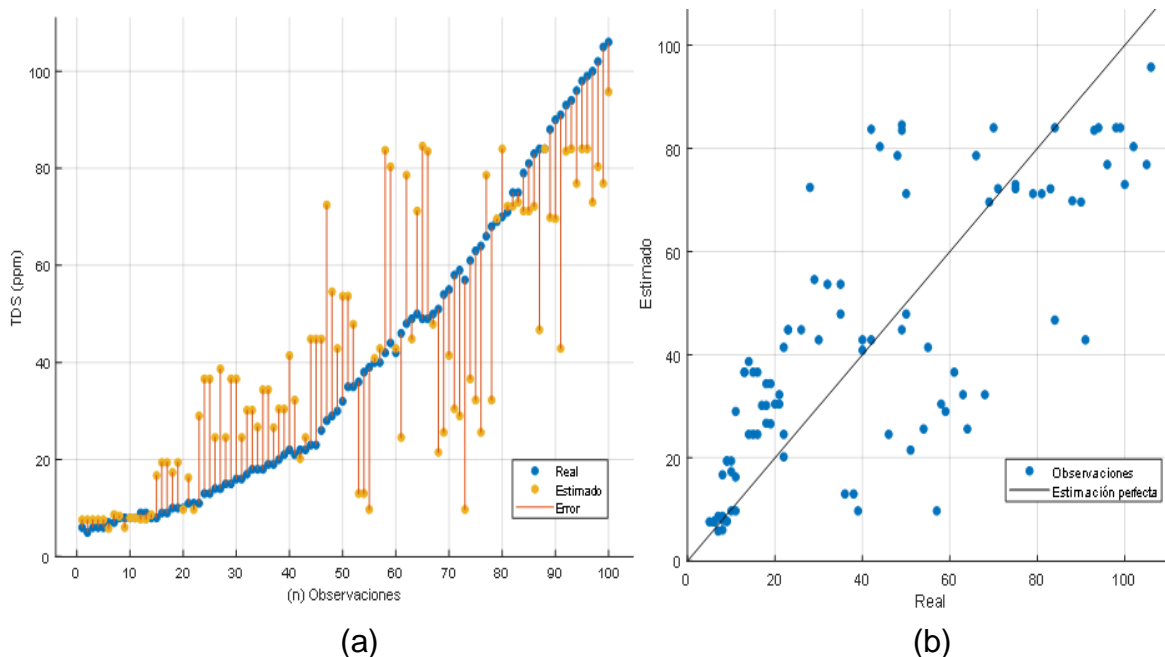


Figura C.4. Estimación TDS modelo AR muestra AL-4: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

En la estimación del TDS para esta muestra puede observarse que, si bien el modelo AR es superior a la RL y MVS, su desempeño es de apenas el 60%, como puede apreciarse en la Figura C.4 (b) cuyos puntos se encuentran muy dispersos y alejados de la línea diagonal de estimación perfecta. Los valores altos de MAE y RMSE también se tribuyen a la penalización que se da por la cantidad y tamaño de los errores.

### C.3. Resultados Estimación Muestra AA-2

Se estimaron los parámetros de pH y TDS para la muestra AA-2 de agua lluvia, los resultados comparativos del desempeño de los modelos aparece en la Tabla C.3.

Tabla C.3. Resultados estimación pH y TDS - muestra AA-2.

Parámetro	pH			TDS		
	RL	AR	MVS	RL	AR	MVS
<b>MAE</b>	0.1301	0.0854	0.1279	8.0819	6.3868	7.8519
<b>RMSE</b>	0.1634	0.1139	0.1635	10.04	8.5318	9.8513
<b>R<sup>2</sup></b>	0.93	<b>0.97</b>	0.93	0.92	<b>0.95</b>	0.93

Los modelos de árboles de regresión presentaron mejores valores en las métricas MAE, RMSE y R<sup>2</sup>, las Figuras C.5 y C.6 contienen los resultados gráficos de estos modelos.

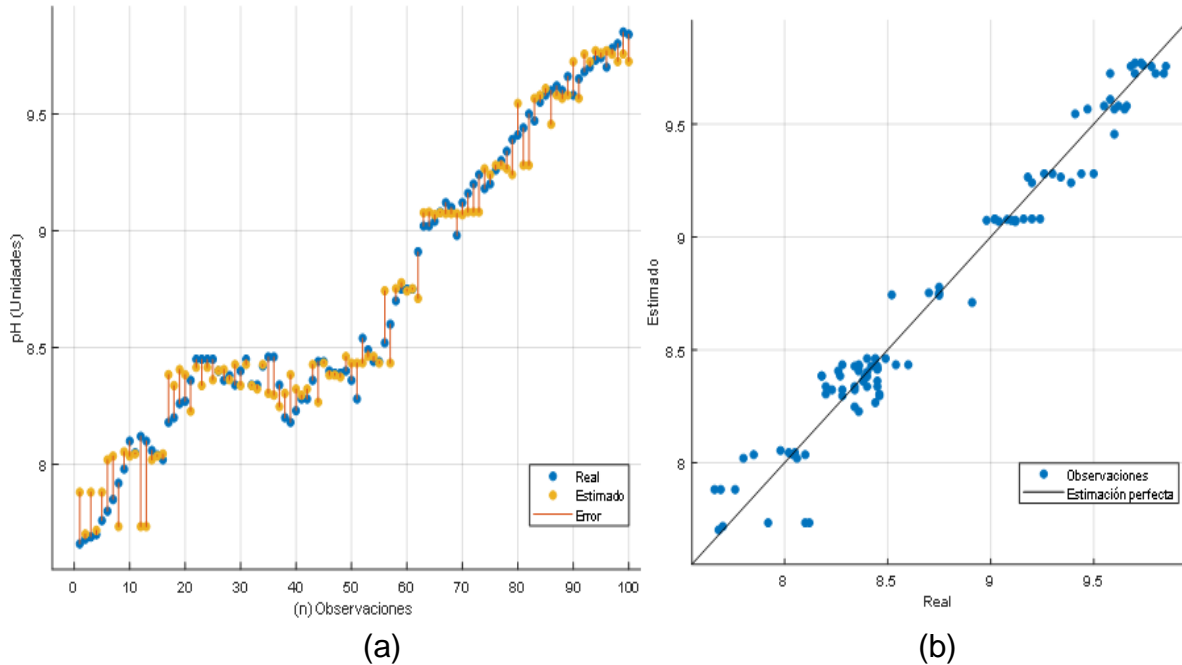


Figura C.5. Estimación pH modelo AR muestra AA-2: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

Para la estimación del pH, el modelo AR presento buenos resultados con un ajuste del 97% de los valores estimados a los valores reales, como se observa en la Figura C.5 (a) y (b).

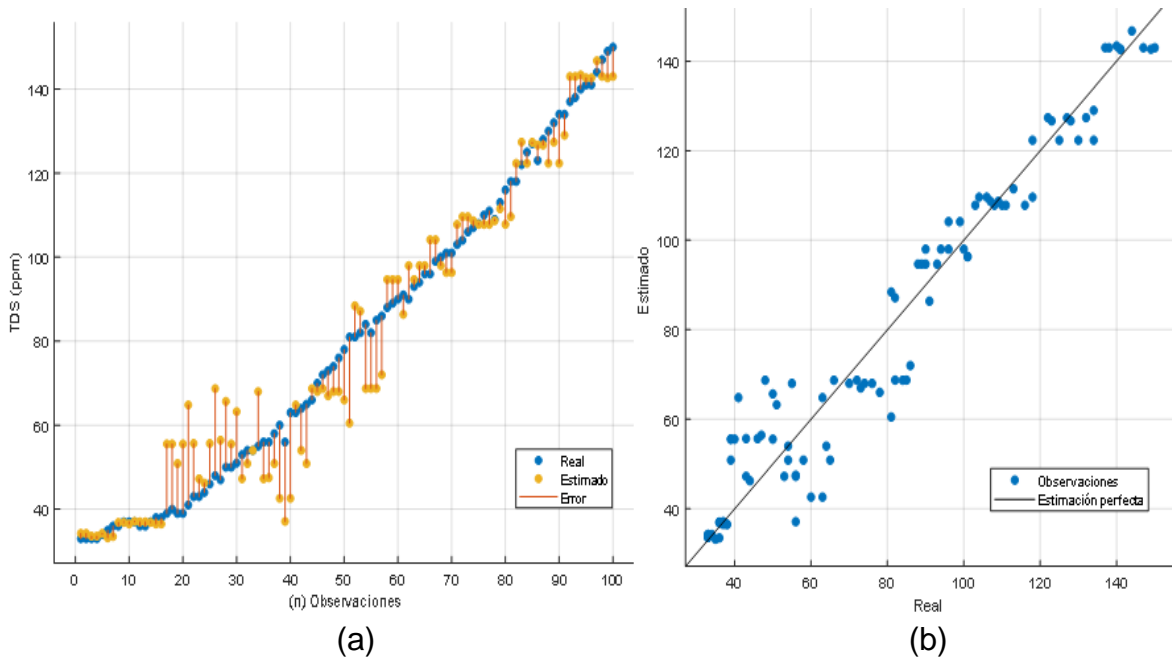


Figura C.6. Estimación TDS modelo AR muestra AA-2: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

En la estimación del TDS el modelo AR también presentó un buen desempeño con valores pequeños de MAE y RMSE lo que indica una buena cercanía de los valores estimados a los reales y que se aprecia en la Figura C.6 (b).

## C.4. Resultados Estimación Muestra AA-4

De manera similar se aplicó el procedimiento para la estimación del pH y TDS sobre la muestra AA-4, los resultados de los modelos se muestran en la Tabla C.4.

Tabla C.4. Resultados estimación pH y TDS - muestra AA-4.

Parámetro	pH			TDS		
	RL	AR	MVS	RL	AR	MVS
MAE	0.1589	0.1061	0.1614	13.209	8.4228	11.884
RMSE	0.2047	0.1578	0.2106	16.361	12.609	14.902
R <sup>2</sup>	0.88	<b>0.93</b>	0.87	0.88	<b>0.93</b>	0.90

De acuerdo a los resultados de la Tabla C.4 se observa que el mejor desempeño fue para el modelo AR en la estimación del pH y el TDS. Las figuras C.7 y C.8 muestran el desempeño de los modelos AR.

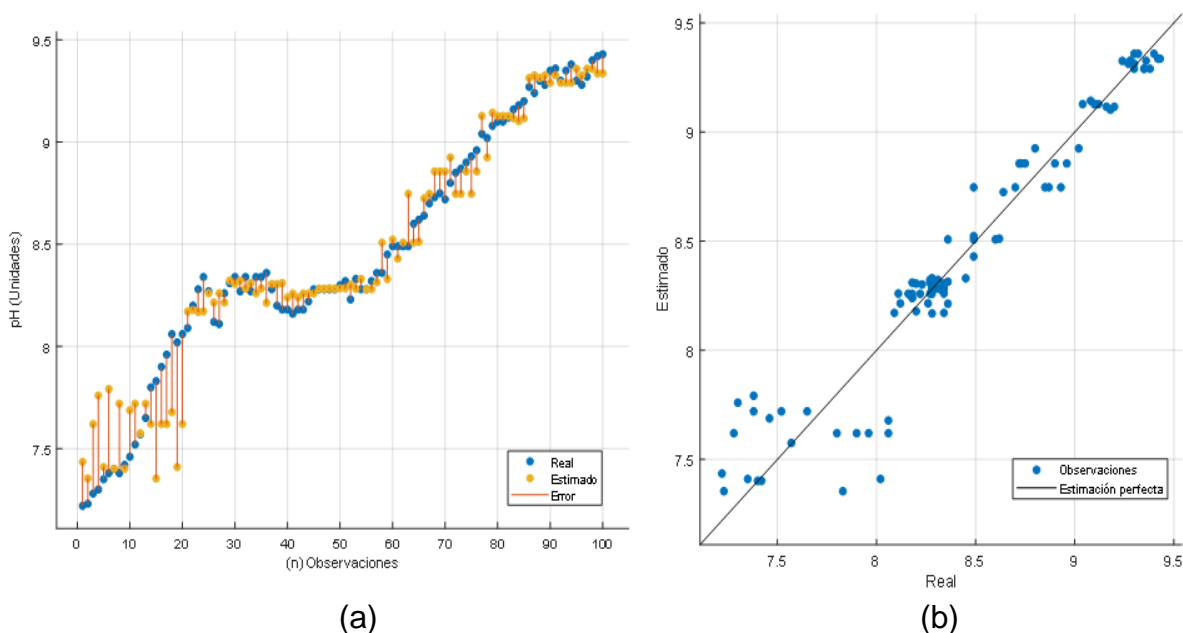


Figura C.7. Estimación pH modelo AR muestra AA-4: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

En la Figura C.7 (a) y (b) se observa que el modelo AR presenta un buen desempeño con estimaciones muy cercanas a los valores reales y a la línea diagonal de estimación perfecta, justificando los valores de MAE, RMSE y R<sup>2</sup>.

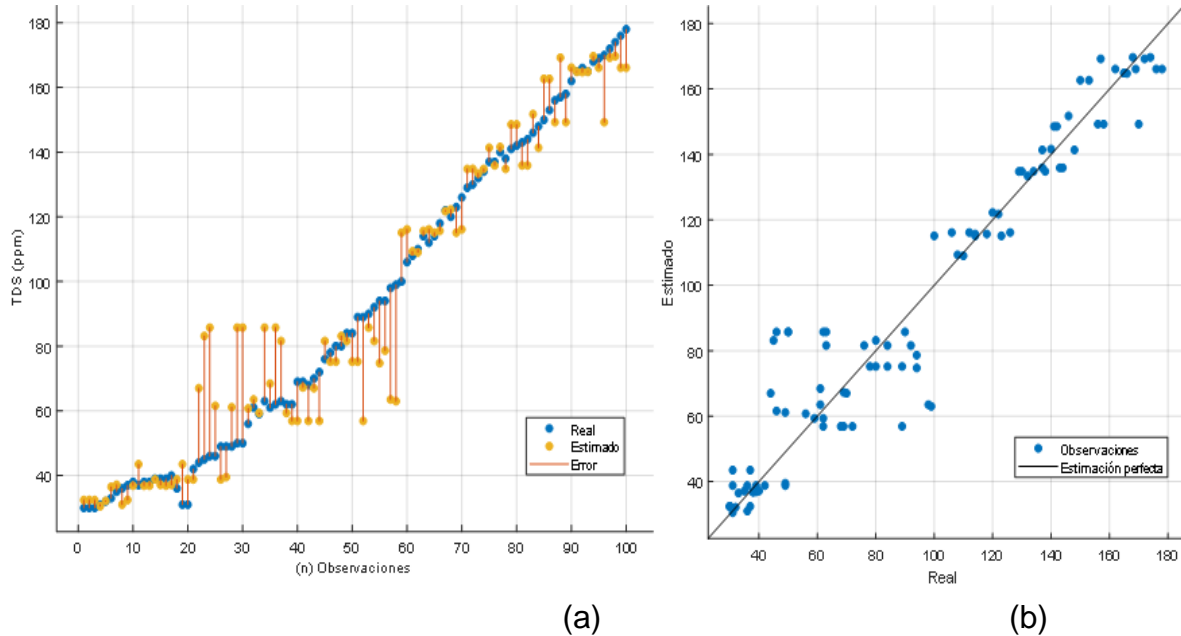


Figura C.8. Estimación TDS modelo AR muestra AA-4: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

En la estimación del TDS también se observa un buen ajuste del modelo a los datos reales con un porcentaje del 93%. Las unidades de MAE y RMSE se da en las mismas unidades de la variable de salida, por lo que los valores obtenidos de 8.4 y 12.6 representan una cercanía considerable de las estimaciones y los datos medidos.

## C.5. Resultados Estimación Muestra AE-1

Finalmente se estimaron los parámetros de pH y TDS sobre las muestras restantes de agua envasada, los resultados de las estimaciones sobre la muestra AE-1 aparecen en la Tabla C.5.

Tabla C.5. Resultados estimación pH y TDS - muestra AE-1.

Parámetro	pH			TDS		
	RL	AR	MVS	RL	AR	MVS
<b>MAE</b>	0.1530	0.0901	0.1465	13.162	13.77	12.965
<b>RMSE</b>	0.1858	0.1148	0.2611	16.125	16.125	15.937
<b>R<sup>2</sup></b>	0.92	<b>0.97</b>	0.84	0.92	0.92	<b>0.93</b>

En la predicción de los valores de pH, el modelo de arboles de regresión tuvo mejores resultados. Para la estimación del TDS el modelo de máquinas de vectores de soporte fue superior seguido por valores muy cercanos de desempeño de los modelos de RL y AR. Los resultados de los modelos AR y MVS se muestran en las Figuras C.9 y C.10.

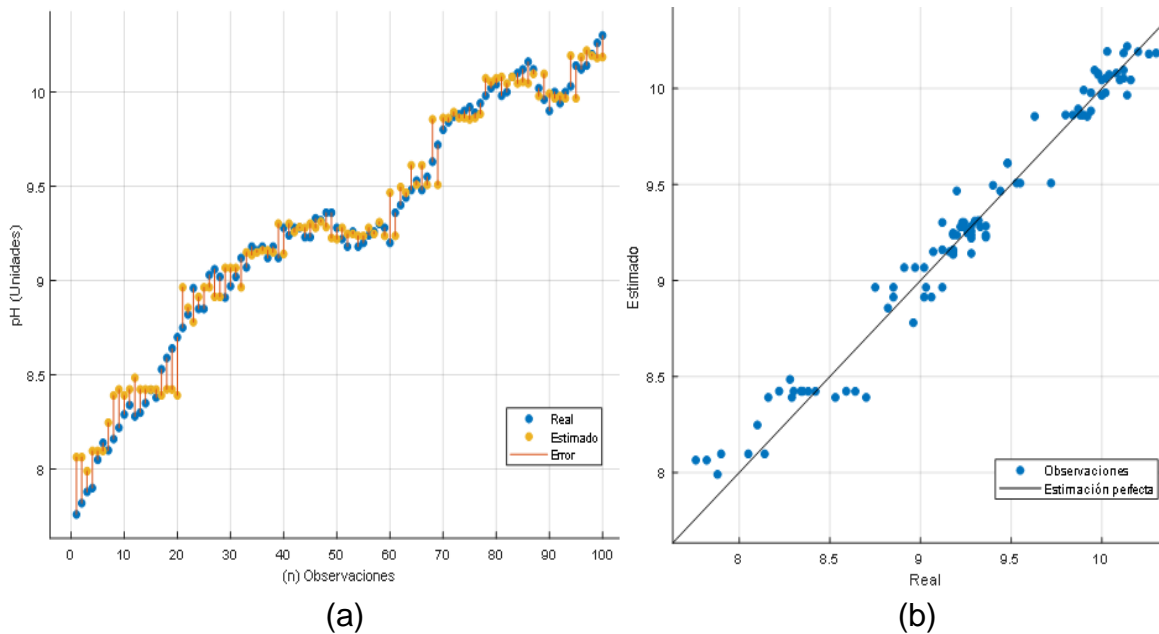


Figura C.9. Estimación pH modelo AR muestra AE-1: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

En la Figura C.9 (a) y (b) se observa un buen desempeño del modelo AR en la estimación del pH con un ajuste del 97% lo que representa un poco dispersión y cercanía de los valores estimados a la línea diagonal de estimación perfecta.

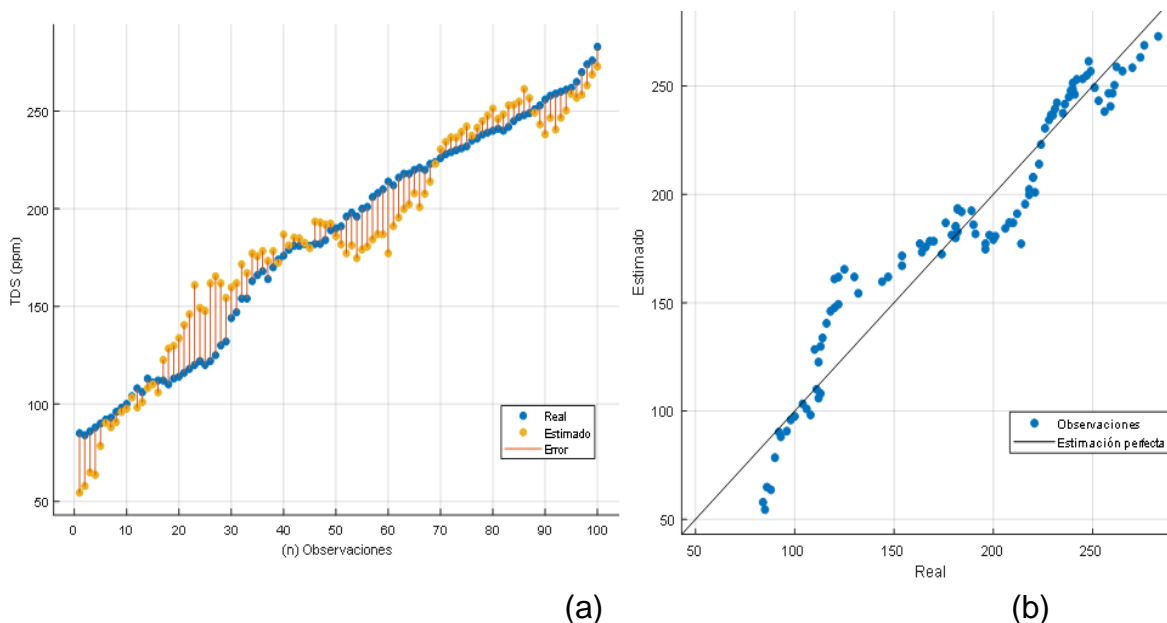


Figura C.10. Estimación TDS modelo MVS muestra AE-1: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

Para la estimación del TDS el modelo MVS presenta un ajuste del 93% de los valores estimados a los reales como se aprecia en la Figura C.10, con alguna leve dispersión de los puntos pero sobre la línea diagonal

## C.6. Resultados Estimación Muestra AE-3

Para la muestra AE-3, estimaron los parámetros de pH y TDS, los resultados comparativos de los modelos se muestran en la Tabla C.6.

Tabla C.6. Resultados estimación pH y TDS – muestra AE-3.

Parámetro	pH			TDS		
	RL	AR	MVS	RL	AR	MVS
MAE	0.1237	0.0919	0.1444	11.597	9.7613	9.6674
RMSE	0.1520	0.1197	0.2248	14.345	12.292	11.545
R <sup>2</sup>	0.92	<b>0.95</b>	0.83	0.92	0.94	<b>0.95</b>

De acuerdo a los valores de la Tabla C.6, los modelos de mejor desempeño fueron el AR para la estimación del pH y MVS para la estimación del TDS. Las Figuras C.11 y C.12 presentan los resultados de los modelos respectivamente.

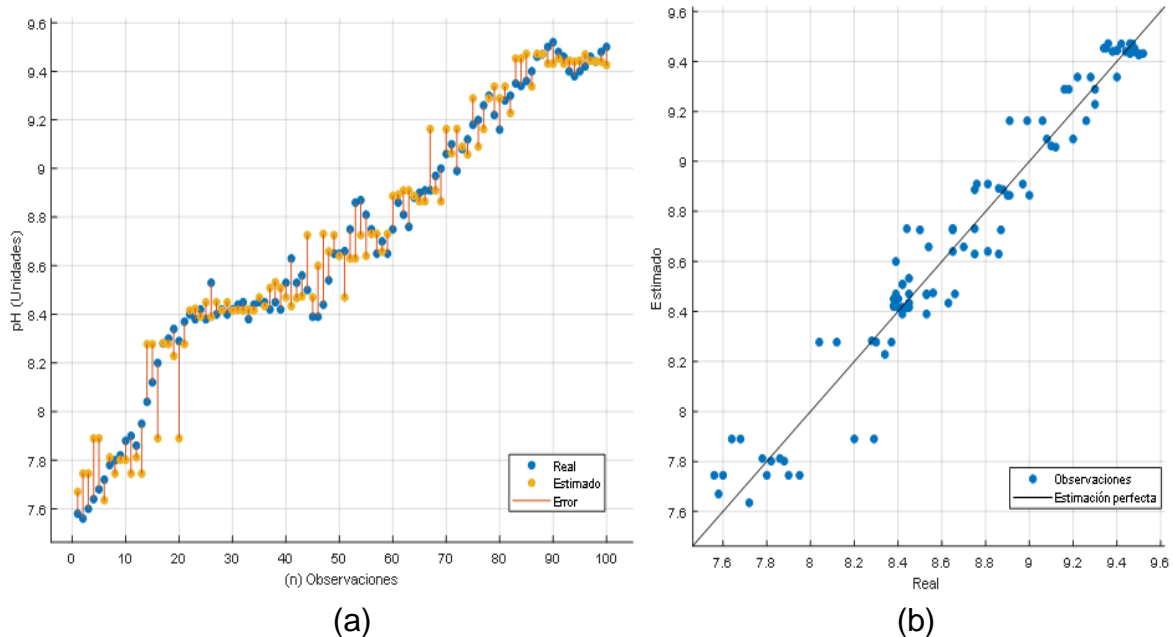


Figura C.11. Estimación pH modelo AR muestra AE-3: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

En la Figura C.11 se observa que los valores estimados se encuentran cercanos a los valores reales y a la línea diagonal de estimación perfecta, esto concuerda con los valores de desempeño obtenidos por el modelo AR con un ajuste del 95% y valores pequeños menores a 1 en las métricas de MAE y RMSE.

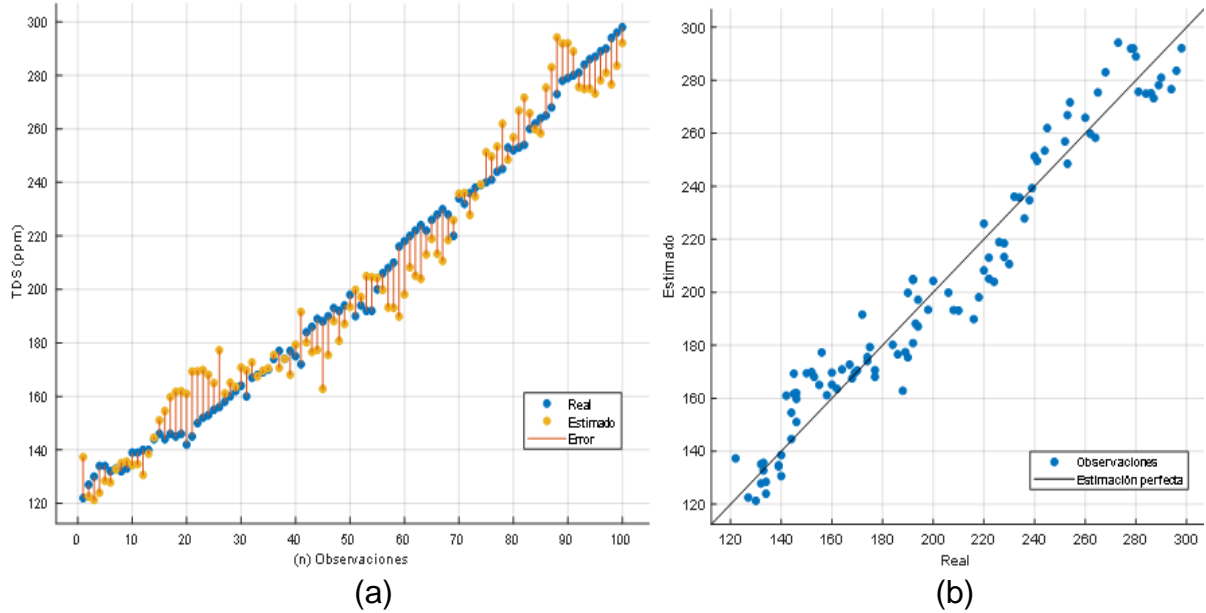


Figura C.12. Estimación TDS modelo MVS muestra AE-3: (a) Resultados modelo, (b) Diagrama de dispersión valores reales vs estimados.

En la Figura C.12 se observan los valores estimados cercanos sobre la línea diagonal que corresponde al ajuste del modelo de un 95% de los valores estimados a los datos reales.

## Anexo D

# Modelo Estimación del ICA y Procedimiento de Validación

## D.1. Características del Modelo para la Estimación del ICA

Los modelos para la estimación del ICA se construyeron a partir de las muestras AL-3, AA-1 y AE-2. De los resultados obtenidos, el modelo que presentó un mejor desempeño corresponde al modelo de árboles de regresión de la muestra AA-1 (**AR-AA1**) con valores de 0.98 para  $R^2$ , 0.9257 para RMSE y 0.6627 para MAE, lo que representa un ajuste del 98% de los valores estimados a los valores del ICA reales. La arquitectura del modelo (**AR-AA1**) se muestra en la Figura D.1.

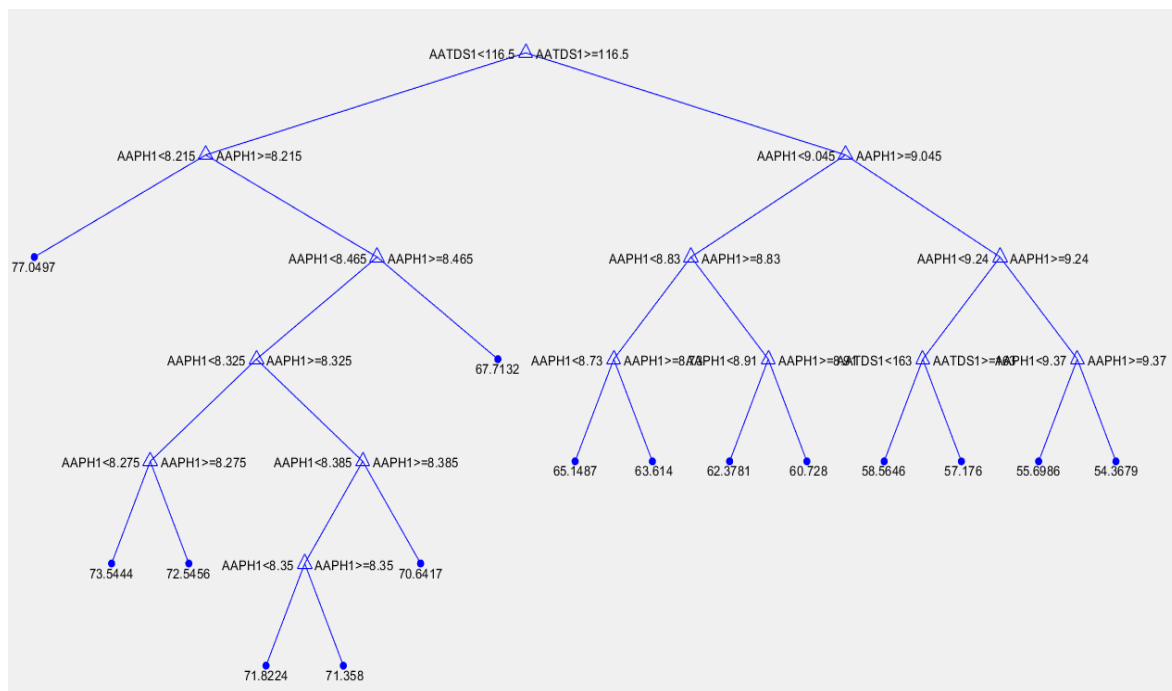


Figura D.1. Arquitectura del modelo (AR-AA1) Estimación ICA.



La estructura del modelo de árboles de regresión cuenta con 29 nodos totales, de los cuales 15 corresponden a nodos terminales, en la figura se puede apreciar cada nodo con su respectiva regla de decisión.

## D.2. Procedimiento para la Estimación de Nuevos Datos

Una vez realizada la estimación del ICA con las respectivas muestras en la herramienta *Regression Learner* de Matlab, se almacenaron los modelos de mejores resultados incluido el modelo seleccionado para la validación con nuevas muestras (**AR-AA1**) Figura D.2.

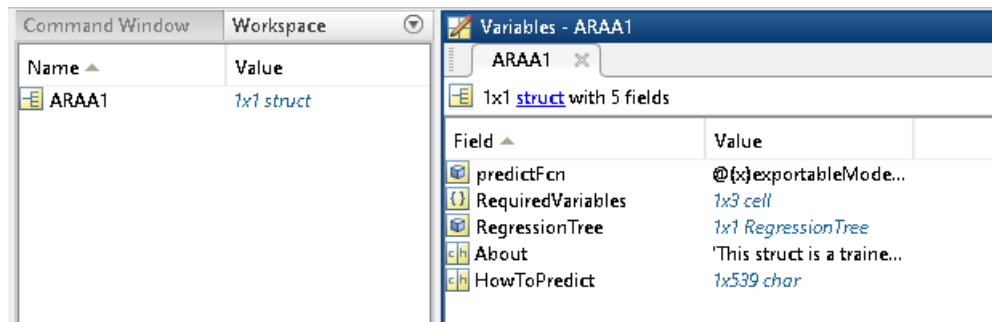


Figura D.2. Modelo AR-AA1 para validación.

Para estimar nuevos valores, el modelo requiere como entradas tres parámetros que son el pH, TDS y temperatura, si bien los modelos se construyeron con 100 observaciones, el número de datos por parámetro puede variar, Figura D.3.

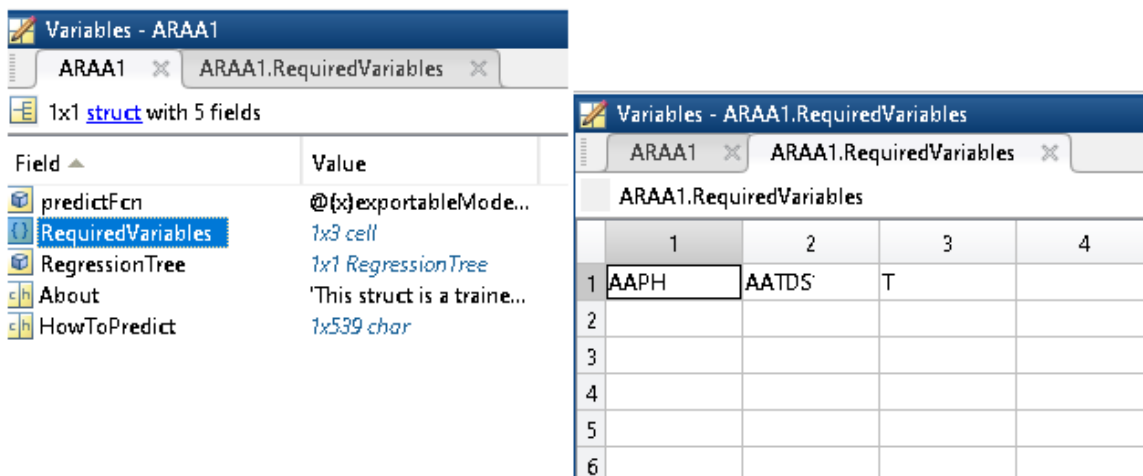


Figura D.3. Parámetros de entrada modelo AR-AA1.

Una vez se tengan los nuevos datos de entrada en el espacio de trabajo con el comando:  $yfit = ARAA1.predictFcn(T)$  donde T corresponde a la variable que

contiene los nuevos valores, se obtienen los valores estimados del índice ICA Figura D.4.

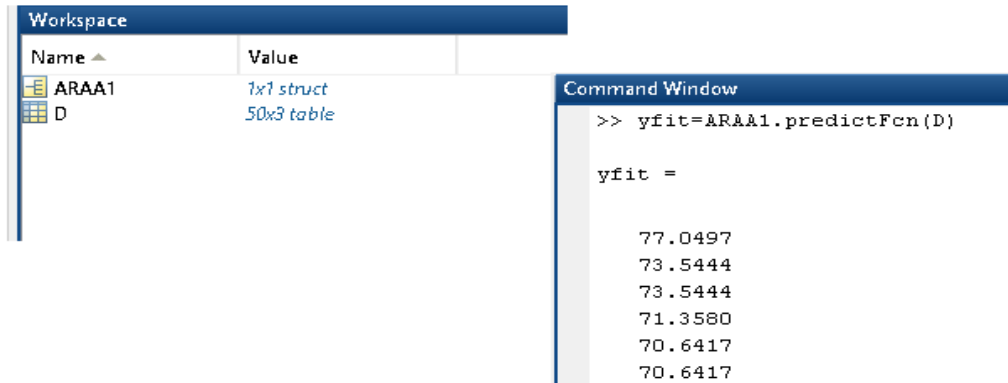


Figura D.4. Parámetros de entrada modelo AR-AA1.

Una vez obtenidos los valores estimados estos se grafican en Matlab junto a los valores calculados del ICA. El código utilizado para interpolar los valores del pH, TDS y temperatura necesarios para el cálculo del ICA se muestran en la Figura D.5.

```

INDICE DE CALIDAD

%INDICE DE PH
x=ALPH3;      %ingresar pH
y=0.10.*(x<4)+((4<=x)&(x<7)).*(0.026284149*exp(x*0.520025))+1.*((7<=x)&(x<=8))+((8<x)&(x<=11))
IpH=y*0.4*100;
WIpH=IpH(:,1) %salida

%INDICE TDS
x3=[0 50 100 150 200 250 300 350 400 450 500];
y3=[80 88 85 79 71 67 60 52 48 39 30];

p1=polyfit(x3,y3,3);

xp1=ALTDS3; %ingresar tds
yp1=polyval(p1,xp1);

ITDS=yp1*0.4;
WTDS=ITDS(:,1) %salida

%INDICE DE TEMPERATURA

x4=[-10 -5 0 5 10 15 20 25 30];
y4=[55 74 93 72 44 29 20 18 10];

p2=polyfit(x4,y4,6);
xp2=T;%ingresar T
yp2=polyval(p2,xp2);
IT=yp2*0.2;
WT=IT(:,1) % salida
  
```

Figura D.5. Código interpolación de parámetros.

### D.3. Estimación del ICA sin el parámetro temperatura

Se realizaron pruebas para la estimación del índice de calidad, teniendo como entradas los parámetros de pH y TDS, a fin de evaluar el desempeño de los modelos sin incluir el parámetro de temperatura, debido a que este presenta un bajo coeficiente de correlación de la temperatura con estos parámetros. Los resultados de estimación ICA para la muestra AL-3 son:

Tabla D.1. Resultados estimación ICA – muestra AL-3.

Parámetro	ICA		
	RL	AR	MVS
MAE	0.8163	0.8819	1.1597
RMSE	1.4134	1.2529	1.8205
R <sup>2</sup>	0.97	0.97	0.95

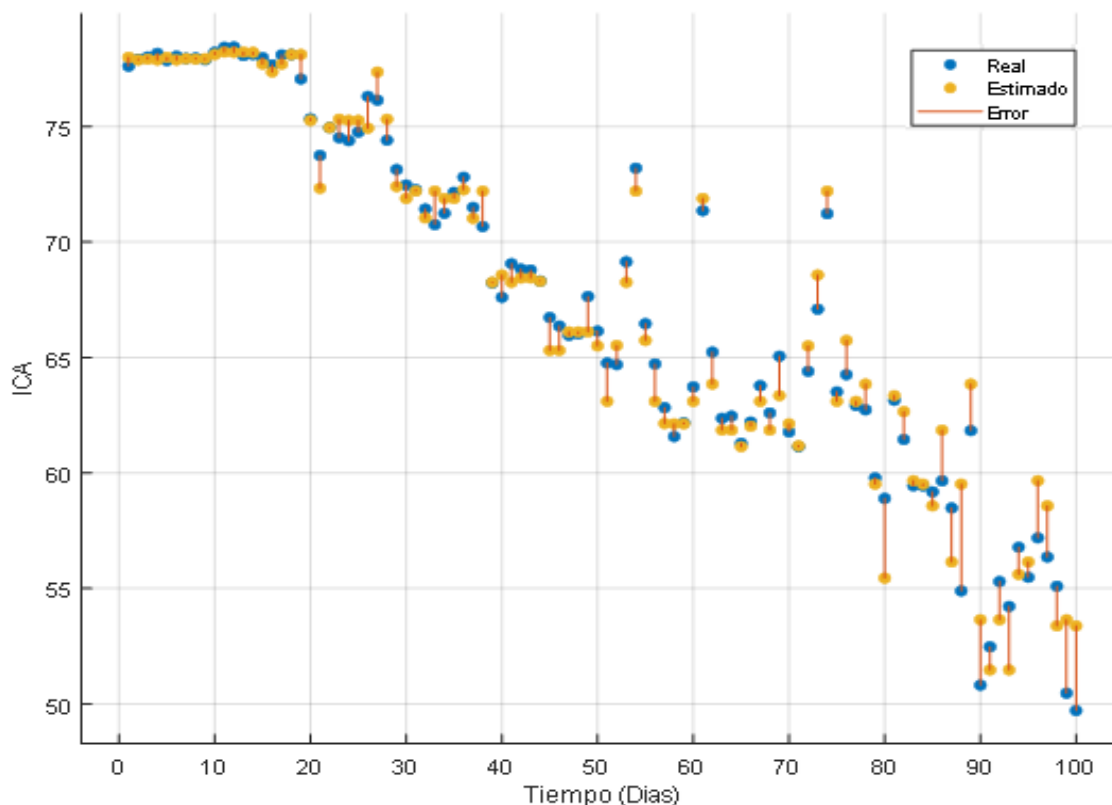


Figura D.6. Estimación ICA modelo AR- muestra AL-3

Los resultados de estimación ICA para la muestra AA-1 son:

Tabla D.2. Resultados estimación ICA – muestra AA-1.

Parámetro	ICA		
	RL	AR	MVS
<b>MAE</b>	0.6724	0.6455	0.9500
<b>RMSE</b>	1.15	0.8611	1.6304
<b>R<sup>2</sup></b>	0.97	0.98	0.95

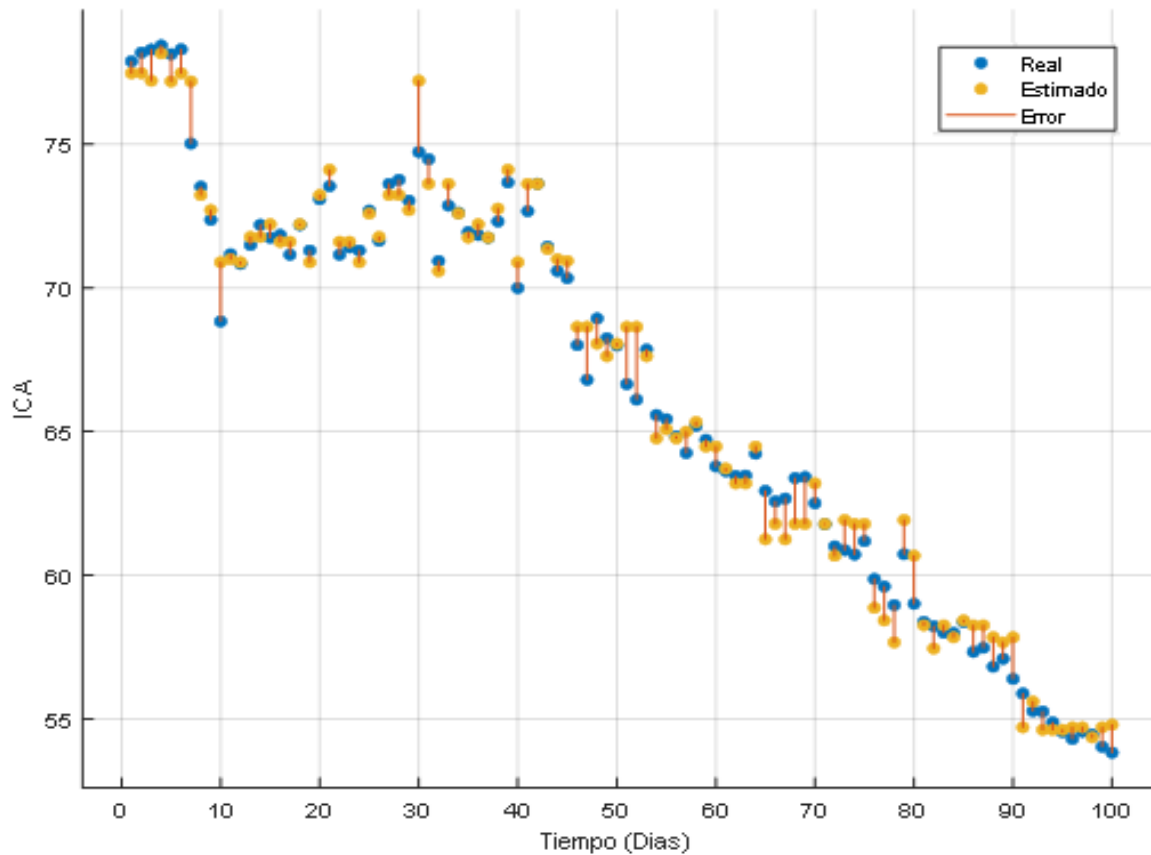


Figura D.7. Estimación ICA modelo AR- muestra AA-1

Los resultados de estimación ICA para la muestra AE-2 son:

Tabla D.3. Resultados estimación ICA – muestra AE-2.

Parámetro	ICA		
	RL	AR	MVS
<b>MAE</b>	0.7884	0.8513	0.9087
<b>RMSE</b>	1.1924	1.0308	1.1295
<b>R<sup>2</sup></b>	0.98	0.98	0.96

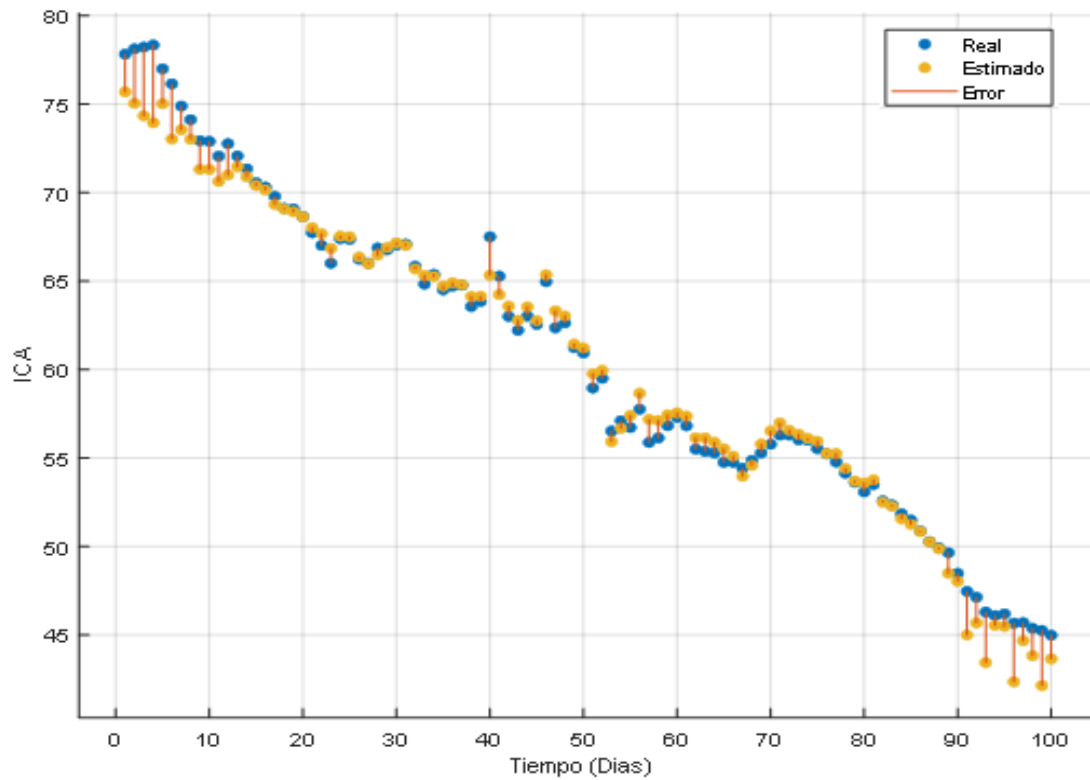


Figura D.0.8. Estimación ICA modelo RL- muestra AE-2

Al comparar los resultados anteriores con los obtenidos en el documento en la estimación del ICA, teniendo como variable de entrada la temperatura en el modelo, se observa que los resultados de desempeño en los dos casos son similares para las métricas de MAE, RMSE y  $R^2$ .

## Datos del Experimento

Ob s.	Muestras agua de acueducto								Muestras agua envasada								Muestras agua lluvia								T
	AAP H1	AATD S1	AAP H2	AATD S2	AAP H3	AATD S3	AAP H4	AATD S4	AEP H1	AETD S1	AEP H2	AETD S2	AEP H3	AETD S3	AEP H4	AETD S4	ALP H1	ALTD S1	ALP H2	ALTD S2	ALP H3	ALTD S3	ALP H4	ALTD S4	
1	7.29	36	7.66	33	7.61	33	7.22	30	7.76	85	7.85	68	7.58	122	7.48	75	7.02	12	7.42	14	7.64	11	7.37	6	22
2	7.45	36	7.68	33	7.68	33	7.23	30	7.82	84	7.9	68	7.56	127	7.46	76	7.06	12	7.43	19	7.66	11	7.34	5	19
3	7.7	38	7.69	33	7.7	33	7.28	30	7.88	86	7.94	68	7.6	130	7.5	79	7.1	14	7.48	19	7.5	11	7.45	6	18
4	7.63	39	7.7	33	7.85	34	7.3	31	7.9	88	8	71	7.64	134	7.52	80	7.16	15	7.52	20	7.55	12	7.58	6	17
5	7.88	41	7.76	34	7.89	35	7.35	32	8.05	90	8.05	72	7.68	134	7.48	82	7.18	16	7.55	20	7.7	12	7.6	6	20
6	7.98	40	7.8	35	7.9	35	7.38	33	8.14	92	8.1	74	7.72	132	7.43	82	7.1	18	7.46	20	7.8	13	7.61	7	18
7	8.16	41	7.85	36	7.92	35	7.4	35	8.1	93	8.16	74	7.78	133	7.59	83	7.04	20	7.4	21	7.82	13	7.73	7	19
8	8.24	40	7.92	36	7.99	36	7.38	36	8.16	96	8.2	75	7.8	132	7.6	84	7.16	20	7.5	22	7.84	14	7.74	8	19
9	8.3	41	7.98	37	8.02	37	7.42	37	8.22	98	8.26	75	7.82	133	7.64	86	7.22	22	7.58	22	7.7	14	7.63	8	20
10	8.53	42	8.1	37	8.1	36	7.46	38	8.29	100	8.28	75	7.88	139	7.55	87	7.2	24	7.6	24	7.68	15	7.82	8	17
11	8.4	43	8.05	37	8.16	37	7.52	37	8.34	104	8.34	74	7.9	139	7.68	86	7.18	26	7.65	25	7.72	15	7.9	8	16
12	8.42	42	8.12	36	8.12	37	7.57	38	8.28	108	8.3	74	7.86	140	7.7	82	7.24	28	7.72	25	7.7	16	7.84	9	16
13	8.36	43	8.1	36	8.18	38	7.65	38	8.3	106	8.32	72	7.95	140	7.73	86	7.26	30	7.76	24	7.56	15	7.89	9	18
14	8.32	44	8.06	37	8.2	38	7.8	39	8.35	113	8.36	75	8.04	144	7.78	87	7.16	32	7.74	26	7.8	18	8.04	8	18
15	8.34	44	8.04	38	8.28	39	7.83	39	8.42	111	8.4	76	8.12	146	7.8	88	7.2	34	7.68	26	7.98	15	8.1	8	19
16	8.34	45	8.02	38	8.34	39	7.9	39	8.38	112	8.42	77	8.2	144	7.87	88	7.28	34	7.78	27	8.02	16	8.12	9	18
17	8.38	45	8.18	39	8.36	40	7.96	40	8.53	112	8.45	79	8.28	146	7.9	88	7.32	36	7.8	28	7.92	17	8.08	9	18
18	8.32	42	8.2	40	8.38	40	8.06	36	8.59	110	8.49	80	8.3	145	7.94	90	7.36	37	7.86	28	7.89	18	8.06	10	18
19	8.38	46	8.26	39	8.28	38	8.02	31	8.64	113	8.5	82	8.34	146	8.02	92	7.4	38	7.9	27	8.06	20	8.1	10	17
20	8.27	46	8.27	39	8.2	39	8.06	31	8.7	114	8.52	80	8.29	142	8.05	94	7.43	38	7.92	29	8.14	20	8.12	10	18
21	8.24	48	8.36	41	8.44	41	8.09	42	8.75	116	8.57	83	8.37	145	8.16	96	7.48	40	8	29	8.22	21	8.16	11	19
22	8.38	52	8.45	43	8.42	42	8.2	44	8.82	118	8.62	86	8.4	150	8.09	98	7.5	41	8.12	32	8.16	20	8.2	11	18
23	8.36	48	8.45	43	8.34	44	8.28	45	8.96	120	8.68	88	8.38	152	8.24	99	7.68	40	8.22	32	8.18	23	8.18	11	19
24	8.38	50	8.45	44	8.26	44	8.34	46	8.85	122	8.6	90	8.42	153	8.2	104	7.7	42	8.3	33	8.2	24	8.26	13	17
25	8.3	50	8.45	46	8.34	45	8.27	46	8.85	120	8.6	93	8.38	155	8.22	107	7.68	44	8.28	35	8.18	24	8.28	13	17

26	8.36	52	8.4	48	8.34	47	8.12	49	9.03	122	8.67	97	8.53	156	8.24	110	7.7	46	8.08	39	8.1	24	8.36	14	17
27	8.26	53	8.36	47	8.34	48	8.11	49	9.06	125	8.7	97	8.4	158	8.24	112	7.82	48	8.26	42	8.12	25	8.46	14	16
28	8.24	55	8.38	50	8.42	49	8.26	49	9.02	130	8.62	102	8.42	160	8.3	114	7.62	50	8.34	43	8.2	26	8.52	15	17
29	8.28	56	8.34	50	8.38	49	8.31	50	8.91	132	8.63	100	8.4	162	8.31	114	7.74	50	8.23	47	8.27	28	8.3	15	17
30	8.18	58	8.4	51	8.3	49	8.34	50	8.97	144	8.6	102	8.42	164	8.32	113	7.66	52	8.38	48	8.3	29	8.32	16	18
31	8.2	61	8.45	53	8.2	52	8.27	56	9.02	147	8.6	106	8.44	160	8.26	120	7.7	54	8.4	50	8.32	30	8.4	16	17
32	8.39	64	8.34	54	8.12	55	8.34	61	9.12	154	8.67	109	8.45	167	8.27	125	7.9	56	8.2	54	8.36	30	8.42	17	18
33	8.28	63	8.34	54	8.18	53	8.27	59	9.07	154	8.74	109	8.38	168	8.27	125	7.8	56	8.17	53	8.4	31	8.38	18	18
34	8.3	67	8.42	55	8.23	55	8.34	63	9.18	163	8.71	110	8.44	169	8.38	126	7.77	52	8.22	53	8.38	32	8.4	18	17
35	8.33	68	8.46	56	8.39	56	8.34	61	9.16	166	8.76	111	8.45	170	8.38	129	7.95	58	8.3	52	8.32	34	8.44	18	18
36	8.33	69	8.46	56	8.28	56	8.36	62	9.18	168	8.74	110	8.45	174	8.31	126	8	60	8.28	58	8.28	43	8.46	19	19
37	8.34	70	8.34	58	8.32	56	8.28	63	9.12	164	8.74	112	8.42	177	8.42	130	8.4	59	8.3	60	8.36	46	8.44	19	18
38	8.3	75	8.2	60	8.39	57	8.2	62	9.18	170	8.82	112	8.45	174	8.49	132	8.16	62	8.36	66	8.4	32	8.49	20	19
39	8.23	75	8.18	56	8.44	57	8.18	62	9.12	174	8.8	115	8.42	177	8.34	134	8.26	64	8.38	66	8.56	50	8.52	21	18
40	8.44	76	8.23	63	8.46	62	8.18	69	9.28	176	8.53	127	8.53	175	8.56	144	8.2	66	8.4	68	8.6	55	8.6	22	18
41	8.28	75	8.28	63	8.44	63	8.16	69	9.24	179	8.67	127	8.63	172	8.6	146	8.36	66	8.44	70	8.5	58	8.62	21	19
42	8.23	79	8.28	64	8.49	64	8.18	68	9.28	181	8.85	122	8.53	184	8.62	147	8.26	67	8.32	73	8.52	60	8.5	22	18
43	8.36	80	8.36	65	8.58	64	8.18	70	9.28	181	8.92	122	8.56	186	8.22	147	8.2	68	8.38	80	8.53	64	8.52	22	17
44	8.4	82	8.44	66	8.6	66	8.22	72	9.23	182	8.84	126	8.5	189	8.18	150	8.18	72	8.44	86	8.55	67	8.6	23	18
45	8.42	86	8.44	70	8.44	68	8.28	76	9.23	181	8.89	124	8.39	188	8.23	156	8.2	74	8.49	90	8.66	73	8.54	23	17
46	8.54	91	8.4	72	8.6	70	8.28	78	9.33	182	8.7	120	8.39	190	8.33	162	8.26	72	8.5	98	8.66	78	8.62	26	20
47	8.6	94	8.39	73	8.49	74	8.28	80	9.32	182	8.86	126	8.44	193	8.33	166	8.3	76	8.58	99	8.67	81	8.66	28	22
48	8.49	93	8.39	74	8.49	75	8.28	80	9.36	184	8.86	130	8.54	192	8.45	173	8.28	79	8.56	110	8.69	84	8.68	29	18
49	8.54	95	8.4	76	8.64	75	8.28	84	9.36	189	8.97	134	8.65	194	8.49	177	8.13	80	8.58	108	8.6	72	8.72	30	17
50	8.54	96	8.36	78	8.72	79	8.3	84	9.28	190	8.97	136	8.65	198	8.49	178	8.16	80	8.6	112	8.67	89	8.7	32	19
51	8.62	98	8.28	81	8.68	83	8.32	89	9.22	191	9.12	139	8.66	190	8.45	180	8.18	82	8.75	114	8.76	90	8.69	35	20
52	8.66	110	8.54	81	8.75	84	8.23	89	9.18	196	9.1	139	8.75	194	8.44	182	8.16	84	8.89	122	8.78	96	8.71	35	17
53	8.55	116	8.49	82	8.7	83	8.33	90	9.26	198	9.39	140	8.86	192	8.44	184	8.18	85	8.74	125	8.49	101	8.34	36	16

54	8.7	117	8.44	84	8.7	85	8.28	92	9.18	196	9.33	141	8.87	192	8.46	186	8.2	89	8.38	128	8.24	112	8.28	38	16
55	8.68	120	8.44	82	8.81	86	8.28	94	9.2	200	9.33	142	8.81	200	8.44	190	8.28	90	8.4	130	8.63	108	8.2	39	18
56	8.72	121	8.52	85	8.81	88	8.32	94	9.24	201	9.23	143	8.75	206	8.39	193	8.32	91	8.89	134	8.74	114	8.6	40	18
57	8.75	123	8.6	86	8.86	90	8.36	98	9.26	206	9.39	146	8.65	208	8.44	194	8.34	89	8.9	136	8.86	120	8.7	40	19
58	8.69	123	8.7	88	8.82	91	8.36	99	9.3	208	9.36	150	8.7	210	8.49	198	8.45	94	9.07	140	8.96	122	8.92	42	18
59	8.72	125	8.75	89	8.89	89	8.45	100	9.28	210	9.28	154	8.65	216	8.44	196	8.56	97	9.1	146	8.9	128	8.9	44	18
60	8.78	128	8.75	90	8.9	90	8.49	106	9.2	214	9.23	156	8.75	218	8.5	202	8.68	99	9.14	151	8.78	130	8.71	42	18
61	8.8	129	8.75	91	8.94	93	8.49	108	9.36	212	9.28	158	8.86	220	8.58	204	8.98	100	9.1	155	8.31	126	8.42	46	17
62	8.8	130	8.91	90	8.88	95	8.49	110	9.4	216	9.39	159	8.81	222	8.62	206	8.9	101	9.06	158	8.67	132	8.65	48	18
63	8.79	131	9.02	93	8.9	95	8.49	114	9.44	218	9.39	160	8.76	224	8.52	202	9.18	106	9.07	164	8.86	136	8.6	49	19
64	8.74	132	9.02	94	8.94	96	8.6	112	9.48	218	9.4	162	8.88	222	8.46	210	9.25	108	9.07	166	8.85	140	8.74	50	18
65	8.82	135	9.04	96	8.96	98	8.62	114	9.53	220	9.44	163	8.9	226	8.59	213	9.12	110	9.14	169	8.92	145	8.82	49	19
66	8.86	137	9.08	96	9.02	100	8.64	118	9.48	221	9.46	164	8.91	228	8.67	218	9	113	9	173	8.86	148	8.76	49	17
67	8.85	138	9.12	99	9	102	8.7	122	9.55	220	9.49	165	8.91	230	8.7	216	9.29	116	9.08	175	8.74	150	8.67	50	17
68	8.8	137	9.1	100	9.06	105	8.73	120	9.63	223	9.44	166	8.97	228	8.75	219	9.39	118	9.14	178	8.82	152	8.4	51	17
69	8.78	140	8.98	101	9.1	106	8.75	123	9.72	224	9.38	167	9	220	8.8	220	9.32	120	9.1	179	8.63	155	8.45	54	18
70	8.84	140	9.12	101	9.1	107	8.72	126	9.8	226	9.32	168	9.06	234	8.9	219	9.46	122	9.14	180	8.85	157	8.56	55	19
71	8.9	141	9.16	103	9.04	109	8.8	129	9.84	228	9.28	168	9.1	232	9	224	9.26	126	9.12	184	8.9	159	8.48	58	18
72	8.95	142	9.2	104	9.13	110	8.85	130	9.87	229	9.26	171	8.99	236	8.92	223	9	130	9.1	186	8.65	162	8.24	59	19
73	8.96	145	9.24	106	9.16	110	8.87	132	9.88	230	9.29	172	9.08	238	9	226	9.07	132	9.08	191	8.48	163	8.07	57	18
74	8.98	142	9.18	107	9.2	112	8.9	134	9.9	231	9.3	170	9.12	239	9.06	230	9.07	131	9.1	194	8.24	163	8.26	61	18
75	8.92	148	9.2	108	9.18	114	8.93	137	9.92	232	9.32	174	9.18	240	9.16	228	9.1	137	9.11	200	8.7	166	8.6	63	19
76	9.03	149	9.26	110	9.23	112	8.96	137	9.89	235	9.35	176	9.2	241	9.2	232	9.28	140	9.12	201	8.65	168	8.44	64	18
77	9.06	150	9.3	111	9.27	115	9.04	140	9.94	236	9.4	178	9.26	244	9.25	234	9.23	144	9.12	206	8.75	169	8.65	66	17
78	9.1	151	9.34	109	9.2	117	9.02	138	9.98	238	9.45	178	9.3	245	9.2	236	9.2	146	9.14	210	8.75	170	8.6	68	18
79	8.96	153	9.39	113	9.25	119	9.08	141	10.0 2	239	9.5	182	9.22	253	9.18	240	9.39	147	9.16	217	8.98	171	8.65	69	17
80	9.06	158	9.41	116	9.29	120	9.1	142	10.0 4	240	9.52	183	9.16	252	9.27	242	9.28	149	9.23	220	9.02	173	8.91	70	20
81	9.12	158	9.44	118	9.31	122	9.1	143	9.98	241	9.48	184	9.28	253	9.3	244	9.35	152	9.26	223	8.7	175	8.65	71	19



82	9.12	160	9.5	118	9.35	124	9.12	144	10	240	9.56	186	9.3	254	9.36	242	9.39	158	9.24	225	8.8	180	8.64	75	20
83	9.16	162	9.47	122	9.37	126	9.16	146	10.0 8	242	9.6	190	9.35	260	9.32	255	9.39	156	9.45	231	8.97	182	8.81	75	17
84	9.18	162	9.55	125	9.45	127	9.18	148	10.1	245	9.67	192	9.34	262	9.28	253	9.46	160	9.49	236	8.97	187	8.75	79	16
85	9.16	158	9.58	127	9.4	128	9.2	150	10.1 2	247	9.7	194	9.36	264	9.32	255	9.6	161	9.66	238	8.98	190	8.82	81	16
86	9.2	164	9.6	123	9.48	129	9.27	153	10.1 6	248	9.72	196	9.4	265	9.35	260	9.65	162	9.8	240	8.9	194	8.8	83	18
87	9.18	166	9.62	128	9.43	129	9.24	156	10.1 2	249	9.78	198	9.46	268	9.4	263	9.7	161	9.54	245	8.97	200	8.7	84	18
88	9.22	169	9.6	130	9.5	131	9.3	157	10.0 2	251	9.8	200	9.47	273	9.36	264	9.96	165	10.0 7	252	9.28	198	8.97	84	19
89	9.2	170	9.66	132	9.48	134	9.28	158	9.96	253	9.84	201	9.5	278	9.32	265	10.0 7	168	9.96	263	8.7	206	8.65	88	18
90	9.26	171	9.58	134	9.52	136	9.35	162	9.9	256	9.96	206	9.52	279	9.3	267	9.96	170	9.86	268	9.65	210	8.65	90	18
91	9.3	173	9.65	134	9.56	140	9.36	165	10	258	10.1	206	9.48	280	9.32	268	9.8	176	9.88	270	9.46	212	8.7	91	18
92	9.36	176	9.68	137	9.57	142	9.3	166	9.94	259	10.1 4	210	9.46	281	9.34	270	9.94	174	10	281	9.22	208	8.86	93	17
93	9.34	178	9.7	138	9.52	144	9.35	165	10	260	10.2	216	9.4	284	9.4	271	10	184	10.1	279	9.26	218	8.9	94	18
94	9.38	175	9.73	140	9.6	140	9.38	168	10.0 3	261	10.2 3	214	9.38	286	9.32	269	10.0 2	187	10.0 6	283	9.02	220	9.07	96	19
95	9.41	178	9.74	141	9.63	147	9.3	169	10.1 4	262	10.2	218	9.4	287	9.37	275	10.1	186	9.91	285	9.12	224	9.02	98	18
96	9.43	179	9.7	141	9.64	150	9.28	170	10.1 2	265	10.2 8	218	9.42	289	9.4	278	10.1 2	192	9.98	286	8.97	226	8.86	99	18
97	9.4	180	9.78	144	9.7	154	9.32	172	10.1 4	270	10.2 6	220	9.46	290	9.42	280	10.1	196	10	282	9.02	230	8.75	100	18
98	9.42	181	9.8	147	9.68	155	9.4	174	10.2	274	10.3	224	9.44	294	9.41	284	9.98	199	10.1	290	9.12	234	8.86	102	17
99	9.44	183	9.85	149	9.7	158	9.42	176	10.2 6	276	10.2 8	226	9.48	296	9.45	282	9.95	204	10.1 2	293	9.54	236	9.07	105	18
100	9.46	181	9.84	150	9.72	158	9.43	178	10.3	283	10.3	227	9.5	298	9.43	286	10.1	206	10.2	292	9.6	238	9.05	106	19