

# Diseño e Implementación de un Sistema Experto para Supervisión y Control de una Planta de Tratamiento de Aguas Residuales de Lodos Activados

Enrique Fernández

Facultad Politécnica

Universidad Nacional de Asunción

San Lorenzo - Paraguay

enrifernandez@gmail.com

Diego P. Pinto-Roa

Facultad Politécnica

Universidad Nacional de Asunción

San Lorenzo - Paraguay

dppintoroa@gmail.com

**Abstract**—La gestión incorrecta de las aguas residuales puede provocar diversos problemas ambientales, como la contaminación de aguas utilizadas para la potabilización y su posterior consumo, o la presencia de sustancias que provocan graves perjuicios a la vida acuática. Los problemas ambientales generados por estas aguas residuales, impulsaron la necesidad de encontrar métodos de tratamientos capaces de ayudar a la depuración natural del medio ambiente, con el objeto de restablecer las condiciones originales. El aumento de los vertidos, unido a las cada vez más importantes restricciones legales, han obligado a dedicar un gran esfuerzo en la mejora y desarrollo de procesos de tratamientos, así como a su control para optimizar sus rendimientos. La introducción en las leyes de limitaciones en los vertidos de Nitrógeno y Fósforo, añadidos a las ya existentes de materia orgánica y sólidos en suspensión, ha aumentado la complejidad de los procesos de tratamientos de aguas. Esta complejidad ha provocado la necesidad de sistemas de controles más desarrollados, que permitan la monitorización y el control de los nuevos procesos utilizados. Los nuevos requerimientos incrementan el número de posibles problemas que pueden aparecer durante el funcionamiento de una planta depuradora, por lo que se hace todavía más necesaria la disponibilidad de herramientas como la monitorización con analizadores automáticos y el control de procesos para tratar adecuadamente los residuos. El presente estudio se enmarca dentro del desarrollo y utilización de herramientas de monitorización y control sobre procesos de tratamientos biológicos. En ella se plantea la implementación de un sistema experto como el elemento superior de supervisión y control en un sistema distribuido de una planta de tratamiento de efluentes. Esta arquitectura puede producir un sistema de control autónomo y capaz de adaptarse a los diferentes problemas que pueden aparecer en una planta real y son capaces de manejar las situaciones que aparecen durante la operación, diferentes a los sistemas de control convencionales, que tienen un comportamiento correcto solo en situaciones normales.

## I. INTRODUCCIÓN

Las Plantas de tratamiento de aguas residuales, PTAR, han adquirido en los últimos años un gran auge debido a la necesidad de depuración de los numerosos deshechos que, generados por las aglomeraciones urbanas y las instalaciones industriales, son vertidos a los cauces hídricos. Las legislaciones actuales imponen límites estrictos en la calidad del agua vertida por

las plantas depuradoras, siendo el objetivo de estas poder cumplirlos.

El diseño de este tipo de plantas depuradoras no representa en la actualidad un gran desafío, ya que los principios de funcionamiento de las mismas son ampliamente conocidos y su construcción se lleva a cabo de forma extendida en todo el mundo. Sin embargo, en la mayoría de los casos, el mantenimiento y control de las mismas es realizado, casi en su totalidad, por técnicos expertos que, en base a mediciones realizadas sobre las variables de la planta, corrigen y reajustan los parámetros del sistema para un funcionamiento satisfactorio. El diseño del sistema de control de estos tipos de planta depuradoras normalmente no se realiza teniendo en cuenta la dinámica compleja y las perturbaciones de carga inherentes a esta clase de procesos biológicos. Estas situaciones no sólo han de ser detectadas y diagnosticadas rápidamente, sino que también deben tenerse en cuenta en el modelo del proceso y así poder usarlo en la predicción y estimación de variables no disponibles y en el propio proceso de control.

Para el caso de una PTAR con proceso de lodos activos, obedece a una dinámica claramente no lineal y la respuesta de los sistemas no lineales depende de las condiciones iniciales, así como de la magnitud, la configuración y la forma de las entradas. Ello provoca la ausencia de un cuerpo teórico bien definido capaz de analizar y controlar su comportamiento.

Además, los cambios en las condiciones ambientales a los que está sometido un reactor aerobio, que contiene a los lodos activados en una PTAR, tales como cambios de temperatura, humedad, etc., hacen que este proceso sea variable en el tiempo.

A esta complejidad en el comportamiento dinámico se une la demanda de unas especificaciones de funcionamiento cada vez más exigentes. No sólo se requiere un comportamiento transitorio y estacionario satisfactorio, sino que se busca el mejor comportamiento posible respecto a un cierto índice de calidad final del agua depurada.

Para hacer frente a todas estas circunstancias, la teoría

clásica de control resulta insuficiente y es necesaria la aplicación de técnicas de control avanzados. La alternativa que se toma en este estudio es el del machine learning por su nombre en inglés.

El problema se enfoca en particular en una fabrica de chacinados que tiene una PTAR funcional pero con muchas limitaciones de control y operación que se realiza de manera sencilla, es por ello que los indicadores de calidad del agua de salida final se encuentra fuera de los rangos establecidos por el ente regulador.

La solución del problema de la PTAR se enfoca en dos ámbitos, en la infraestructuras civiles anexas, y el desarrollo tecnológico, que consiste en la implementación de los equipos e instrumentos, el control de los procesos y el enlace con el sistema experto que predice el comportamiento de una de las variables del proceso, la Demanda Química de Oxígeno, DQO, que es un indicador del la calidad del agua del efluente final, logrando así la estabilidad y robustez del sistema de depuración.

La *Contribución* que realiza este trabajo es a la depuración eficiente de aguas residuales a partir de la construcción de un sistema experto que se encargue de gestionar el correcto y económico funcionamiento de una planta de tratamiento de aguas residuales de lodos activados.

Desarrollando la arquitectura tecnológica necesaria para la gestión de una planta de tratamiento de aguas residuales; Implementando la supervisión y control con todos los equipos necesarios para la adquisición, procesamiento y actuación final del proceso de tratamiento y; Desarrollando e implementando un sistema experto, basado en machine learning, tomando datos de las variables del proceso, para que ayude al trabajo del operador de planta ya sea en operación normal o en situaciones de funcionamiento incorrectos

## II. PLANTEAMIENTO DEL PROBLEMA

A continuación se muestra la figura 1 la planta de tratamiento de aguas residuales con la que cuenta la fabrica de chacinados, nuestro caso de estudio, que consta de un pozo primario donde ingresa el agua residual proveniente de la fabrica, para luego ingresar a un reactor biológico a través de un filtro o criba, en el reactor biológico ocurre la depuración del agua residual por medio del lodo activado que posee microorganismos que consumen la materia orgánica y la degradan, para luego pasar a los sedimentadores secundarios que separan el lodo activado del agua depurada para una parte de estos lodos retornar y mantener estabilizado el reactor biológico y por otro lado purgar los lodos en exceso. Luego el agua depurada pasa por un filtro de arena y a un laberinto de cloración para ser vertido al cauce hídrico.

*1) Problema particular:* Para este caso de estudio, se tiene que las demandas de los productos finales son variables en el tiempo por ende la producción varían considerablemente y acompañando a esto la utilización del agua para los distintos procesos, por lo que el agua residual producida tiene grandes variaciones diarias con lo que el aporte de las composiciones, concentraciones y descargas son muy variables, lo que hace

que la PTAR sufra muchas variaciones, no opere normalmente y salga de los parámetros establecidos por el ente regulador.

- Sistema de Control:

El sistema de control actual es mínimo ya que cuenta con solamente en el pozo primario con una boya de control de nivel on-off que acciona una bomba sumergible y envía a la criba para luego ir al reactor biológico, a partir de allí el agua que ingresa va desplazándose por el reactor biológico a los sedimentadores hasta el vertido final por gravedad.

- Limitaciones de Control:

El control es muy básico ya que posee solo un control on- off de nivel, en caso de ingreso de agua en gran cantidad este tipo de control solo se encarga de bombejar el agua al reactor provocando un ingreso de gran cantidad de composiciones y concentraciones afectando considerablemente el reactor biológico donde se encuentran los microorganismos encargados de la depuración del agua. Por otro lado, el sistema de PTAR está en dependencia total del operador y su experiencia para afrontar las situaciones, por ende, no tiene el control en caso de muchas perturbaciones externas, como vertidos de pesticidas, venenos, ácidos y otros que puedan dañar a los microorganismos y a la dinámica de la PTAR como el agua de lluvia en mucha cantidad. Otras de las limitaciones es el dimensionamiento hidráulico de la PTAR que puede soportar hasta un caudal de punta limitado.

Para poder visualizar el proceso general de la PTAR se realiza un diagrama de cañerías e instrumentación de acuerdo a la norma P and ID ISA 5.1 [1], que es utilizado para entender mejor el proceso y cómo la instrumentación está interconectada. Como se puede observar en la figura 2, se tiene un controlador de nivel on-off que acciona una bomba sumergible en el pozo primario que envía la alimentación fresca de agua residual al reactor biológico y luego sigue el proceso de depuración hasta el vertido final.

- 2) *Indicadores del control de proceso por lodos activados:*

Como ya se mencionó los procesos de lodos activados son sistemas complejos donde se producen muchas reacciones y donde los microorganismos están compitiendo por los nutrientes y sustratos para su crecimiento y reproducción, cualquier des balance puede ocasionar grandes fallas y por ende contaminaciones en los cauces hídricos, es por ello que existen unos indicadores de control de lodos activados que muestran el estado en que se encuentra el licor de mezcla y las concentraciones y nutrientes necesarios, realizando estos en laboratorios mediante análisis químicos o de reacciones para sus determinaciones. Los parámetros de control de lodos activados se muestra en la tabla I.

- 3) *Indicadores de la calidad del Agua:* De acuerdo al padrón de calidad de las aguas [24] del territorio nacional establecido en la resolución Nro. 222, de la secretaría del ambiente SEAM, las aguas son clasificadas en 4 clases que van desde la clase 1 que son las más exigentes hasta la clase 4 la menos exigentes, para nuestro caso los indicadores

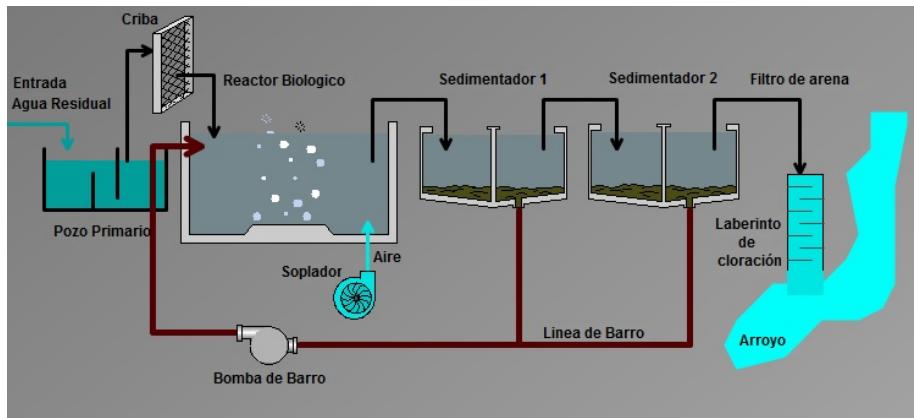


Fig. 1. Diagrama de la planta de tratamiento de agua residuales. Elaboración propia.

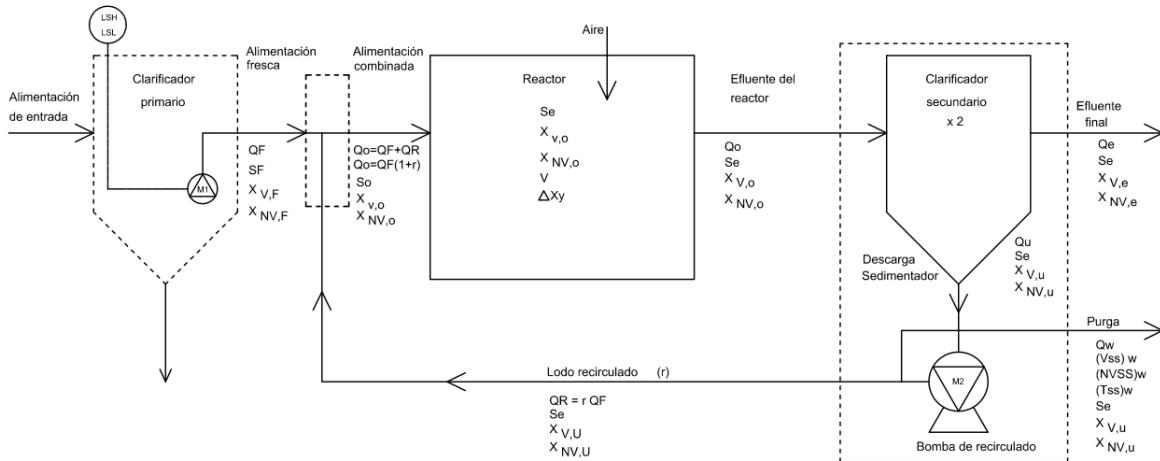


Fig. 2. Diagrama de proceso de la PTAR. Elaboración propia.

Tabla I

INDICADORES DE CONTROL DEL PROCESO DE TRATAMIENTO POR LODOS ACTIVADOS, [25]

Etapa del proceso	Tipo de pruebas	Responsable
Descarga entrada	Medición de flujo, $Q$	Operador
Descarga entrada	Demanda Biológica Oxígeno, DBO	Laboratorio
Descarga entrada	Demandra Química Oxígeno, DQO	Laboratorio
Descarga entrada	Nitrógeno, N	Laboratorio
Descarga entrada	Fósforo, P	Laboratorio
Descarga entrada	Potencial Hidrógeno, pH	Operador, lab.
Descarga entrada	Grasas y aceites, GyA	Laboratorio
Descarga entrada	Solidos Suspensidos Totales, SST	Laboratorio
Reactor biológico	Solidos Totales, ST	Laboratorio
Reactor biológico	Solidos Volátiles, SV	Laboratorio
Reactor biológico	Oxígeno disuelto, O <sub>2</sub>	Operador
Reactor biológico	Prueba de sedimentación, TS	Operador
Reactor biológico	Edad de lodos, X <sub>y</sub>	Operador, lab.
Reactor biológico	Índice Volumétrico Lodos, IVL	Operador, lab.
Reactor biológico	Comida/Microorganismos, F/M	Operador, lab.

la navegación, armonía paisajística y para los usos menos exigentes. Indicadores para el vertido al cuerpo receptor, clase 4. Estos valores de referencia, se indican en la tabla II y es la referencia hasta donde pueden llegar los niveles de contaminación de las aguas que se vierten en los cauces hídricos.

Tabla II  
INDICADORES DE CONTROL DE CALIDAD DEL EFLUENTE FINAL. [24]

Parámetros	Responsables
Oxígeno disuelto: superior a 2mg/l	Lab., operador
pH de 5 a 9	Lab., operador
DBO 5d 20C, inferior a 50 mg/l	Laboratorio
DQO, inferior a 150 mg/l	Laboratorio
Temperatura, inferior a 40 oC	Operador
Sedimentables, 1 ml/l, test de 1 hora, cono Imhoff	Operador
caudal max. 1,5 veces, caudal mín. cuerpo receptor	Operador
aceites vegetales y grasas animal hasta 50mg/l	Laboratorio
Nitrógeno total 4 N	Laboratorio
Fósforo total 4 P	Laboratorio
Coliformes fecales 4000 NMP/100ml	Laboratorio

se referencian para la clase 4 que son aguas utilizadas para

### III. TRABAJOS RELACIONADOS

Varios trabajos proponen estudios sobre como controlar ciertas variables o todas las variables intervientes en los procesos de tratamientos de aguas residuales, uno de ellos es el Environmental decision support systems (EDSSs), [6], es una sinergia entre técnicas de inteligencia artificial con métodos numéricos y estadísticos, sistemas de información geográfico y ontología ambiental, el cual está dirigido a problemas ambientales no lineales de hasta 2do orden de dificultad, que lo que propones es tomar todas estas herramientas y utilizarlas con todas las variables que intervienen en el proceso de una PTAR, dando aproximaciones de los valores de estas variables y la tendencia que van tomando, el enfoque es amplio en contraste al trabajo realizado que toma solo las variables que el operador puede medir en el proceso. Por otro lado esta otro estudio que es, Intelligent decision making in wastewater treatment plant SCADA system: La propuesta de automatización [8] de una planta de tratamiento de agua residuales envuelve la utilización de pequeños sistemas de controles como un controlador lógico programable o PLC, con los sensores y actuadores finales que se encuentran en campo, cuya lógica de funcionamiento está establecido en la programación del PLC, que se enlaza con un software de supervisión SCADA con el que se puede controlar remotamente todo el sistema, el control se basa en lógica difusa, fuzzy logic, este estudio es muy semejante encuanto a la utilizacion de sistemas distribuidos con la diferencia que utiliza la logica fuzzy y no se ocupa de ninguna variable del proceso biológico solo los caudales. Por otro lado esta el Adaptive model based control for wastewater treatment plants: El [9], se basa en un modelo adaptativo que se ocupa de optimizar el control de la aireación en el reactor biológico aerobico de modo a mantener la cantidad necesaria de oxígeno que requieren los microorganismos para que puedan disolver el amoniaco del agua residual. Este se basa en el modelo original llamado WOMBAT que predecía a partir de un nivel de O<sub>2</sub> el posible nivel de amoniaco que se tenía que luego se realimentaba para ver si coincidía con la realidad, llegó a ayudar bastante al proceso pero no siempre era muy acertado, debido a esto se realizó una variación llamandole Adaptative WOMBAT que es una modificación del original utilizando la misma ecuación solo que introduciéndole un algoritmo de predicción y con menos variables. Los resultados fueron mejorados con respecto al original, el caso de este estudio solo se ocupa de una variable que es el Oxígeno disuelto y se trata de una optimizacín dentro del proceso biológico, para nuestro estudio se utilizo esta variable como una de las variables de entrada dentro de la predicción. El otro estudio es el Control de una planta de tratamiento de aguas residuales mediante redes neuronales: En [11], el modelo neuronal propuesto busca aprender el comportamiento de los técnicos expertos para, posteriormente, reproducirlo en un modelo simulado de la planta, también buscá la optimización de las acciones de control, mediante redes neuronales, para que fueren a la planta a trabajar en una zona de rendimiento óptimo, en el estudio realizado se opto por análisis de patrones

y clasificación dentro de lo que son las acciones de corrección por medio de una variable indicadora de la calidad del agua.

### IV. PROPUESTA DE SOLUCIÓN

En esta sección se plantean soluciones inherentes al proceso de tratamiento, que son las infraestructuras necesarias para mejorar los valores de los indicadores de salida del efluente final y el desarrollo tecnológico en cuanto al proceso, teniendo en cuenta los nuevos anexos de infraestructuras. Finalmente se tiene un esquema general de la solución, y el sistema experto que se expone en detalle en la siguiente sección.

#### A. Anexos de infraestructuras

A continuación se muestra la figura 3, donde se observan los anexos de infraestructura que se colocarón para poder reducir los sólidos en las aguas residuales.

Se incluye a la configuración original, un tanque para sostener los caudales de punta y para homogeneizar las concentraciones de los contaminantes, y luego ingresar a un tanque circular en un tratamiento físico de flotación por aire disuelto para extraer una alta cantidad de sólidos por medio del aire y de polímeros.

También se colocan dos tanque de retención de Fósforo para poder hacer precipitarlo gracias al Sulfato de Aluminio, y así reducir este valor en las aguas residuales hasta un máximo de 4P, volviendo nuevamente al tratamiento biológico de lodos activados y el proceso de sedimentación, finalizando se le agrega un filtro de arena para las últimas retenciones de sólidos antes de la desinfección y vertido.

Con los nuevos agregados se procede a la diagramación del proceso de la PTAR, iniciando con el dimensionamiento de los equipos y sensores, en base a los análisis realizados en los procesos unitarios, y siguiendo con el diagrama del proceso general, con lo cual se tiene la figura 4.

#### B. Adquisición de datos y accionamientos

A partir de la figura 4, se tiene los equipos que realizarán la adquisición de datos y los accionamientos necesarios que ayudará a poder realizar la correcta elección del controlador lógico programable, PLC, que se utilizará para el control y procesamiento de estas señales.

A cada instrumento de entrada se le asigna una etiqueta única con la que será identificado en todo el proceso de implementación y el tipo de señal de entrada y salida que se utlizará en la programación del controlador.

#### C. Control, supervisión y registro, SCADA

En cuanto al controlador lógico programable, PLC, debe tener una programación para que pueda realizar la recolección de los datos enviados por lo equipos, para poder realizar los lazos de controles y para poder registras algunas variables del proceso. Es por ello que se coloca en forma genérica las parte de la programación interna que lleva este controlador. Los bloques de programación se observan en la figura 5.

Para la realización de las pantallas de interfaz entre el usuario y todo el sistema, se tomo como base [12] y a continuación se detallan las pantallas para el sistema.

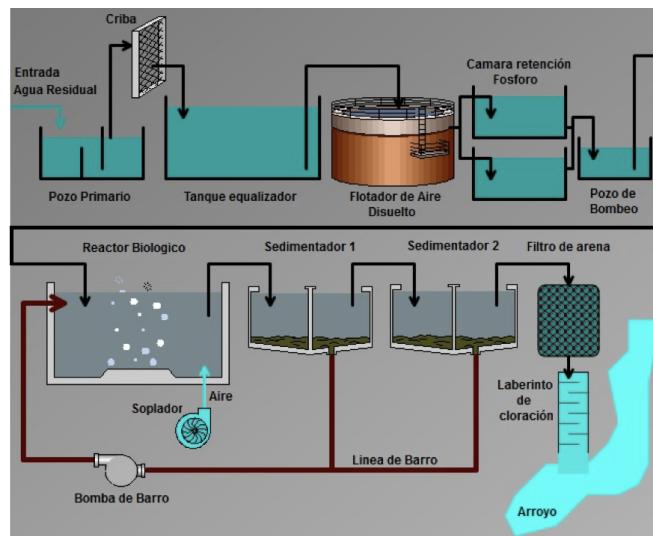


Fig. 3. Diagrama de la PTAR. Elaboración propia.

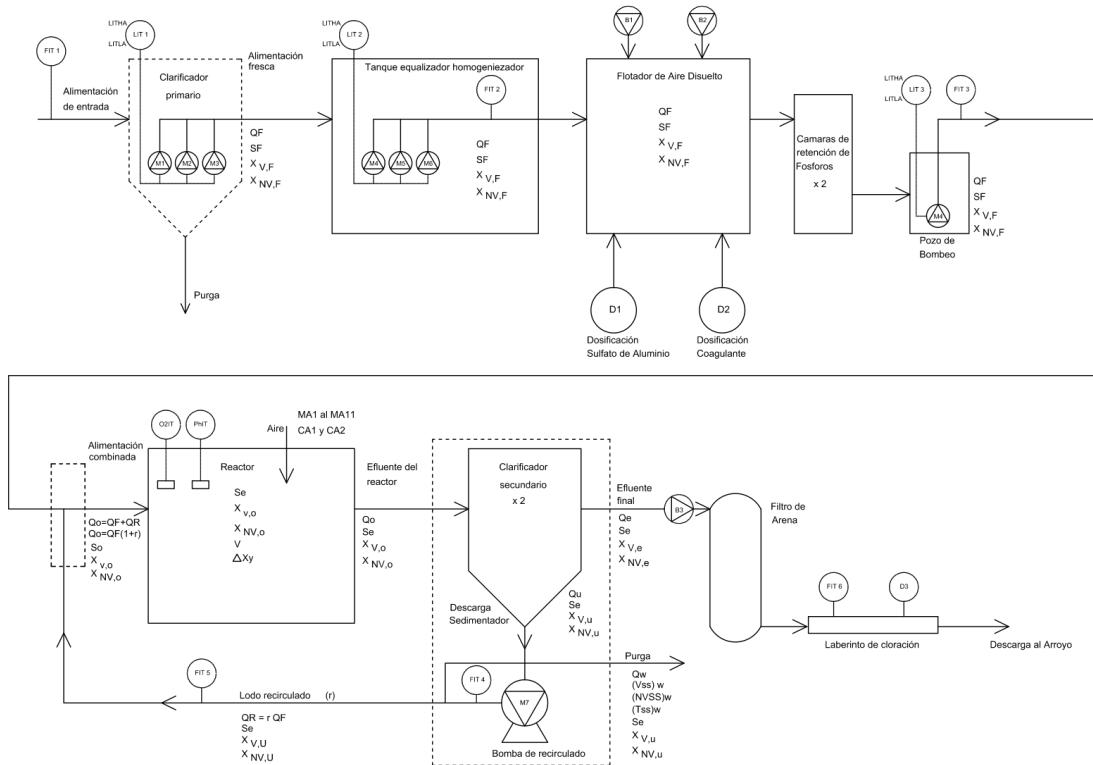


Fig. 4. Diagrama del proceso de la PTAR. Elaboración propia.

Menú general, Vista general de la PTAR, Vista general del DAF, Vista del pozo primario, instrumentación y actuadores, Vista del tanque homogeneizador, instrumentación y actuadores, Vista del DAF, instrumentación y actuadores, Vista del reactor biológico, instrumentación y actuadores, Vista de sedimentadores, instrumentación y actuadores, Vista de salida efluente final, instrumentación y actuadores.

En cuanto al registro y base de datos, el software que se utiliza para realizar las interfaces HMI, ya cuenta con un apoyo de bases de datos de SQL, con lo que los registros de datos, que vienen del proceso, se pueden colocar en archivos o bien en bases de datos para poder utilizarlo posteriormente. Las principales variables que se tomarán para poder almacenarlos son:

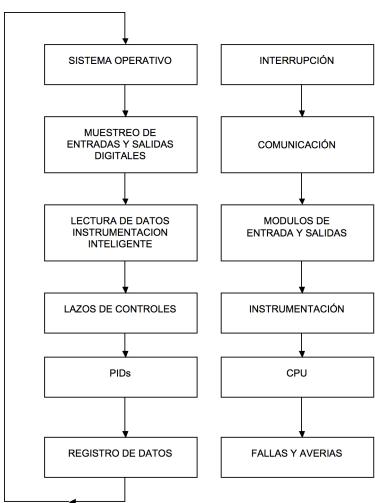


Fig. 5. Programación PLC. Elaboración propia.

Caudal de entrada primario; Oxígeno disuelto en reactor biológico; pH en reactor biológico; Caudal de salida tanque homogeneizador; Caudal de recirculado de lodos activados; Caudal de purga de lodos activados; Caudal de agua de salida de efluente final; pH de agua de salida efluente final; Oxígeno disuelto de agua de salida efluente final; Tiempo de sedimentación de lodos activados del reactor biológico y Temperatura de agua de salida efluente final.

#### D. Arquitectura de conexión y SCADA

Se presenta en las figura 6 la arquitectura de conexión entre los distintos instrumentos y equipos involucrados en el proceso de tratamiento y en la figura 7 la visualización del HMI dentro del SCADA.

#### V. MODELADO DEL PREDICTOR DE DQO Y VALIDACIÓN DEL MODELO

Se realizó un análisis de las variables del proceso, cuales son las correlaciones entre ellas y con la variable que se quiere predecir. Con esto se procede a realizar varias pruebas con la herramienta de machine learning, el support vector machine, SVM, con kernel no lineales y el algoritmos genético, dio el mejor comportamiento para un modelo de predicción de la variable requerida.

#### A. Ambiente Experimental

Se presenta las principales informaciones sobre el ambiente usado para el experimento. Además de explicar como fue construido la tabla de datos también será detallado el ambiente computacional utilizado en la construcción del modelo y la herramienta de prueba.

1) *Ambiente computacional:* Los experimentos fueron ejecutados en una maquina Intel Core i5 de 2,7 GHz y 8 GB de memoria RAM en el sistema operativo macOS Sierra version 10.12.1 y el desarrollo fue hecho en RapidMiner Studio 7.3.000. Variables de entradas y construcción de tabla de variables: Las variables tomadas para la experimento fueron:

Tabla III  
MEDICIÓN DE VARIABLES Y ANÁLISIS POR LABORATORIO DEL DQO.  
ELABORACIÓN PROPIA.

T	O2	pH	TS	DQO
29,3	2,5	6,98	550	33
28,8	3,2	7,01	570	22
28,4	2,4	6,65	550	40
27	4,2	6,1	600	19
24,3	4,3	6,17	600	16
26,4	3,9	5,08	400	4
25	4,9	5,65	400	29
26,9	4,1	5,15	350	13
27,3	3,4	7,04	150	364
27,8	3,6	7,2	175	490
27	2,8	7,19	250	361
28,5	2,4	7,02	275	517
28,1	4,2	7,12	310	320
28,9	2,4	7,16	275	155
27,1	3,6	6,8	225	69
29,6	2,6	7,21	225	113
28,5	3,9	7,14	400	131
27,6	2,6	6,94	425	387
29,1	2,4	6,89	375	316
28,3	2,4	7,05	400	266
29,7	2,3	6,96	375	253
29	2,8	6,73	425	262
29,7	3,1	6,95	450	207
30,3	4,8	7,23	425	160
30	2,9	6,74	400	157
28,4	3,3	7,13	400	80
29,4	2,6	6,59	400	40

- 1) Temperatura (T), del agua de salida final.
- 2) Oxígeno disuelto (O2), del licor de mezcla, reactor biológico.
- 3) Potencial hidrógeno (pH), del agua de salida final.
- 4) Tiempo de sedimentación (TS), del licor de mezcla, reactor biológico.
- 5) Demanda química de Oxígeno, (DQO), del agua de salida final.

Para la validación y simulación del modelo fue realizado una tabla con las mediciones de las variables tomadas del proceso.

Para las mediciones y los análisis de estas variables, se procedió a tomar muestras del agua de salida final y del licor de mezcla, todos los días durante 4 semanas, con lo que se llegó a 26 muestras. Estas muestras de agua de salidas son sometidas a mediciones de pH y temperatura por medio de equipos [22] que miden este tipos de variables y están calibrados y trazados de acuerdo a estandares, también con el agua de salida se realiza la prueba de laboratorio para determinar su DQO, que se hace por medio de un equipo [23] que posee los viales con líquidos para determinación de DQO, un reactor de temperatura para la digestión del líquido con la muestra de agua y por ultimo un equipo fotoespectrómetro que mide la determinación de la DQO, este equipo esta homologado y calibrado de acuerdo a estandares A continuación se muestra en la tabla III los valores obtenidos.

Con estos datos se realiza la correlaciones entres las variables, tabla IV, para tener una certeza de que tan correlacionadas se encuentras las variables entre si y con la variable

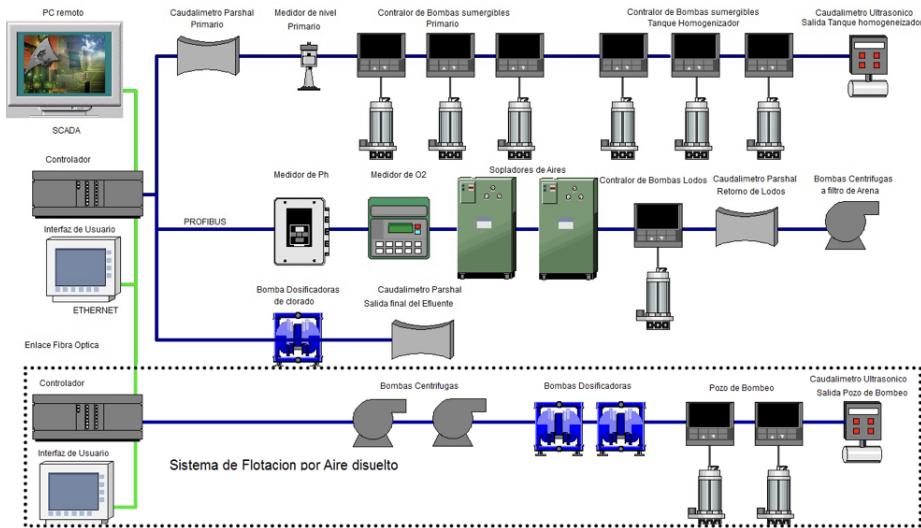


Fig. 6. Conexiones entre intrumentación inteligente, controlador y HMI. Elaboración propia.



Fig. 7. Pantalla HMI. Elaboración propia

Tabla IV  
CORRELACIÓN ENTRE LAS VARIABLES Y LA DQO. ELABORACIÓN PROPIA.

	DQO	pH	TS	O2	T
DQO	x				
pH	0,54	x			
TS	-0,58	-0,3	x		
O2	-0,3	-0,42	0,13	x	
T	0,2	0,6	-0,1	-0,5	x

de salida, DQO.

2) *Modelado del predictor y datos de entrenamiento:* Para que fuese posible realizar las pruebas con diferentes métodos de predicción, fue necesario realizar una arquitectura de modelado dentro del programa RapidMiner [21], que sirvió de herramienta de prueba para encontrar el óptimo entre los métodos. El modelado resultante se observa en la figura 8.

Se trabajo con el SVM con AG por que arrojaba el error más bajo y donde se ataca un problema de regresión con valores continuos.

Los datos para el entrenamiento fueron tomados de la tabla III que se utilizó para realizar los ajustes en los parámetros del SVM a partir de la iteración del AG para encontrar los mejores parámetros para el SVM, el núcleo utilizado fue el de RFB (radial function basics) que arrojo los siguientes valores que se muestra en la figura 9.

Este modelo de núcleo k, arrojo un rendimiento como se muestra a continuación, figura 10.

Como se puede observar se tiene un error relativo entre 47.5 y 27 porcento del valor a predecir que es nuestro caso es el DQO.

3) *Parámetros evolutivos:* Se optimizaron los modelos con Validation Set (separar el dataset en 2 sets, training set (70 porcentaje de los datos) y test set (30 porcentaje restante)).

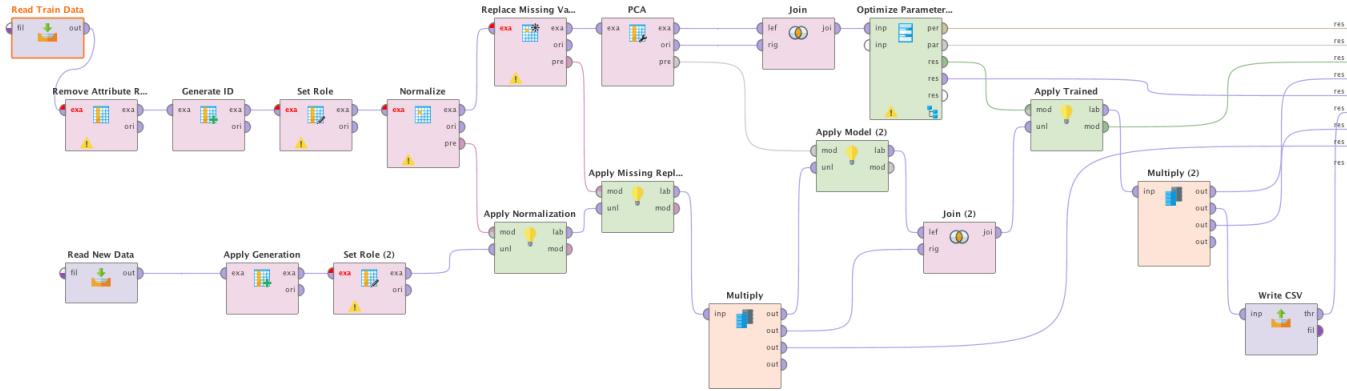


Fig. 8. Modelado en RapidMiner para pruebas con la herramienta de machine learning. Elaboración propia.

**Kernel Model**

```
Total number of Support Vectors: 43
Bias (offset): -158.240
w[pc_1] = 153053.728
w[pc_2] = -80818.410
w[pc_3] = 71295.948
w[Oxigeno Disuelto] = 339367.934
w[Potencial Hidrogeno] = 725772.528
w[Temperatura] = 645010.028
w[Solidos T] = 285845.198
number of classes: 2
number of support vectors: 43
```

Fig. 9. Modelo del núcleo, k. Salida del software RapidMiner.

**PerformanceVector**

```
PerformanceVector:
root_mean_squared_error: 217.090 +/- 0.000
absolute_error: 162.475 +/- 143.979
relative_error: 47.50% +/- 27.06%
squared_error: 47128.073 +/- 55556.358
```

Fig. 10. Rendimiento del SVM. Salida del software RapidMiner

El GA procesó 50 generaciones de 50 individuos cada uno. Selección por torneo binario y probabilidad de mutación de 1 a 5 porciento. Se busca optimizar el error relativo.

Se muestra la figura 11, sobre los conjuntos de parámetros de nuestro modelo que arroja los mejores valores de C y  $\gamma$  para el SVM.

**ParameterSet**

```
Parameter set:
Performance:
PerformanceVector [
----root_mean_squared_error: 217.090 +/- 0.000
----absolute_error: 162.475 +/- 143.979
****relative_error: 47.50% +/- 27.06%
----squared_error: 47128.073 +/- 55556.358
]
SVM (3).gamma = 1423.213019413879
SVM (3).C = 773.6251506535501
```

Fig. 11. Parameter Set. Salida del software RapidMiner

**4) Datos de pruebas:** Para los datos de pruebas y validación se utilizaron mediciones y datos anteriores de más de 1 año que se tenían registrados para los valores de TS y O<sub>2</sub>, para los valores de pH y T, se utilizaron registros de análisis de laboratorios y se pudo obtener 5 juegos de valores de estas variables, que se observa en la tabla V.

**Tabla V**  
**DATOS DE PRUEBAS Y VALIDACIÓN. ELABORACIÓN PROPIA.**

O2	pH	T	TS	DQO
2	6.86	30.8	660	133
2.6	6.68	29.5	640	93
2.2	6.98	29.9	620	122
3.2	5.6	27	500	43
4.2	5.08	26.4	470	4

**5) Resultados experimentales:** Para los resultados experimentales se utilizaron los datos de entrenamientos y los datos de pruebas.

Se introducen los datos de entrenamientos y se procede a encontrar el porcentaje de error, esto se observa en la figura 12.

Con el resultado del error, se tiene que el predictor para este caso es muy preciso ya que arroja un error de 0.0012 porciento, esto indica que si la medida de las variables reales del proceso se acercan a los valores de entrenamientos se infiere que el error tiende a ser bajo, asegurando así que el predictor cumple con las expectativas del proceso y confiabilidad.

Para los datos de pruebas los resultados se observan en la figura 13:

El error en estos 5 juegos de datos es de 0.12 porciento que es un error bajo y está dentro de la tolerancia que se manejan en los análisis laboratoriales que es del 1 porciento.

**6) Implementación del predictor:** Para la implementación del predictor se observa en el modelado que se adquieren los valores de las variables de un fichero input.csv y el resultado de la predicción se envía a un fichero output.csv, que son ficheros de texto. Con una aplicación en Java que se ejecuta desde la consola del sistema, se toma el fichero de modelado del RapidMiner y adquiriendo estos datos desde el fichero input.csv que es proveído desde el SCADA donde se registran las 4 variables del proceso, se realiza la predicción y se crea el fichero ouput.csv, que el SCADA lo vuelve a leer y lo registra para luego ser visualizado dentro de las variables del proceso de la PTAR. Al tener el fichero output.csv registrado dentro

	O2	pH	T	TS	id	prediction(DQO)	DQO	% Error DQO
0.06451612903225803	0.6289752650176679	0.8552631578947371	0.5643564356435643	1.0	133,09968551476000	133		
0.25806451612903225	0.5653710247349822	0.6842105263157896	0.5445544554455446	2.0	93,10043912036690	93,0		
0.12903225806451615	0.6713780918727916	0.7368421052631579	0.5247524752475248	3.0	122,09970403556700	122,0		
0.41055718475073035	1.0	0.5445046439628484	0.504950495049505	4.0	116,30043912036600	116		
0.7741935483870968	0.0	0.276315789473684	0.37623762376237624	10.0	4,09970064688715	4		
0.0	0.5618374558303886	0.8815789473684212	0.0	16.0	426,89968551738000	427		
0.22580645161290316	0.6713780918727916	0.6578947368421054	0.45544554455445546	17.0	33,0998778436310	33		
0.45161290322580644	0.68197895865723	0.5921052631578949	0.4752475247524752	18.0	22,09989273532090	22		
0.1935483870967741	0.5547703180212015	0.5394736842105262	0.45544554455445546	19.0	40,09970403607310	40		
0.7741935483870968	0.3604240282865511	0.35526315789473684	0.504950495049505	20.0	19,10048487320290	19		
0.8064516129032256	0.385159010607067	0.0	0.504950495049505	21.0	16,10043957802230	16		
0.6774193548387094	0.276315789473684	0.3069306930693069	0.22.0	4,09970203182399	4			
1.0	0.20141342756183755	0.09210526315789466	0.3069306930693069	23.0	29,1000794243510	29		
0.41055718475073036	0.024734982332155576	0.3421052631578946	0.25742574257425743	24.0	13,10043133026400	13		
0.5161290322580644	0.6925795053003533	0.3947368421052633	0.0594059405940594	25.0	363,89970064688700	364		
0.5806451612903225	0.749116607738516	0.46052631578947384	0.08415841584158416	26.0	489,90048487197800	490		
0.3225806451612902	0.745583038869258	0.35526315789473684	0.15841584158415842	27.0	360,89970203182300	361		
0.1935483870967741	0.6855123674911658	0.5526315789473685	0.18316831683168316	28.0	516,89972707605300	517		
0.7741935483870968	0.720840565371025	0.5000000000000002	0.21782178217821782	29.0	319,90009151924500	320		
0.1935483870967741	0.734982332155477	0.6052631578947367	0.18316831683168316	30.0	155,09972302700900	155		
0.5806451612903225	0.6077738515901059	0.36842105263157915	0.133663366336633637	31.0	69,10043133026400	69		
0.25806451612903225	0.7526501766784451	0.6973684210526319	0.133663366336633637	32.0	113,09970403607300	113		
0.6774193548387094	0.7279151943462896	0.5526315789473685	0.03069306930693069	33.0	131,09972302448400	131		
0.25806451612903225	0.6572438162544171	0.4342105263157897	0.3316831683168317	34.0	386,90045102371000	387		
0.1935483870967741	0.6395759717314486	0.6315789473684214	0.28217821782178215	35.0	315,90043957802200	316		
0.1935483870967741	0.6395759717314486	0.6315789473684214	0.28217821782178215	36.0	265,90048487197800	266		
0.16129032258064507	0.6643109546036042	0.701526315789473604	0.28217821782178215	37.0	252,90007849219800	253		
0.3225806451612902	0.5830388692579507	0.618421052631579	0.3316831683168317	38.0	261,90043910974100	262		
0.4193548387096774	0.660773851590106	0.7105263157894737	0.3564356435643564	39.0	206,89968551476000	207		
0.9677419354838708	0.7597173144876326	0.7894736842105265	0.3316831683168317	40.0	159,89970403556700	160		
0.35483870967741926	0.5865724381625442	0.7500000000000001	0.3069306930693069	41.0	157,09970403607300	157		
0.4838709677419353	0.724381625441696	0.5394736842105262	0.3069306930693069	42.0	80,09989778438310	80		
0.25806451612903225	0.5335689045936395	0.6710526315789472	0.3069306930693069	43.0	40,0996851738030	40		
Promedio					395,4997233307	395,0	0,126388178987341	

Fig. 12. Datos de entrenamientos con porcentaje de error del predictor. Salida de la aplicación del modelado.

Oxígeno Disuelto	Potencial Hidrógeno	Temperatura	Sólidos T	id	prediction(DQO)	DQO Medido	% Error DQO
0.06451612903225803	0.6289752650176679	0.8552631578947371	0.5643564356435643	1.0	133,0996855	133,0	
0.25806451612903225	0.5653710247349822	0.6842105263157896	0.5445544554455446	2.0	93,10043912	93,0	
0.12903225806451615	0.6713780918727916	0.7368421052631579	0.5247524752475248	3.0	122,099704	122,0	
0.45161290322580644	0.1831578947368421	0.35526315789473684	0.40594059405940597	9.0	43,09970404	43,0	
0.7741935483870968	0.0	0.276315789473684	0.37623762376237624	10.0	4,099700647	4,0	
Promedio					395,4997233307	395,0	0,126388178987341

Fig. 13. Datos de pruebas y porcentaje de error del predictor. Salida de la aplicación del modelado.

del SCADA ya se pueden realizar la retroalimentación para poder cambiar parámetros de control dentro de la dinámica de la planta de tratamiento de efluentes.

En la figura 14 se observa el esquema de implementación del predictor .

#### B. Mejoras introducidas por el predictor

- Alta velocidad de respuesta en control de contaminación al cauce hídrico.
- Tiempos bajos en mantenimiento de los valores de las variables de calidad del agua en las referencias deseables.
- Ahorro por gestión de análisis de laboratorios.
- Alta eficiencia en el proceso de depuración gracias a la estabilidad y robustes del parámetro de retroalimentación.

#### VI. CONCLUSIÓN

Para las conclusiones finales se referencia el trabajo realizado a la contribución propuesta inicialmente.

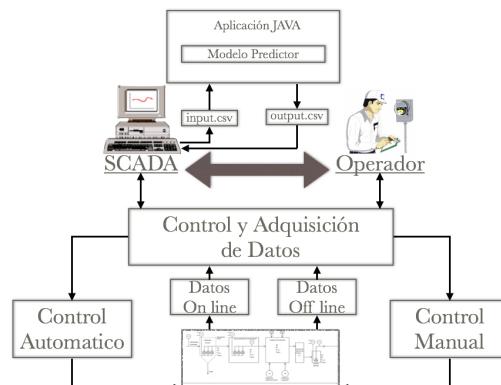


Fig. 14. Implementación del predictor al control del proceso de la PTAR. Elaboración propia.

En cuanto a desarrollar una arquitectura tecnológica dentro de la infraestructura para poder gestionar las variables del proceso, se llegó a la realización del dimensionamiento y la selección de todos los equipos de mediciones y de accionamiento requeridos para los procesos unitarios, y la implementación de estos equipos en cuanto enlaces con el sistema de control, desarrollando así una arquitectura para todo el proceso.

Para la supervisión y control, se implementó un sistema SCADA que con el controlador lógico programable, PLC, se encargaron de realizar el procesamiento de las señales, la adquisición de los datos del proceso y el control final a través de los lazos de controles para cada unidad del proceso a través de la programación del PLC, también se logró eficiencias y optimizaciones en el procesos de lodos activados gracias a la introducción de equipos de medición de Oxígeno disuelto, que contribuyeron a reducir costos energéticos.

Para el desarrollo e implementación de un sistema experto, basado en machine learning, con datos de las variables del proceso, se logró con mucho éxito desarrollar un predictor de unas de las variables que indica la calidad del agua, gracias a un detallado trabajo de recolección y análisis de datos de variables del proceso de las aguas residuales y el desarrollo de un modelo basado en SVM con AG que fue entrenado y probado con variables reales. Se obtuvo resultados experimentales con bajos porcentajes de errores que hacen que este modelo sea confiable en cuanto a sus predicciones. Se pudo enlazar este modelo con el sistema SCADA de modo a introducir una retroalimentación en la parte de control y ayuda para el operador.

Como última mención, el trabajo en forma indirecta hace énfasis a la concientización de la utilización racional del agua en todos los niveles de las actividades humanas y como la tecnología bien utilizada mitiga el impacto sobre los recursos naturales .

## REFERENCIAS

- [1] NORMAS ANSI / ISA S 5.1- 1984 (R1992), ISA, 1992.
- [2] Adaptive Control, K.J. Astrom, B. Wittemark, 1995., Addison-Wesley
- [3] Handbook of Intelligent Control: Neural, Fuzzy and Adaptive Approaches, D.A. White, D.A. Sofge, 1992., Multiscience Press
- [4] Intelligent control System: Theory and Applications, Eds. Gupta and N. Sinha, 1996. IEEE Press
- [5] IFAC, Artificial Intelligence in Real Time Control, E. Camacho and M. Arahal, Valencia, Spain, 1994., Neural Network based Adaptive Control
- [6] Designing and building real environmental decision support systems, Poch, Manel and Comas, Joaquim and Rodríguez-Roda, Ignasi and Sanchez-Marre, Miquel and Cortés, Ulises, Environmental Modelling & Software, 19, 9, 857–873, 2004, Elsevier
- [7] Application of expert systems in wastewater treatment, Mikosz, J, 2001
- [8] Intelligent Decision Making in Waste Water Treatment Plant SCADA System, Nikolić, Vlastimir and Ćojbašić, Žarko and Ćirić, Ivan and Petrović, Emina, Facta Universitatis Series: Automatic Control and Robotics, 9, 1, 69–77, 2010
- [9] Adaptive model based control for wastewater treatment plants, Niet, Arie and Vrugt, Maartje and Korving, Hans and Boucherie, Richard J, 2011, Exeter University
- [10] Sistema experto para el tratamiento de aguas residuales (SECTRARES), López, Yoan Martínez, Revista Ingeniería Agrícola, 4, 3, 51–55, 2017
- [11] Control de una planta de tratamiento de aguas residuales mediante redes neuronales, Chamizo, Juan Carlos Peguero, 2003, Universidad de Extremadura

- [12] ISA-TR77.60.40, Fuel Power Plant Human Machine Interface (HMI), ISA, 1996.
- [13] Models in advanced wastewater treatment plant control, Artículo, Peter A. VANROLLEGHEM, BIOMATH Department for Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure, 2003., Colloque Automatique et Agronomie Montpellier
- [14] Las Maquinas de soporte Vectorial, SVM, Gustavo Betancourt, 2005., Scientia et Technica
- [15] The nature of statistical learning theory, V.N. Vapnik., 1995., New York: Springer-Verlag
- [16] Genetic Algorithm for Support Vector Machines Optimization in Probability of Default Prediction, Wolfgang Hrdle and Dedy Dwi Prastyo, exposición, Humboldt Universität zu Berlin, 2013, Center for Applied Statistics and Economics
- [17] Text Classification using String Kernels, H. Lodhi, C. Saunders, J. Shawe-Taylor, N. Cristianini, and C. Watkins, 2002, Journal of Machine Learning Research
- [18] A GA based feature selection and parameters optimization for support vector machines, Cheng Lung Huang and Chieh Jen Wang, 2006., IEEE Communications Magazine
- [19] Genetic Algorithms for Support Vector Machine Model Selection, Stefan Lessmann, Robert Stahlbock and Sven F. Crone, 2006., International Joint Conference on Neural Networks
- [20] Prediccion de series temporales usando maquinas de vectores de soporte, Juan D. Velasquez, Yris Olaya and Carlos J. Franco, Artículo, Facultad de Ingeniería, Universidad de Chile, 2010., Revista de Ingeniería
- [21] RapidMiner 7.3 Data Science Platform, <https://rapidminer.com/products/studio/>, 2016.
- [22] Equipos de medición de ph, O<sub>2</sub>, temperatura, Manual, 2014., HACHT COP.
- [23] COD reagent, reactor and Multiparameter Photometer, Manual, 2012., HANNA COP.
- [24] Resolución Nro. 222, Padrón, Secretaría del Medio Ambiente, 2002., SEAM
- [25] Operación y Mantenimiento de Plantas de Tratamiento de Aguas Residuales con el Proceso de Lodos Activados, Armando Marín Ocampo and Manuel Osés Pérez, Manual de procedimientos, Dirección de Operación de Plantas de Tratamiento de Aguas Residuales (DOP), 2013, Comisión Estatal del Agua de Jalisco