



**APLICACIÓN DE MÁQUINAS DE SOPORTE VECTORIAL
PARA LA PREDICCIÓN DEL VOLUMEN DE FLUIDO
DESPLAZADO POR UNA BOMBA
ELECTROSUMERGIBLE**

Presentado para obtener el título de

**MAGÍSTER EN MATEMÁTICAS APLICADAS Y
CIENCIAS DE LA COMPUTACIÓN**

Jose Alberto Murcia Navas

Dirección:

Alexander Caicedo

Edgar Jose Andrade

Universidad del Rosario

Escuela de Ingeniería, Ciencia y Tecnología

Maestría en Matemáticas Aplicadas y Ciencias de la Computación

ABSTRACT

La explotación y producción de crudo en los campos petroleros en Colombia presenta varios retos, entre estos, la confiabilidad en la medición de los fluidos producidos y la capacidad de respuesta frente a caídas de producción.

La determinación del volumen de fluido que es desplazado por un equipo de levantamiento artificial (Bomba BES) es fundamental para realizar un control de producción adecuado. Ya que, con estos valores es posible proyectar la producción diaria del campo e identificar pérdidas de eficiencia en los equipos de levantamiento artificial, lo cual tiene asociada una caída de la producción debido a que llegan menos volúmenes de crudo y agua a las estaciones de recolección y tratamiento, y un incumplimiento con el pronóstico de producción del día.

Con el desarrollo de este proyecto es posible estimar el volumen de fluido que aporta o desplaza un equipo BES en función a sus variables y condiciones operativas. El modelo de aprendizaje de máquina entrenado corresponde a máquinas de soporte vectorial (SVM) y puede ser usado en sistemas multifásicos agua-crudo con poca producción de gas. Para el entrenamiento del modelo se tuvo en cuenta los registros de los 11 sensores instalados en los equipos BES tanto en fondo como en superficie, y se lograron resultados confiables para aquellos que desplazan fluidos con poca o nula presencia de gas.

English

The exploitation and production of crude oil in the oil fields in Colombia present several challenges, among them, the reliability in the measurement of the fluids produced and the capacity to respond to production drops.

The determination of the volume of fluid that is displaced by artificial lift equipment (BES pump) is essential to carry out adequate production control. With these values, it is possible to project the daily production of the field and identify efficiency losses in the artificial lift equipment, which is associated with a drop in production since fewer volumes of crude oil and water arrive at the collection stations and treatment, and a non-compliance with the day's production forecast.

With the development of this project, it is possible to estimate the volume of fluid that a BES unit contributes or displaces based on its variables and operating conditions. The trained machine learning model corresponds to support vector machines (SVM) and can be used in multiphase oil-water systems with little gas production. For the training of the model, the records of the 11 sensors installed in the BES equipment both on the bottom and on the surface were considered, and reliable results were achieved for those that displace fluids with little or no presence of gas.

TABLA DE CONTENIDO

CAPÍTULO 1 INTRODUCCIÓN	1
CAPÍTULO 2 OBJETIVOS	3
1.1 OBJETIVO GENERAL	3
1.2 OBJETIVOS ESPECÍFICOS.....	3
CAPÍTULO 3 PROBLEMA Y JUSTIFICACIÓN.....	4
CAPÍTULO 4 MARCO TEÓRICO Y ESTADO DEL ARTE.....	6
CAPÍTULO 5 METODOLOGÍA	9
CAPÍTULO 6 RESULTADOS Y DISCUSIÓN	12
CAPÍTULO 7 CONCLUSIONES Y RECOMENDACIONES.....	13
REFERENCIAS.....	15
APÉNDICE.....	16

LISTA DE TABLAS

Tabla 1. Muestra del conjunto de datos de uno de los equipos BES, usado en el modelo.	16
Tabla 2. Descripción de los parámetros operativos y de yacimiento usados en el modelo, los valores corresponden al instante de medición.	17
Tabla 3. Volúmenes de fluido total liquidado usados para el entrenamiento del modelo	18
Tabla 4. Volúmenes de fluido total liquidado usados para la evaluación del modelo.....	19
Tabla 5. Métricas de los dos modelos entrenados.....	19
Tabla 6. Correlación de Pearson para el modelo SVM.....	20
Tabla 7. Características del fluido movilizado por las bombas BES Serie 400.....	20

LISTA DE FIGURAS

Figura 1. Ilustración del preprocesamiento del conjunto de datos de entrenamiento.	21
Figura 2. Gráfico cruzado usando el conjunto de datos de evaluación ($R^2=0.93$).	22
Figura 3. Gráfico cruzado usando el conjunto de datos de entrenamiento ($R^2=0.92$).	22
Figura 4. Intervalos de confianza del modelo SVM.	22
Figura 5. Distribución normal del error para el conjunto de datos de entrenamiento.	23

Capítulo 1

INTRODUCCIÓN

Con la implementación de los sistemas de levantamiento artificial (ALS), la producción de crudo, agua y gas está sujeta al desempeño de estos sistemas. Uno de ellos son las bombas electrosumergibles (BES) y corresponde a nuestro caso de estudio. En la actualidad la liquidación de los volúmenes de agua, crudo y gas de un pozo productor se realiza alineándolo de manera independiente a un tanque vacío (o con nivel conocido), y luego se miden el volumen del fluido producido que desplaza durante 10 horas o menos. Una vez transcurrido este tiempo, se cierra el tanque y se determina el volumen total (crudo y agua) que aportó el pozo y que podría aportar en 24 horas para definir la producción diaria.

Realizar lo anterior requiere infraestructura como tanques y vasijas como separadores bi/tri-fásicos, lo cual limita el número de pruebas que se pueden realizar a los pozos productores del campo en el término de un mes. En algunos campos de producción de petróleo suele realizarse una sola prueba o medición de volumen de fluido total por mes para cada equipo y sólo hasta esa instancia se puede conocer si hubo pérdida de eficiencia del equipo BES y por ende una menor producción de crudo del campo. Así mismo, los tiempos de respuesta por parte de los profesionales de control de producción del campo no son los deseados, debido a que toma días justificar las caídas de producción, cuando éstas se presentan.

A partir de esta situación, vamos a desarrollar el presente proyecto haciendo uso de las variables de operación del equipo BES que se pueden monitorear en tiempo real, y algunas condiciones del yacimiento con el fin de entrenar un modelo de máquinas de soporte vectorial (SVM) que nos permita estimar el volumen de fluido desplazado por un equipo BES. La decisión de usar SVM además de ser un modelo muy potente y versátil, capaz de realizar clasificación lineal o no lineal y regresión[1], obedece a que el conjunto de observaciones de nuestra variable objetivo es pequeño (32 en total) lo cual lo hace oportuno para el presente trabajo.

Si el modelo funciona, los profesionales de control de producción del campo petrolero podrán reducir significativamente el tiempo invertido para determinar donde se genera la caída o incremento de producción, mejorando así la gestión frente a la operación del equipo BES y la optimización de los tiempos de respuesta, lo que permitirá asegurar el cumplimiento diario del pronóstico de producción del campo. Además, impactará positivamente en las finanzas del campo petrolero, ya que el tiempo invertido en identificar donde se presenta una caída de producción tiene asociado un volumen de fluido que no llega a superficie y una consecuente pérdida económica correspondiente a ese volumen de crudo que no se vende y hacía parte del pronóstico de producción del día.

Teniendo en cuenta todo esto, se presenta el modelo de máquinas de soporte vectorial de regresión, que permite estimar el volumen de fluido total (agua y crudo) que desplaza un equipo BES.

Capítulo 2

OBJETIVOS

1.1 Objetivo general

Desarrollar un modelo de máquinas de soporte vectorial para la predicción del volumen de fluido que desplaza una bomba electrosumergible (BES) a partir de sus variables de operación y condiciones del yacimiento donde se encuentra instalado.

1.2 Objetivos específicos

- Definir un modelo de máquinas de soporte vectorial (SVM) de regresión, para estimar el volumen de fluido total que puede desplazar un equipo BES serie 400, instalado en fondo de un pozo productor de petróleo.
- Determinar el desempeño del modelo entrenado, en la predicción del volumen de fluido producido o movilizado por el equipo BES.
- Sugerir ajustes al modelo entrenado.

Capítulo 3

PROBLEMA Y JUSTIFICACIÓN

Una de las tareas claves de los ingenieros de control de producción de un campo petrolero consiste en establecer los pronósticos de producción. Son con estas promesas de producción con las cuales se gestionan recursos económicos en pro del desarrollo del campo o se suspenden los proyectos de maduración y expansión [2].

Tradicionalmente se han usado métodos de simulación dinámica, análisis de curvas de declinación (DCA) y tablas de correlación presentadas en los manuales técnicos de referencia para el diseño de bombas sumergibles. Sin embargo, estos métodos resultan dispendiosos y son muy sensibles al error humano debido a la complejidad del sistema en sí y al número de variables del proceso que se debe tener en cuenta.

Los separadores de prueba – también conocidos como sets de pruebas - siguen siendo el equipo de liquidación de volúmenes convencional y más confiable para los pozos productores de crudo. En ellos se puede separar el agua, el crudo y el gas de manera independiente y posteriormente liquidar los volúmenes. Sin embargo, en campos petroleros maduros con un alto número de pozos productores provistos de sistemas de levantamiento artificial se requiere de varios separadores y tanques de prueba, lo cual representa una inversión de alto costo, y es por ello por lo que generalmente se cuenta con pocos sets de pruebas y se limitan las mediciones o liquidaciones de volúmenes de fluido total de los pozos a una (1) medición por mes. Así mismo, el transporte de los separadores de prueba sigue siendo un problema y lo que limita su uso [3].

Así las cosas, identificar en cuál de los equipos BES instalados en un campo petrolero se presenta la caída o incremento de aporte de crudo que permite cumplir o incumplir con el pronóstico de producción del día, es el reto.

Ya que, en la actualidad, identificar el equipo BES que no está aportando el volumen esperado puede tomar entre tres y cinco días, teniendo en cuenta que se siguen las siguientes etapas:

- Revisiones de desempeño de los equipos BES de los pozos. Esto se realiza con los proveedores de las bombas de levantamiento artificial y en esta etapa se revisan las variables de operación que pudieran estar asociadas a la pérdida de eficiencia del equipo.
- Toma de muestras de los pozos identificados con pérdida de eficiencia del equipo BES, para que el laboratorio analice y determine las fracciones de crudo y agua.
- Reporte de resultados del laboratorio, con los cuales se pueden hacer comparativos respecto a registros anteriores e identificar donde hubo la pérdida o incremento de producción.

Las demoras en identificar el equipo BES que no aporta o moviliza menos fluido, representa un volumen de crudo que no se vende y una pérdida económica promedio diaria de USD \$10,000 que puede ascender a los USD \$ 50,000 de acuerdo con el tiempo que tome identificar donde se está generando la pérdida de producción. Si el modelo de aprendizaje de máquina que se propone en este documento funciona, estas cifras deberán reducirse y parte del dinero ahorrado por este concepto, podrá ser destinado a proyectos de inversión social con las comunidades de la zona de influencia del campo.

Adicionalmente, la eficiencia energética representa uno de los grandes objetivos por parte de las operadoras petroleras en Colombia con el fin de reducir la huella de carbono, y una vez se pueda identificar oportunamente los equipos de levantamiento artificial con bajo desempeño, se podrán tomar decisiones que garanticen el uso responsable de la energía en dichos equipos.

Capítulo 4

MARCO TEÓRICO Y ESTADO DEL ARTE

Con el paso del tiempo, la energía de los reservorios (yacimientos petrolíferos) decrece y no es suficiente para hacer llegar el crudo hasta superficie (lo que se conoce como levantamiento natural). Esta condición afecta la presión de fondo de pozo y la capacidad de movilizar el aceite, lo que se traduce en una caída o pérdida de producción del reservorio. Cuando esto sucede, se recurre a métodos de levantamiento artificial (ALS por sus siglas en inglés) ampliamente usados en la industria de petróleo y gas para mejorar la productividad del reservorio [4].

Algunos de los sistemas de levantamiento artificial más conocidos son: levantamiento artificial por gas (gas lift), bombas de varilla de bombeo (SRP), bombas electrosumergibles (ESP), bombas de cavidades progresivas (PCP), e inyección hidráulica mediante bomba (HP) [5].

En lo que respecta a bombas electrosumergibles, el conjunto del equipo BES instalado en fondo de pozo generalmente incluye la bomba BES, un controlador/separador de gas o una entrada estándar, un sello, un motor sumergible con cable de alimentación conectado que llega a la superficie y un paquete de sensores de fondo de pozo, que registra y comunica a superficie el rendimiento de la bomba y las condiciones de fondo de pozo. El equipo de superficie incluye un controlador de velocidad variables (VSD) con el que se ajusta la frecuencia del equipo [6].

Lo anterior corresponde al conjunto de herramientas que se instalan en fondo del pozo productor con el fin de extraer el crudo desde el yacimiento hasta las facilidades de superficie para su posterior venta. La cuestión ahora es lograr medir los caudales de crudo y agua que aporta el pozo productor.

Las mediciones de las tasas de crudo son importantes para el campo petrolero, ya que permiten evaluar el desempeño de los pozos, diseñar las operaciones de campo, calcular el factor de recobro, monitorear los yacimientos y sugerir cambios en los diseños de las tuberías del sistema de recolección y tratamiento en las instalaciones de superficie [3].

En la actualidad existen formas mecánicas de hacerlo. Una manera es usando separadores de prueba, de los cuales hicimos mención en el capítulo 3, y otra forma es mediante medidores de fluido multifásico que pueden registrar en línea las tasas de flujo de agua, crudo y gas sin necesidad de realizar una separación física de cada fase.

Sin embargo, la precisión de los medidores de fluido multifásico en corrientes con alto GOR (relación gas-aceite) es baja, y si se compara la confiabilidad de sus mediciones frente a un separador de prueba, se evidencia un error de hasta 50% en cortes de agua (BSW) de 60% [7].

Frente a esta situación, el aprendizaje de máquina ha comenzado a desempeñar un papel importante en la industria petrolera y ya se ha comenzado a usar modelos de aprendizaje de máquina para determinar el volumen de petróleo extraído a partir de la inyección de agua [8], inyección de aceite [4], predecir la presión de la tubería o sarta de producción de un pozo [9], y pronosticar la producción de petróleo de todo un campo [10]. Distintos modelos han sido aplicados para estos fines, desde regresión lineal, máquinas de soporte vectorial, hasta redes neuronales y XGBoost. Estos modelos se limitan a 1 o 2 variables explicativas, pero no consideran las mediciones en tiempo real de los parámetros de operación de un equipo BES instalado en fondo de un pozo productor.

Ninguno de estos casos ha hecho una predicción en tiempo real del volumen de fluido que desplaza una bomba electrosomergible desde yacimiento hasta superficie.

Para el caso específico de bombas electrosomergibles, en Colombia esta comercialmente disponible SubPUMP® de IHS, el cual consiste en un software para análisis y diseño de bombas sumergibles y se basa en curvas de correlaciones teóricas de las variables operativas del equipo BES para estimar el volumen de fluido total capaz de movilizar. Para poder usar estas curvas de correlaciones se debe contar con mediciones de variables que no hacen parte del monitoreo diario de los equipos BES y el yacimiento. Incluso si contáramos con dichas mediciones sería necesario pagar por la licencia. Por ello tiene mucha relevancia realizar un modelo con las variables de seguimiento y monitoreo de los equipos BES para llevar a cabo estas estimaciones.

Vale la pena mencionar que el modelo de máquinas de soporte vectorial, seleccionado para este proyecto, es una técnica poderosa para resolver problemas de clasificación, regresión y estimación de funciones. La estrategia del modelo SVM es una técnica para un conjunto de métodos de aprendizaje supervisado que analiza datos y reconoce patrones y también se utiliza para análisis no lineal[8].

Capítulo 5

METODOLOGÍA

1. Selección de una serie de bomba BES para la cual se aplicará el modelo.

Según el fabricante, cada bomba BES presenta información como nombre de la bomba (DN4000, GN4000, etc.), caudal mínimo, caudal máximo, frecuencia en hz, eficiencia máxima de la bomba, potencia del eje y nombre de serie (300, 400, etc.).

En nuestro caso, se seleccionaron cinco equipos BES serie 400 con similar número de etapas.

2. Preparación del conjunto de datos.

El set de datos empleado en este trabajo es parte del registro que se lleva en un campo petrolero de Colombia. Debido a los temas de confidencialidad, los detalles de los datos se han limitado en este documento.

En la [tabla 1](#) se presenta una muestra del conjunto de datos usado, que incluye el registro de las variables operativas de los equipos BES serie 400 instalado en fondo, y variables propias del yacimiento. Todas las tablas y figuras están ubicadas en el apéndice.

La descripción detallada de los parámetros operativos de los equipos BES, usados para entrenar el modelo se muestran en la [tabla 2](#).

El siguiente paso en esta sección consistió en organizar el conjunto de datos, dado que los parámetros de operación de la bomba BES son registrados mediante sensores con frecuencias aleatorias que van desde 10 min hasta 30 min (ver primer panel de la [figura 1](#)). Estos parámetros se miden en tiempo real de operación del equipo. Adicionalmente se eliminaron los valores atípicos o “outliers” (ver segundo panel de la figura 1) y se realizó una interpolación de los datos, de manera que se pudiera contar con datos cada 1 hora, como se muestra en el tercer panel de la figura 1.

Una vez definida una frecuencia de 60 min para los 11 parámetros considerados, se fijó el primer 70% de los datos como el conjunto de datos de entrenamiento (22

observaciones) y el 30% restante se consideró como el conjunto de datos de evaluación (10 observaciones).

Teniendo en cuenta que las pruebas volumétricas oficiales - usando sets de pruebas - de los cinco (5) equipos BES fueron realizadas en promedio una (1) vez por mes para cada equipo, se definió trabajar con los parámetros registrados durante las 24 horas de la fecha en la que se realizó la prueba con la cual se determinó el volumen de fluido total real aportado por el equipo BES. De esta forma, se construyó un vector X de dimensión 264: 11 parámetros, 24 horas de registros (1 cada 60 min) con su respectivo rótulo Y que representa el volumen de fluido total real (variable objetivo) correspondiente al día de las 24 horas de registro para cada equipo BES.

Luego del preproceso planteado, y tomando en cuenta las 22 mediciones de volumen de fluido total real – usando sets de pruebas – que hacen parte del conjunto de datos de entrenamiento (ver [tabla 3](#)), se construyó una matriz de 22 filas (1 fila por cada medición u observación) x 265 columnas para el conjunto de datos de entrenamiento.

3. Caracterización del fluido desplazado por los equipos BES.

Las principales características del fluido de los cinco pozos productores, cada uno con su equipo BES a evaluar, se muestran en la [tabla 7](#). Aunque estos parámetros no fueron tenidos en cuenta para el entrenamiento del modelo, si se consideró oportuno conocer, al menos de manera general, las fracciones de agua, crudo y gas de cada uno de los fluidos movilizados por los equipos BES en cuestión, y que pueden ser tenidos en cuenta para futuros trabajos.

4. Entrenamiento del modelo.

Se usó un modelo de máquinas de soporte vectorial para regresión y se realizó validación cruzada Leave-One-Out.

Leave-One-Out proporciona índices de entrenamiento/prueba para dividir datos en conjuntos de entrenamiento/prueba. Cada muestra se usa una vez como conjunto de prueba, mientras que las muestras restantes forman el conjunto de entrenamiento.

Seguidamente se realizó una búsqueda de los hiperparámetros que mejoraban el desempeño del modelo mediante GridSearchCV.

Adicionalmente, se realizó una regresión lineal usando el conjunto de datos de entrenamiento y se compararon las métricas de error absoluto medio (MAE) y la raíz del error cuadrático medio (RMSE) con respecto al modelo de máquinas de soporte vectorial entrenado.

5. Evaluación del modelo.

El desempeño del modelo fue monitoreado mediante el error absoluto medio (MAE) y la raíz de la media del cuadrado de los errores (RMSE), y comparado usando estas métricas frente a un modelo de regresión lineal.

Adicionalmente, se revisó el grado de desviación que tiene el modelo con respecto a la identidad, mediante un gráfico cruzado de datos de prueba y entrenamiento, la revisión de los intervalos de confianza, la distribución normal del error y el coeficiente de correlación de Pearson.

Capítulo 6

RESULTADOS Y DISCUSIÓN

El modelo de máquinas de soporte vectorial fue construido con el conjunto de datos de entrenamiento y luego usado para predecir el fluido total de producción de los cinco pozos productores de crudo (cinco equipos BES). En la [figura 2](#) se puede evidenciar el resultado, en donde se realiza un gráfico cruzado entre los volúmenes reales de fluido producido (previamente definido como el conjunto de datos de evaluación de acuerdo con la [tabla 4](#)) y los volúmenes estimados por el modelo y existe una buena correlación entre ambos valores, que se denota por un R^2 de 0.93. De manera similar, para el conjunto de datos de prueba se obtuvo un R^2 de 0.92 de acuerdo con la [figura 3](#). Adicionalmente, se determinó un coeficiente de correlación de Pearson de 0.96 para los datos de entrenamiento y 0.97 para los datos de evaluación, y en ambos casos un p-valor que tiende a cero (ver [tabla 6](#)).

Por otra parte, el error absoluto medio (MAE) fue de 181.55 y la raíz de la media del cuadrado de los errores (RMSE) de 13.47. Ambas métricas para el modelo de máquinas de soporte vectorial superaron las registradas por el modelo de regresión lineal, en el cual el MAE fue de 291.16 y la RMSE de 17 (ver [tabla 5](#)).

También se logró para el modelo un intervalo de confianza de 95% entre los valores 1378 y 1536, tal como se muestra en la [figura 4](#).

Así mismo, la [figura 5](#) permite apreciar la distribución normal del error y denota un buen desempeño del modelo en la predicción del volumen de fluido total desplazada por un equipo BES.

Capítulo 7

CONCLUSIONES Y RECOMENDACIONES

Inicialmente se contempló, como alcance del presente proyecto, realizar la predicción de la tasa de aceite en el fluido, lo cual corresponde al producto que se comercializa y representa los ingresos económicos en la industria petrolera. Pero llevar a cabo esto requiere contar no sólo con los registros de los parámetros del equipo BES, sino también con las características fisicoquímicas del aceite en cada uno de los cinco pozos productores y no hubo disponibilidad de dicha información, por lo que consideramos interesante poder registrar densidades del fluido para futuros trabajos.

En este trabajo, abordamos desde un enfoque basado en el aprendizaje automático de máquina, poder predecir la tasa de producción de fluido total (agua y crudo) para datos reales de un campo petrolero colombiano donde se usan bombas electrosumergibles como sistema de levantamiento artificial.

Debido a la poca cantidad de observaciones, lo cual redujo el tamaño de nuestro conjunto de datos, asociado principalmente a la falla temprana de los equipos BES y baja frecuencia de las pruebas de los pozos productores involucrados en este proyecto, la escogencia del modelo se limitó a usar máquinas de soporte vectorial, que corresponde a un modelo sencillo, con pocos parámetros para sintonizar.

Sin embargo, el modelo de máquinas de soporte vectorial entrenado fue capaz de predecir el volumen de fluido total desplazado por un equipo BES con un R^2 de 0.93 tal como lo permite apreciar la [figura 2](#). Además, el modelo permite usar los registros de los equipos BES para pronosticar en tiempo real la producción de los pozos, lo cual representa una enorme ventaja para el grupo de control de producción del campo, con lo cual pueden identificar el pozo o equipo BES con pérdida de eficiencia y, por ende, pérdida de producción. De manera que ahora se puede ofrecer una respuesta mucho más oportuna en los casos donde haya una caída de producción, evitando incurrir en la búsqueda exhaustiva

que se menciona en la definición del problema y logrando vender los barriles que de otra manera no se hubieran identificado a tiempo.

Las predicciones con mayor desviación respecto a los datos reales se presentaron en los dos fluidos con un mayor GOR (relación gas-aceite), lo que permite concluir que el modelo se desempeña mejor cuando hay poco gas en el fluido, y en cambio, un mayor volumen de líquido. Para futuros trabajos valdría la pena involucrar este parámetro en el entrenamiento de los modelos (ver [tabla 7](#)), así como evaluar la variación de la corriente de motor, VSD y THP del pozo productor, que corresponden a parámetros del equipo BES que se ven afectados significativamente por el volumen de gas en el fluido.

REFERENCIAS

- [1] A. Géron, «Chapter 5. Support Vector Machines», en *Hands-On Machine Learning with Scikit-Learn and TensorFlow*, O'Reilly Media, Inc., 2017.
- [2] A. Sheikhoushaghi, N. Y. Gharaei, y A. Nikoofard, «Application of Rough Neural Network to forecast oil production rate of an oil field in a comparative study», *J. Pet. Sci. Eng.*, vol. 209, p. 109935, feb. 2022, doi: 10.1016/j.petrol.2021.109935.
- [3] A. F. Ibrahim, R. Al-Dhaif, S. Elkatatny, y D. A. Shehri, «Applications of Artificial Intelligence to Predict Oil Rate for High Gas–Oil Ratio and Water-Cut Wells», *ACS Omega*, vol. 6, n.º 30, pp. 19484-19493, ago. 2021, doi: 10.1021/acsomega.1c01676.
- [4] N. A. Sami, «Application of machine learning algorithms to predict tubing pressure in intermittent gas lift wells», *Pet. Res.*, vol. 7, n.º 2, pp. 246-252, jun. 2022, doi: 10.1016/j.ptlrs.2021.09.006.
- [5] M. Crnogorac, M. Tanasijević, D. Danilović, V. Karović Maričić, y B. Leković, «Selection of Artificial Lift Methods: A Brief Review and New Model Based on Fuzzy Logic», *Energies*, vol. 13, n.º 7, p. 1758, abr. 2020, doi: 10.3390/en13071758.
- [6] J. D. WOODS y J. F. LEA, *What's new in artificial lift?*, vol. 238, n.º 6. 2017. Accedido: 21 de noviembre de 2022. [En línea]. Disponible en: <https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=128780082&site=eds-live>
- [7] T. Ganat, «Reliability Analysis of Multiphase Flow Measurements in Different Oil Fields», *OIL GAS-Eur. Mag.*, 2016.
- [8] A. Kamari, A. Bahadori, y A. H. Mohammadi, «An Efficient Approach for the Determination of Oil Production Rate During the Water-flooding Recovery Method», *Pet. Sci. Technol.*, vol. 33, n.º 11, pp. 1208-1214, jun. 2015, doi: 10.1080/10916466.2015.1057595.
- [9] Y. Ning, H. Kazemi, y P. Tahmasebi, «A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet», *Comput. Geosci.*, vol. 164, p. 105126, jul. 2022, doi: 10.1016/j.cageo.2022.105126.
- [10] R. Simanjuntak y D. Irawan, «Applying Artificial Neural Network and XGBoost to Improve Data Analytics in Oil and Gas Industry», *Indones. J. Energy*, vol. 4, n.º 1, pp. 26-35, feb. 2021, doi: 10.33116/ije.v4i1.103.

APÉNDICE

Tabla 1. Muestra del conjunto de datos de uno de los equipos BES, usado en el modelo.

<i>Fecha</i>	<i>Presión Casing (psig)</i>	<i>Temperatura Intake (F)</i>	<i>Amperaje Motor (A)</i>	<i>Frecuencia del motor (Hz)</i>	<i>Temperatura del motor (F)</i>	<i>Voltaje (V)</i>	<i>Presión de descarga de la bomba (psig)</i>	<i>Presión Intake de la bomba (psig)</i>	<i>Presión tubing en superficie (psig)</i>	<i>Vibraciones (g)</i>	<i>Amperaje del VSD (A)</i>
10/05/22 23:23	NaN	NaN	NaN	289.9	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10/05/22 23:32	NaN	43.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10/05/22 23:39	NaN	NaN	NaN	NaN	NaN	NaN	453.3	NaN	NaN	NaN	NaN
10/05/22 23:43	NaN	NaN	NaN	NaN	NaN	3750.9	NaN	NaN	NaN	NaN	NaN
10/05/22 23:43	NaN	NaN	NaN	NaN	NaN	NaN	NaN	82.95	NaN	NaN	NaN
...
10/06/22 22:45	NaN	NaN	NaN	289.3	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10/06/22 23:13	NaN	NaN	NaN	NaN	NaN	3770.0	446.2	NaN	NaN	NaN	NaN
10/06/22 23:14	NaN	NaN	NaN	NaN	NaN	NaN	NaN	99.40	NaN	NaN	NaN
10/06/22 23:23	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.44	NaN	NaN
10/06/22 23:34	NaN	NaN	NaN	NaN	NaN	NaN	NaN	67.02	NaN	NaN	NaN

Tabla 2. Descripción de los parámetros operativos y de yacimiento usados en el modelo, los valores corresponden al instante de medición.

<i>Parámetro</i>	<i>Descripción</i>
Fecha	Día/mes/año/hora/segundo de la medición
Presión del casing (psig)	Presión dentro del revestimiento o entre el revestimiento y la tubería en la cabeza del pozo
Temperatura intake (F)	Temperatura registrada en la succión del equipo BES
Amperaje del motor (A)	Consumo de amperaje del motor del equipo BES
Frecuencia del motor (Hz)	Frecuencia o velocidad del motor
Temperatura del motor (F)	Temperatura registrada en el motor
Voltaje (V)	Voltaje del motor del equipo BES
Presión descarga de la bomba (psig)	Presión con la que la bomba BES descarga el fluido hacia superficie.
Presión intake de la bomba (psig)	Presión del yacimiento en el punto de succión de la bomba.
Presión tubing en superficie (psig)	Presión registrada en superficie o cabeza de pozo.
Vibraciones (g)	Vibraciones registradas en la bomba BES. Pueden presentarse por atascamientos en la etapas o desbalance del equipo.
Amperaje del VSD (A)	Consumo de amperaje del variador de frecuencia de la bomba.

Tabla 3. Volúmenes de fluido total liquidado usados para el entrenamiento del modelo

<i>Pozo productor</i>	<i>Fecha de la prueba</i>	<i>Volumen de fluido total liquidado (Bbls)</i>
W1	21/01/22	938
	28/01/22	1656
	14/02/22	1661
	4/03/22	1641
	8/04/22	1631
	8/06/22	1641
	7/07/22	1774
W2	7/06/22	2519
	5/08/22	2538
W3	2/06/22	587
	31/07/22	456
W4	9/12/21	1405
	20/01/22	1397
	28/02/22	1452
	14/03/22	1440
	9/04/22	1437
	15/06/22	1421
	9/07/22	1325
W5	11/03/22	1381
	18/04/22	1250
	30/06/22	1705
	04/08/22	1793

Tabla 4. Volúmenes de fluido total liquidado usados para la evaluación del modelo

<i>Pozo productor</i>	<i>Fecha de la prueba, usando set de pruebas</i>	<i>Volumen de fluido total liquidado (Bbls)</i>
W1	17/08/22	1604
	1/09/22	1605
	2/10/22	1588
W2	6/09/22	2509
	4/10/22	2527
W3	27/08/22	315
	11/10/22	501
W4	9/08/22	1540
	6/10/22	1517
W5	30/09/22	1752

Tabla 5. Métricas de los dos modelos entrenados

<i>Modelo</i>	<i>MAE</i>	<i>RMSE</i>
Regresión lineal	291.16	17.06
SMV	181.55	13.47

Tabla 6. Correlación de Pearson para el modelo SVM

<i>Conjunto datos</i>	<i>Coefficiente correlación Pearson</i>	<i>p - valor</i>
Entrenamiento	0.964	5,39E-13
Evaluación	0.977	1,02E-06

Tabla 7. Características del fluido movilizado por las bombas BES Serie 400

<i>Pozo productor</i>	<i>GOR</i>	<i>%BSW (fracción de agua)</i>
W1	0.455	98
W2	0.221	93
W3	0.522	82
W4	0.061	83
W5	1.247	90

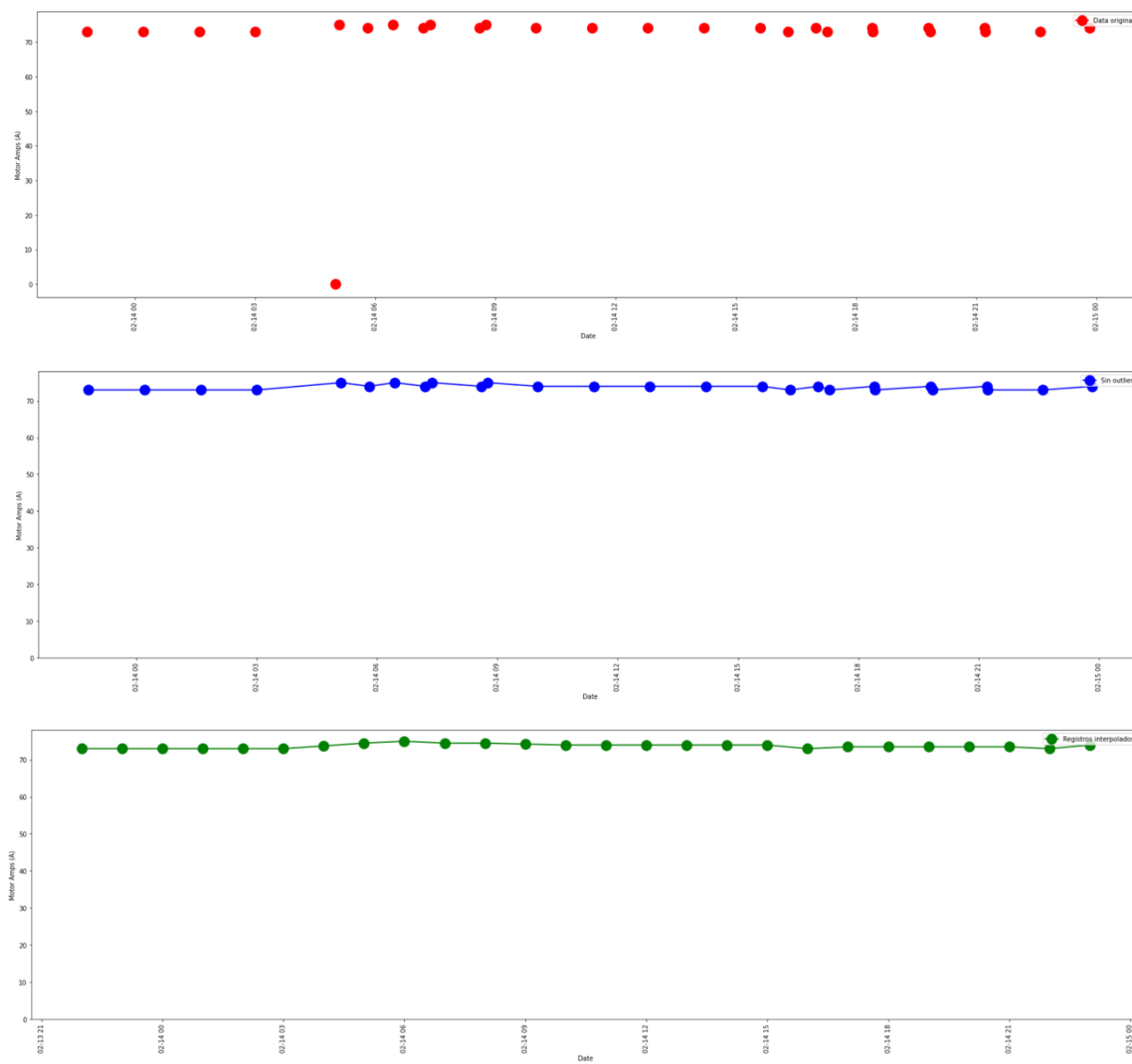


Figura 1. Ilustración del preprocesamiento del conjunto de datos de entrenamiento.

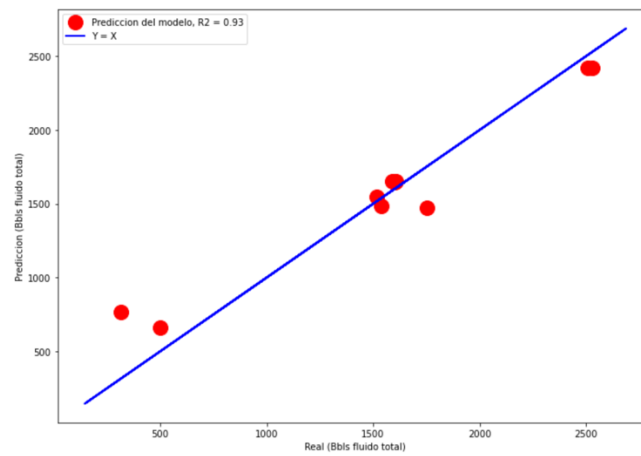


Figura 2. Gráfico cruzado usando el conjunto de datos de evaluación ($R^2=0.93$).

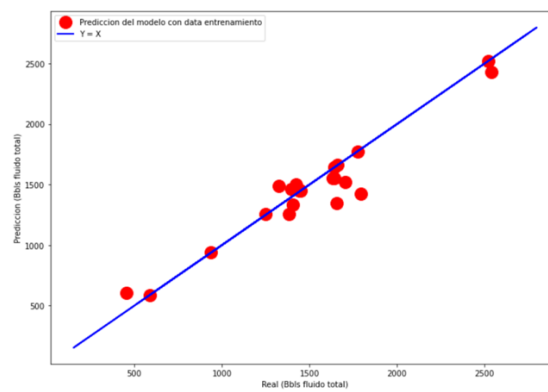
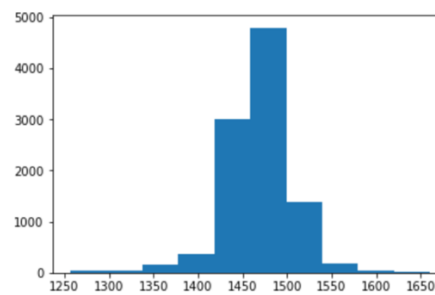


Figura 3. Gráfico cruzado usando el conjunto de datos de entrenamiento ($R^2=0.92$).



95.0 confidence interval 1378.3913107467367 and 1535.83

Figura 4. Intervalos de confianza del modelo SVM.

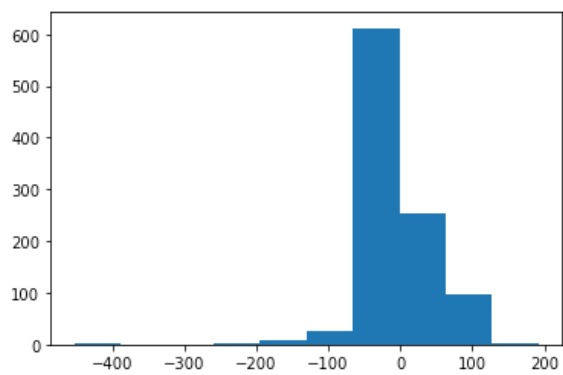


Figura 5. Distribución normal del error para el conjunto de datos de entrenamiento.