

DIFUSIÓN CIENTÍFICA

Ciencias Sociales y Humanas,
Ingeniería, Ciencia, Tecnología
e Innovación



Nthe

Publicación del Consejo de Ciencia y Tecnología del Estado de Querétaro

Directorio

PRESIDENTE

M. V.Z. FRANCISCO DOMÍNGUEZ SERVIÉN

VICEPRESIDENTE

M. EN C. JOSÉ CARLOS ARREDONDO VELÁZQUEZ

DIRECTOR GENERAL

M. EN A. RAÚL ITURRALDE OLVERA

SECRETARIO

LIC. MAURICIO PALOMINO HERNÁNDEZ

DERECHOS DE AUTOR Y DERECHOS CONEXOS. Año 12 num. 35, enero - abril de 2021.

Nthe es una publicación cuatrimestral editada por el Consejo de Ciencia y Tecnología del Estado de Querétaro (CONCYTEQ): calle Luis Pasteur Sur núm. 36, col. Centro; CP 76000; tel. (442) 214 3685; www.concyteq.edu.mx; nthe@concyteq.edu.mx. Editor responsable: Felipe de Jesús Esperón Valenzuela. Reserva de derechos al uso exclusivo núm. 04-2018-111410321700-203; ISSN 2007-9079, ambos otorgados por el Instituto Nacional del Derecho de Autor. Responsable de la última actualización de este número: Felipe de Jesús Esperón Valenzuela: calle Luis Pasteur Sur núm. 36, col. Centro; CP 76000. Fecha de última modificación: abril de 2021

Nthe ha sido aprobada para su inclusión en el Índice del Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal (LATINDEX)

Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor de la publicación.

Se autoriza la reproducción total o parcial de los contenidos e imágenes de la publicación siempre y cuando se cite la fuente.

DIRECTOR DE LA REVISTA ELECTRÓNICA NTHE

DR. MARCO ANTONIO CARRILLO PACHECO

EDICIÓN Y DISEÑO DE LA PUBLICACIÓN

LIC. FELIPE DE JESÚS ESPERÓN VALENZUELA

CORRECCIÓN DE ESTILO

DRA. MARÍA LUISA ÁLVAREZ MEDINA
LIC. MONSERRAT ACUÑA DELGADO

Nthe, Publicación del Consejo de Ciencia y Tecnología del Estado de Querétaro:
<http://nthe.mx/>

Luis Pasteur Sur núm. 36
Col. Centro, CP 76000
Tel. 52 (442) 214 3685 / 212 7266, ext. 105
Querétaro, Qro., México

Consejo editorial

Investigadores nacionales

Dr. Alejandro Manzano Ramírez

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, unidad Querétaro

Dr. Flora Mercader Trejo

Universidad Politécnica de Santa Rosa Jáuregui

Dr. Sergio Barrera Sánchez

Instituto Tecnológico de Estudios Superiores de Monterrey, campus Querétaro

Dr. Martha Cruz Soto

Universidad del Valle de México, campus Querétaro

Dr. Gabriela Calderón Guerrero

Facultad de Psicología, Universidad Autónoma de Querétaro

Dr. Víctor Castaño Meneses

Centro de Física Aplicada y Tecnología Avanzada de la UNAM, campus Juriquilla

Dr. Rolando Salinas García

Unidad Multidisciplinaria de Estudios Sobre el Trabajo, Universidad Autónoma de Querétaro

Dr. Miguel Martínez Madrid

Instituto Mexicano del Transporte, SCT

Dr. Daniel Hiernaux Nicolás

Facultad de Ciencias Políticas y Sociales, Universidad Autónoma de Querétaro

Dr. Gabriel Corral Velázquez

Facultad de Ciencias Políticas y Sociales, Universidad Autónoma de Querétaro

Dr. Salvador Echeverría Villagómez

Centro Nacional de Metrología

Dr. Alberto Traslosheros Michel

Universidad Aeronáutica en Querétaro

Dra. Alejandra Urbiola Solís

Facultad de Contaduría y Administración, Universidad Autónoma de Querétaro

Dr. Jöns Sánchez Aguilar

Instituto Tecnológico de Querétaro, Tecnológico Nacional de México

Dr. Octavio López Millán

Instituto Tecnológico de Hermosillo

Dra. Marcela Hernández Romo

Universidad Autónoma Metropolitana, unidad Iztapalapa

Mtra. Martha Díaz Muro

Instituto Tecnológico de Hermosillo, Tecnológico Nacional de México

Dr. Sergio Sandoval Godoy

Centro de Investigación en Alimentación y Desarrollo, Sonora

Dr. Martín Alfonso Gutiérrez López

Universidad Autónoma de Querétaro

Dr. Jorge Raúl Palacios Delgado

Universidad del Valle de México, campus Querétaro

Dra. María Guadalupe Ordaz Cervantes

Universidad Autónoma de Querétaro

Mtra. Eva Leticia Ortiz Ávalos

Universidad Nacional Autónoma de México

Dra. Adriana del Carmen Gallegos Melgar

CONACYT - COMIMSA

Dra. Ana Isabel Moreno Calles

Universidad Nacional Autónoma de México, Escuela Nacional de Estudios Superiores, Morelia.

Dra. Olivia Solís Hernández

Universidad Autónoma de Querétaro

Externo

Dr. Víctor Hugo Muciño Quintero

Universidad de West Virginia, EUA



La revista electrónica *Nthe* es financiada en su totalidad por el Consejo de Ciencia y Tecnología del Estado de Querétaro en el marco del presupuesto autorizado el día 21 de diciembre de 2020. (Publicado en el periódico oficial del estado de Querétaro, La Sombra de Arteaga).



SECRETARÍA DE EDUCACIÓN



Análisis predictivo del comportamiento de un actuador magnético

Armando Erik Rivera Paz¹, John Alexander Villada Villalobos², Luis Alberto Cáceres Díaz^{3*}

¹POSGRADO Centro de Tecnología Avanzada CIATEQ A.C. unidad San Luis Potosí, San Luis Potosí, México.

²Cátedras CONACYT - Centro de Ingeniería y Desarrollo Industrial (CIDESI), Querétaro, México.

³CONACYT - CIATEQ A.C. Centro de Tecnología Avanzada, unidad San Luis Potosí, San Luis Potosí, México.

¹aerp36@gmail.com, ²luis.caceres@ciateq.mx

Resumen

Los actuadores son dispositivos con la capacidad de recibir y convertir energía que realiza un efecto de interacción de este con un medio externo. Durante la fabricación de un actuador, se genera mucha información que permite optimizar los procesos y calidad del dispositivo; sin embargo, el nivel de complejidad y de correlación de los datos dificulta la toma de decisiones. El ajuste de los procesos se basa en el conocimiento del efecto-resultado del proceso mismo. Las nuevas tecnologías permiten generar y almacenar información que describe el comportamiento del proceso. En este trabajo se implementó una metodología de análisis estadístico, aplicando regresiones lineales simples y múltiples para evaluar el efecto de variables de ensamble sobre características críticas de fabricación de los actuadores magnéticos. Esto permitió generar un modelo con la capacidad de predecir en un 87.5 % la relación entre los valores medidos durante el proceso de ensamble y el resultado de una característica de funcionamiento del actuador.

Palabras Clave: actuador magnético, regresión lineal múltiple, predicción

Abstract

Actuators are devices with the ability to receive and convert energy that interacts with an external environment. During the manufacture of an actuator, a lot of information is generated that allows optimizing the processes and quality of the device; however, the level of complexity and correlation of the data makes it difficult to make decisions. The adjustment of the processes is based on the knowledge of the effect-result of the process itself. New technologies make it possible to generate and store information that describes the behavior of the process. In this work a statistical analysis methodology was implemented applying simple and multiple linear regressions to evaluate the effect of assembly variables on critical manufacturing characteristics of magnetic actuators. This allowed generating a model able to predict in 87.5 % the relationship between the values measured during the assembly process and the result of an operating characteristic of the actuator.

Keywords: magnetic actuator, multiple linear regression, forecasting

Artículo arbitrado

Recibido:

15 de marzo de 2021

Aceptado:

20 de abril de 2021

Introducción

Los actuadores son dispositivos que tienen la capacidad de recibir y convertir energía, la cual realiza un efecto de interacción con un medio externo. Los actuadores electromagnéticos son usados en componentes de seguridad, sistemas de inyección y comprensión para vehículos automotores tipo SUV, camioneta y pick-up (Kocabicak et al., 2020; Wang et al., 2020). El principio de funcionamiento de un actuador electromagnético solenoide (a partir de aquí llamado solamente actuador magnético) se basa en la ecuación de Lorenz, la cual describe la relación de la magnitud de la fuerza generada en función de su campo magnético (Gutiérrez et al., 2017; Fundación Endesa., 2020). El ensamble de un actuador magnético se realiza en diferentes pasos, resumidos a continuación:

Subensamble A: enrollado del hilo conductor alrededor del núcleo magnético.

Subensamble B: ensamble de hilo-núcleo dentro de carcasa; ensamble imán varilla dentro de carcasa; ensamble de actuador magnético completo. Al final de cada proceso de ensamble se realiza una evaluación que determina si el ensamble cumple con las tolerancias requeridas. Los resultados de cada evaluación (aprobada y no aprobada) son registrados en una base de datos que se almacena en tiempo real. Los datos medidos de cada proceso de ensamble y de cada variable contiene información relevante para mejorar el proceso y para identificar las fallas que pueda tener el actuador magnético producido (Ismail et al., 2019). El problema radica en que la cantidad de datos es extensa y la relación de codependencia que existe entre las variables aumenta la complejidad de los análisis.

Para garantizar un proceso estable se requiere de una metodología que permita la toma efectiva de decisiones, la cual pueda estar basada en el análisis de la información obtenida durante la fabricación de las piezas.

Uso de ciencia de datos para análisis de proceso.

Como se mencionó, los diferentes procesos de ensamble y las evaluaciones posteriores generan bases de datos diarias que pueden generar dos tipos de variables: 1) variables por atributos que clasifican la respuesta de acuerdo con un grupo establecido, y 2) variables continuas que pueden obtener cualquier valor dentro de un dominio definido (valores enteros, decimales, fracciones, etc.). El uso de los datos para generar indicadores o también llamados *Insights* permite la toma de decisiones. Por su parte, está probado que el uso de técnicas de procesamiento de datos es efectivo para resolver problemas de aprendizaje de máquina (Li et al., 2017).

Exploración de datos y preprocesamiento.

El proceso de exploración de los datos es importante para garantizar que el algoritmo utilizado sea implementado de forma correcta. Dentro de este se logra visualizar la forma en que se estructura la información y entender los tipos de datos (o también llamados “observaciones”) obtenidos del proceso de fabricación. Una vez visualizada y entendida la base de datos, se procede a realizar un preprocesamiento de la información, en el cual se eliminan todas las observaciones que contengan valores no representativos del proceso. Estos valores se pueden filtrar con el uso de técnicas estadísticas o algoritmos automatizados (o bien con ambas); ya sea detectando valores extremadamente fuera de los rangos convencionales (valores normalmente fuera del límite de capacidad del equipo de medición) u observaciones con valores indefinidos (se identifican como “NaN”) y, por último, valores iguales a 0. Estos valores se pueden generar por diferentes causas (periodos de calibración, fabricación de prototipo, periodos de mantenimiento, etc.) y no deben ser considerados dentro del análisis.

En este trabajo se implementó una metodología de análisis estadístico y se utilizó el método de regresiones lineales múltiples para evaluar el efecto de variables de ensamble sobre características críticas de funcionamiento de los actuadores magnéticos.

Método

Características de funcionamiento de un actuador.

La última evaluación que se realiza al actuador magnético antes de definirse como producto confirma que este cumpla con las características de funcionamiento requeridas en campo. Estas evaluaciones son comparadas de acuerdo con el rango de tolerancia definido previamente.

En la Figura 1, se muestra el diagrama general de un actuador magnético lineal solenoide, en donde se aprecian los componentes que representan las variables de ensamble y características críticas. De las primeras, se consideraron en este trabajo la profundidad de bobina, altura del ancla, resistencia eléctrica y temperatura. De las segundas, se contemplaron el valor de altura de pin extraído y retraído, promedio tiempo muerto (tiempo de reacción), tiempo de cambio y fuerza de retracción. En la misma figura, se representa el embobinado del actuador para generar el campo magnético de forma de líneas diagonales.

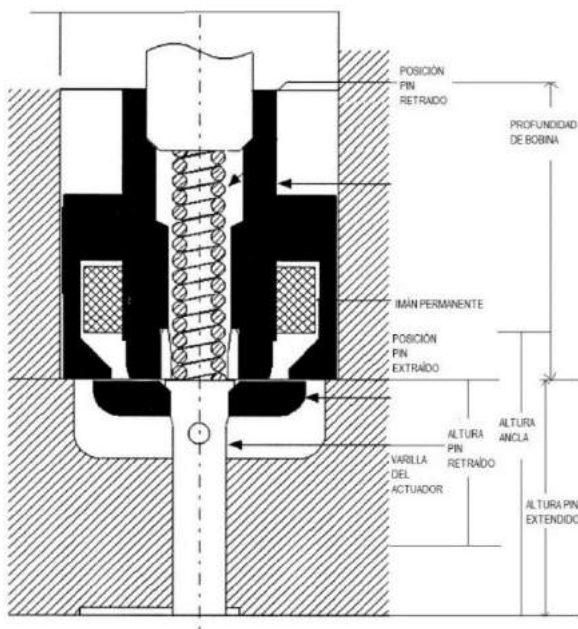


Figura 1. Diagrama general de un actuador magnético lineal solenoide. Fuente: imagen adaptada de Howe (2000).

La posición de pin extraído y retraído se genera cuando el magneto cambia de posición. La altura

del ancla se define desde la cara superior del magneto hasta la posición la cara superior de la varilla del actuador. Dichos parámetros se optimizan durante el proceso de fabricación

Recolección y tratamiento de datos. Los datos fueron recolectados utilizando sistemas de medición tales como transductores mecánicos, sistemas de visión y sensores fotoeléctricos. La base de datos que se recolectó durante un periodo de 6 meses para la variable de interés altura de la varilla en posición Extraída (pin), contaba con una dimensión de 142 columnas y alrededor de 80 mil datos. Sin embargo, se observó que 133 columnas mostraban información no relevante para el modelo (Variables categóricas) y solo 9 aportaban información para generar el modelo. Como resultado, se creó una base de datos que incluyera solo las variables de interés. Se realizó un pre-procesamiento de los datos, para descartar valores atípicos o datos faltantes de la base de datos.

Análisis de correlación de los datos. Se empleó el método de correlación para poder explicar la relación lineal que presentan los datos. Existen diferentes indicadores que permiten explicar la relación entre las variables, entre los cuales se puede mencionar la Correlación de *Pearson* y la de *Spearman* y de *Jackson* (Hernández et al., 2018). Para este estudio se utilizó el método de correlación de *Pearson*, en el cual se explica el nivel de relación lineal en cada análisis para determinar el grado de correlación entre las variables predictoras respecto de una variable de interés. El cual se define mediante la fórmula (Cheng et al., 2017):

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - x_{prom}) * (y_i - y_{prom})}{\sqrt{\sum_{i=1}^n (x_i - x_{prom})^2 * \sum_{i=1}^n (y_i - y_{prom})^2}}$$

Modelo de regresión lineal. Para poder plantear este modelo de regresión se analizaron los valores obtenidos en la correlación de *Pearson*. La evaluación del modelo de regresión lineal se realizó a través de los residuos, los cuales son el resultado de la diferencia entre el valor real de la observación y el valor de predicción el modelo. El análisis de los residuos no solo permite visualizar la variación entre predicción y valor real, sino que

también aporta información para confirmar que se cumplen las condiciones necesarias para generar un modelo de regresión lineal (Amat, 2016).

Modelo de regresión múltiple. En la literatura se define como el modelo que ajusta modelos lineales entre variables dependientes y variables independientes. Este se representa mediante la siguiente fórmula (García et al., 2018):

$$y_i = \beta_0 + \beta_1 * x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon_i$$

Donde y_i es la variable dependiente (en nuestro caso la variable de interés altura de pin), β_0 el valor constante ordenado al origen y $\beta_{1,2,\dots,i}$ representan los valores de los coeficientes de regresión parcial, que se determinan a partir de las variables predictoras ($x_{1,2,i}$) (Montero R, 2016).

Para evaluar la importancia de cada variable, se evaluó el valor de ajuste de bondad para modelo simple de acuerdo con la siguiente fórmula:

$$R^2 = 1 - \frac{\sum(y_{i_predicho} - y_i)^2}{\sum(y_i - y_{promedio})^2}$$

Este parámetro indica qué porcentaje de datos pueden ser explicados por el modelo. De igual forma, se puede determinar el grado de importancia de cada uno de los predictores en el modelo. Para los lineales múltiples, se utilizó la siguiente fórmula:

$$R^2_{ajustado} = R^2 - (1 - R^2) \frac{n - 1}{n - p - 1}$$

Resultados y discusión

En la Figura 2, se presenta la matriz de coeficientes de correlación de *Pearson* (R) entre las variables de ensamble y las características críticas. El valor de correlación entre cada variable se presenta también como un mapa de calor que permite identificar las contribuciones más importantes entre las medidas (Usharani et al., 2020).



Figura 2. Matriz y mapa de calor de coeficientes de correlación de *Pearson* R. Fuente: elaboración propia.

Se identifica en la barra de colores que desde $R=-1$, en color rojo, representa una correlación lineal negativa, lo cual significa que la relación que existe entre las variables es 1 a 1 o perfecta, pero en este caso es de forma inversa; es decir, cuando una variable aumenta la otra disminuye. Los valores de $R=0$ están representados en color blanco en el mapa de calor. Esto significa que no hay alguna dependencia entre las variables de interés o correlacionadas. Por su parte, los valores de $R=1$, en color azul, muestran una correlación lineal positiva perfecta, esto quiere decir que cuando una variable aumenta en una unidad de magnitud la de correlacionada aumenta en la misma cantidad.

La diagonal que se observa con valores de $R=1$ es de esperarse porque se está analizando la correlación de la variable con ella misma. En este trabajo, se estudió el valor de la altura de pin en configuración extendida o también llamado valor pin fuera. Se observó el efecto de la altura del ancla y de la contribución de cada variable sobre el comportamiento del pin. El efecto de correlación sobre la altura de pin se analiza en la línea horizontal del mapa de calor, en donde se identificó que la correlación con resistencia eléctrica (RE), temperatura, valor pin dentro

(VPD), promedio tiempo muerto (ptR), tiempo de cambio (tC), fuerza de retracción (FR), altura de ancla (AA) y profundidad de bobina (PB) son 0.28, 0.074, 0.36, 0.65, 0.67, -0.044, 0.87 y -0.28, respectivamente.

La Tabla 1 muestra el reporte del primer modelo de regresión lineal implementado para la variable de interés altura de pin (y) en función de la variable de ensamble altura de ancla. Este modelo es una regresión lineal implementada por el método de mínimos cuadrados ordinario (OLS) para una sola correlación (Df). El total de la población de datos empleado fue de 89431.

Tabla 1. Valores del primer modelo de regresión lineal.

Reporte de primer modelo de regresión				
Variable dep:	y	R-cuadrada:	0.759	
Modelo:	OLS	Adj. R-cuadrada:	0.759	
Método:	min cuadrados	F-estadística:	281600	
N° Observ	89431	Prob (F-statistic):	0	
Df modelo:	1	rmse	0.02340	
Inter conf	[0.025 0.975]			
	coef	std err	t	P> t
Const	-5.1696	0.073	-71.254	0
Ancla	1.0602	0.002	530.614	0

Fuente: elaboración propia.

Se trabajó con un intervalo de confianza estándar entre 2.5% y 97.5% de los datos. Se obtuvo una estimación de bondad del ajuste (R^2) de 0.759.

El valor del error de ajuste (rmse) fue de 0.0234, esto define cuanto se está desviando el valor del promedio de predicción respecto a los valores reales. El coeficiente parcial de regresión para el ancla fue de 1.0602. Este valor se emplea en la implementación del segundo modelo en el que incluye la contribución de todas las variables al comportamiento del pin del actuador magnético.

Después de identificar la correlación real entre la variable altura de pin con el ancla se implementó el método de regresión múltiple, en donde se agregó cada variable en función de las iteraciones. Esto se llevó a cabo considerando el método *step-forward* que implica evaluar cada iteración agregando los predictores de manera secuencial y revisando su efecto sobre la altura de pin, para saber si cada una de ellas tenía o no alguna contribución (Zhang, 2016).

En la Tabla 2 se presenta el resultado de los coeficientes parciales de la regresión lineal múltiple, obtenidos por el método de *step-forward*. Para este análisis se tuvo en cuenta la contribución de las variables de ensamble desde el coeficiente de correlación de *Pearson* mayor al menor.

Debido a la gran cantidad de datos generados en las líneas de producción y dado que la información contenía mucha dispersión, se tomó la decisión de establecer cuál era la correlación que existía entre las variables de evaluación para poder definir un primer método de regresión y esto se consideró teniendo en cuenta el valor de R.

Tabla 2. Coeficientes parciales de la regresión lineal múltiple.

Resultado	PB	ptR	tC	AA	VPD	RE	FR
Cambio	R2	Iteración	R	Coeficientes parciales de regresión lineal múltiple			
	0.76	1	0.87	1.0602			
0.010	0.77	2	0.67	0.0196 0.9477			
0.000	0.77	3	0.65	-0.0065 0.0254 0.9470			
0.079	0.85	4	0.36	0.0271 0.8684 0.5116			
0.003	0.85	5	0.28	0.0339 0.8485 0.5397 -0.0460			
0.023	0.87	6	-0.28	-0.3745 0.0320 0.8639 0.4044 -0.0449			
0.000	0.87	7	-0.04	-0.3739 0.0323 0.8611 0.4063 -0.0449 -0.0004			

Fuente: elaboración propia.

En la tabla se presenta el valor de cambio, el coeficiente de correlación R2, la iteración correspondiente, el coeficiente de correlación de *Pearson* en orden descendente y los coeficientes parciales de correlación. Las interacciones se establecieron en orden, según el peso que tiene cada variable al modelo de regresión múltiple. Esa decisión se tomó con base en la matriz de correlaciones de *Pearson* (ver Figura 1). Se muestran los resultados de predicción de altura de pin empleando el modelo de regresión lineal múltiple, en el cual se tienen en cuenta las contribuciones no solo la altura de ancla sino también la profundidad de bobina, el promedio de tiempo de reacción, el tiempo de cambio, valor de pin dentro, resistencia eléctrica y fuerza de retracción.

En la Figura 3 se muestra la predicción de los datos considerando los modelos de regresión lineal sencillo y múltiple. En la Figura 3a) se observó como el modelo de regresión simple ajustado a los valores medidos muestran una dispersión considerable respecto de la predicción y solamente

tiene en cuenta el efecto del valor de ancla sobre la altura de pin extraído. La confiabilidad de la predicción en este caso depende de la contribución del valor de la variable de ensamble altura de ancla. Si se tiene una concentración grande de puntos en la parte superior de la regresión lineal se obtendrá una gráfica de residuos más disparsa, lo cual no es recomendable para ajustes normales.

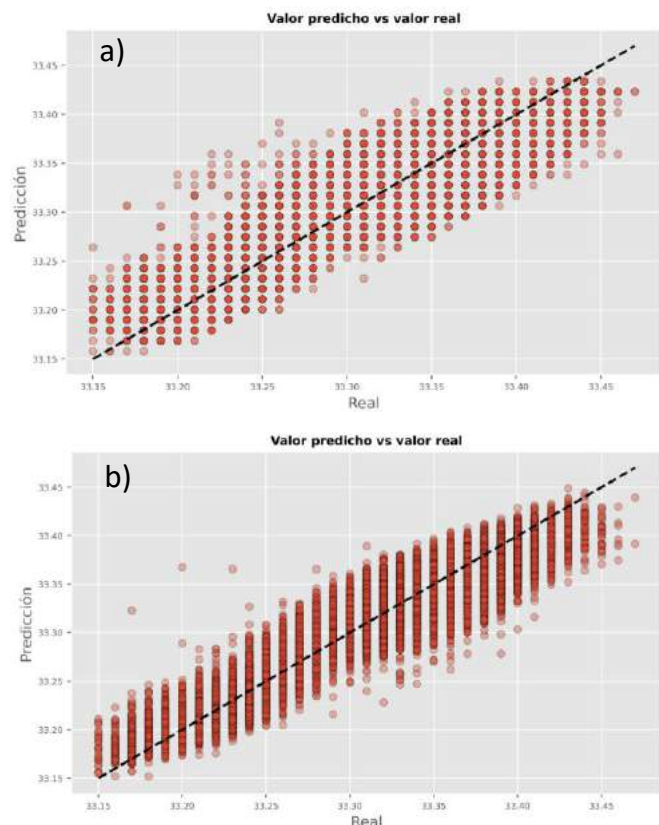


Figura 3. Valores de predicción de datos en función del valor real. Fuente: elaboración propia.

En la Figura 3b) se muestran los resultados de predicción de altura de pin empleando el modelo de regresión lineal múltiple, en el cual se tiene en cuenta el efecto de todas las variables. Para este caso hubo una aproximación mucho más real y se tienen los valores más cercanos a la media. Con esto se genera un modelo que pueda predecir de forma más exacta los valores de la característica del pin.

La Figura 4 presenta la distribución de histogramas de residuos de los modelos empleados; regresión lineal simple en a) y múltiple en b). La gráfica de distribución segmentada en la Figura 4a) muestra que el ajuste no es adecuado y que los datos no se

agrupan de una forma normal. Hay distintas frecuencias alrededor de la media. Todos los valores predichos iguales al valor real indican que hay un valor 0 en los residuos y que el modelo está funcionando de forma correcta. La densidad de histogramas de residuos mostrados en la Figura 4b) corresponden a los valores obtenidos aplicando el modelo de regresión lineal múltiple e indican que la mayoría de los datos están dando alrededor de la media. En este caso al considerar todos los predictores contribuyendo a la altura de pin, el modelo explica de mejor manera el comportamiento del pin y los residuos tienen una distribución más normal.

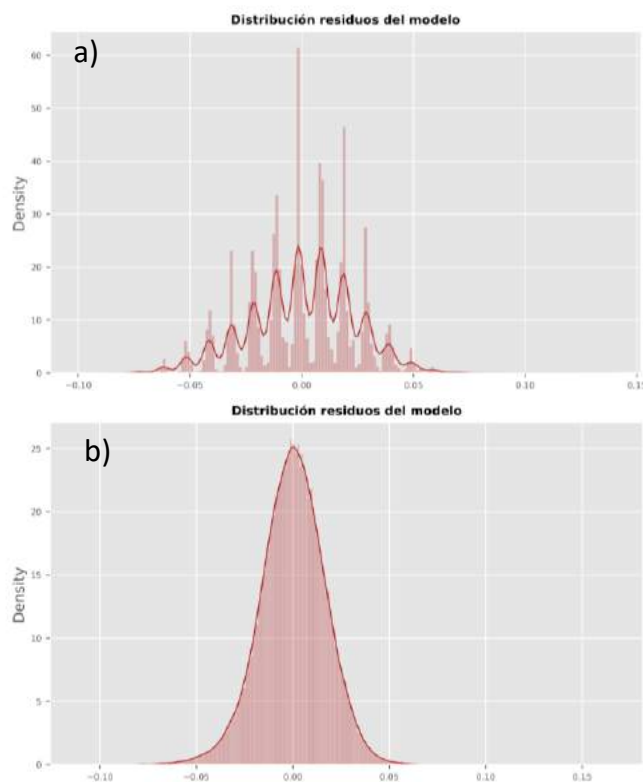


Figura 4. Densidad de distribución de residuos. Fuente: elaboración propia.

En la Figura 5 se observan los residuos de los datos obtenidos del modelo en función de la predicción realizada para los dos casos de modelo de regresión empleados. En a) el modelo de regresión simple hace que los residuos se distribuyan entre -0.10 y 0.14, lo cual muestra que hay una varianza importante.

El valor predicho para la altura de pin considerando solamente el efecto del ancla muestra

los datos de forma seccionada y esto hace que en la imagen se vean grupos verticales de datos, lo cual es un vicio del modelo y no representa la realidad de las medidas. Las condiciones ideales para hacer regresiones indican que los valores de residuo siempre deben ser aleatorios. En nuestro análisis los valores de residuos presentan tendencias.

Al aplicar el modelo de regresión lineal múltiple, las contribuciones de todas las variables en la fabricación del actuador hacen que la predicción sea más confiable y con menor varianza. Aunque el modelo constituye un sistema de entrenamiento al ajustar las variables, es importante evitar procesos de sobre ajuste.

que los valores están dispersos alrededor del residuo central 0. Sin embargo, hay cierta cantidad de datos que se pueden descartar. No obstante, al tratarse de medidas reales tomadas en un proceso de ensamble, el modelo se debe entrenar y cuidar que no se encuentre sobre ajustado. Las 7 iteraciones realizadas integrando de manera sistemática las contribuciones de cada variable, permiten tener un modelo de regresión y predicción confiable y puede emplearse para análisis con otras variables que tengan codependencia de hasta 8 grados de libertad (Fernández et al., 2018).

Conclusiones

En este trabajo se implementó de forma exitosa una metodología para el análisis de datos y predicción del comportamiento de un actuador magnético a partir de sus variables de ensamble y características críticas medidas en tiempo real. Las principales contribuciones son las siguientes;

- Inicialmente se aplica el modelo de regresión lineal simple para estimar la bondad del ajuste y definir la pertinencia de aplicación del modelo de regresión lineal múltiple.
- Se estableció una matriz de correlación de dependencia a través de factores *Pearson* entre variables de ensamble y críticas que permiten definir los parámetros de entrada para el modelo *step-forward*.
- El modelo de regresión lineal múltiple considera la contribución de las variables de ensamble y características críticas y permite predecir en un 87.5% y explicar la relación entre valores medidos durante el proceso de ensamble.
- Se identificó la manera de hacer predicciones más eficientes de tal forma que permite optimizar sistemas con el apoyo de análisis de datos.

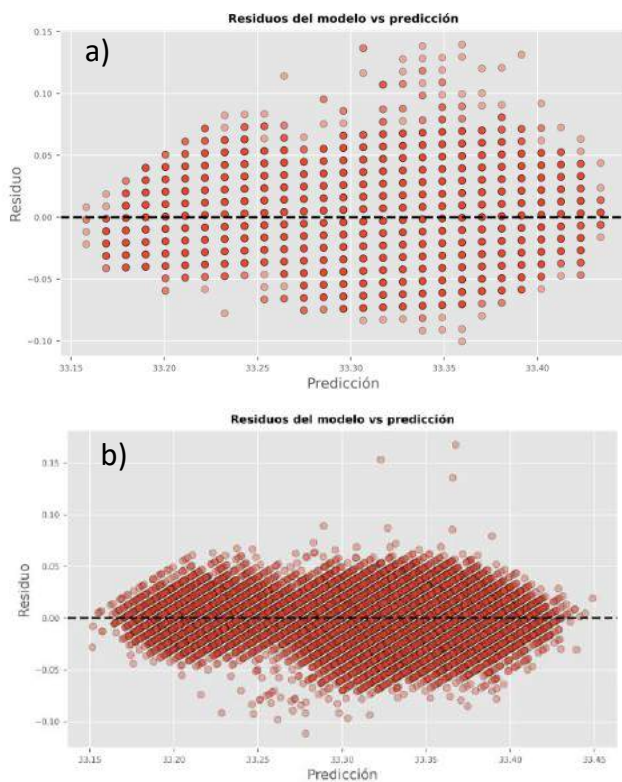


Figura 5. Dispersión de los residuos del modelo.
Fuente: elaboración propia.

Los residuos no se deben ajustar forzando a disminuir la varianza porque el modelo quedaría sobre ajustado, lo cual ocurre cuando durante el entrenamiento de un sistema de aprendizaje automático se plantea únicamente la optimización de una variable (varianza, por ejemplo) y entonces el sistema queda con algoritmos anómalos y reproduce patrones que no son generales (Gonzalo, 2020). A pesar de que en la figura 5b) se observa

Agradecimientos

A ETO Magnetic por la oportunidad de realizar trabajo de investigación con el equipo.

A CIATEQ Centro de Tecnología Avanzada A.C. por el apoyo para el desarrollo de este proyecto.

También los autores agradecemos al programa de investigadores Cátedras CONACyT por el soporte científico brindado para el desarrollo de este trabajo.

Referencias

- Amat, J., (2016). Correlación lineal y Regresión lineal simple. Creative Commons Attribution 4.0 International License. Recuperado el día 20 de febrero de 2021 de https://www.cienciadedatos.net/documentos/24_correlacion_y_regresion_lineal
- Khan, S., Liu, X., Shakil, K. A. & Alam, M. (2017). A survey on scholarly data: From big data perspective. *Information Processing & Management*, 53(4), 923-944. DOI: <https://doi.org/10.1145/3136625>
- Fernandez-Palacin, F., Lopez-Sanchez, M. A. & Munóz-Márquez, M. (2018). Stepwise selection of variables in DEA using contribution loads. *Pesquisa Operacional*, 38(1), 31-52. DOI: <https://doi.org/10.1590/0101-7438.2018.038.01.0031>
- Fundación Endesa. (2020). Electromagnetismo. Endesa Fundación. Recuperado el día 11 de marzo de 2021 de <https://www.fundacionendesa.org/es/recursos/a201908-que-es-el-electromagnetismo>.
- García, J., Molina, J. M., Berlanga, A., Patricio, M. Á., Bustamante, Á. L. & Padilla, W. R. (2018). *Ciencia de Datos: Técnicas Analíticas y Aprendizaje Estadístico. Un enfoque práctico*. Alfaomega, Tarragona.
- Gonzalo, A. (2020). ¿Qué es el sobreajuste u overfitting y por qué debemos evitarlo?. *Machine Learning para todos*. Recuperado el día 05 de marzo de 2021 de <https://machinelearningparatodos.com/que-es-el-sobreajuste-u-overfitting-y-por-que-debemos-evitarlo/>
- Gutiérrez Hinestroza, M. D. V. & Iturralde Kure, S. (2017). *Fundamentos básicos de instrumentación y control*.
- Hernández Lalinde, J. D., Espinosa Castro, J. F., Peñaloza Tarazona, M. E., Fernández González, J. E., Chacón Rangel, J. G., Toloza Sierra, C. A., ... & Bermúdez Pirela, V. J. (2018). Sobre el uso adecuado del coeficiente de correlación de Pearson: definición, propiedades y suposiciones.
- Howe, D. (2000). Magnetic actuators. *Sensors and Actuators A: Physical*, 81(1-3), 268-274. DOI: [https://doi.org/10.1016/S0924-4247\(99\)00174-0](https://doi.org/10.1016/S0924-4247(99)00174-0)
- Ismail, A., Truong, H. L. & Kastner, W. (2019). Manufacturing process data analysis pipelines: a requirements analysis and survey. *Journal of Big Data*, 6(1), 1-26. DOI: <https://doi.org/10.1186/s40537-018-0162-3>
- Kocabicak, Z. K. & Demir, U. (2020). Design and optimization of an electromechanical actuator for the latch of a foldable vehicle seat. *Materials Testing*, 62(7), 749-755. DOI: <https://doi.org/10.3139/120.111539>
- Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J. & Liu, H. (2017). Feature selection: A data perspective. *ACM Computing Surveys (CSUR)*, 50(6), 1-45. DOI: <https://doi.org/10.1145/3136625>
- Montero, R. (2016). Modelos de regresión lineal múltiple. Documentos de Trabajo en Economía Aplicada. Universidad de Granada.
- Usharani, R. & Shanthini, A.. (2020). Machine Learning Approaches For Predicting Patient Severity Levels In T2dm Complications Neuropathy. *European Journal of Molecular & Clinical Medicine*.
- Wang, S., Weng, Z. & Jin, B. (2020). A Performance Improvement Strategy for Solenoid Electromagnetic Actuator in Servo Proportional Valve. *Applied Sciences*, 10(12), 4352. DOI: <https://doi.org/10.3390/app10124352>
- Zhang, Z. (2016). Variable selection with stepwise and best subset approaches. *Annals of translational medicine*, 4(7). DOI: [10.21037/atm.2016.03.35](https://doi.org/10.21037/atm.2016.03.35)