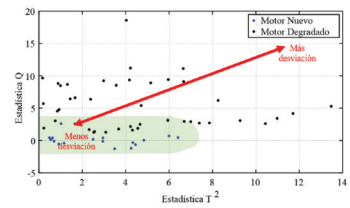


Estadística multivariable para detección de anomalías: aplicación en un turborreactor



Multivariate statistics for anomaly detection: application in a turbojet



Salma Salazar-Martínez¹, Luis Takano-De-La-Cruz³, Igor Loboda², Francisco Villarreal-Valderrama¹, Diana Hernández-Alcantara³, Luis Amézquita-Brooks¹

¹ Universidad Autónoma de Nuevo León, Av. Universidad SN, San Nicolás de los Garza, Nuevo León, México.

² Instituto Politécnico Nacional, Escuela Superior de Ingeniería Mecánica y Eléctrica, Unidad Culhuacán, Ciudad de México, México.

³ Universidad de Monterrey, Av. Ignacio Morones Prieto, San Pedro Garza García, Nuevo León, México.

DOI: <https://doi.org/10.6036/10921> | Recibido: 13/mar/2023 • Inicio Evaluación: 09/mar/2023 • Aceptado: 08/jun/2023

To cite this article: SALAZAR-MARTÍNEZ, Salma; TAKANO-DE-LA-CRUZ, Luis; LOBODA, Igor; VILLARREAL-VALDERRAMA, Francisco; HERNANDEZ-ALCANTARA, Diana; AMÉZQUITA-BROOKS, Luis. MULTIVARIATE STATISTICS FOR ANOMALY DETECTION: APPLICATION IN A TURBOJET. DYNA. March-April 2024. Vol. 99, n.2, pp.208-214. DOI: <https://doi.org/10.6036/10921>

ABSTRACT

- Although the computational power of embedded systems has increased in recent years, these systems are increasingly being taxed with more tasks. This raises the interest for computationally lean algorithms which are able of rendering process operation more efficient and reliable. This is particularly relevant in the case of flight computers for autonomous aircraft. Fault detection, isolation and identification assist in management strategies to improve both predictive maintenance and operational safety. This article combines a principal component-based representation with multivariate statistics to detect and isolate anomalies in a process. The resulting algorithm is computationally lean and was validated with respect to experimental measurements in a turbojet before and after years of operation. The results show that the developed algorithm is capable of successfully determining the fouling components in the turbojet.
- Key Words:** Fault detection, PCA, multivariate statistics, sensor fusion, process monitoring.

RESUMEN

Aunque el poder computacional de los sistemas embebidos ha aumentado en los últimos años, estos sistemas cada vez son responsables de un número mayor de tareas. Esto aumenta el interés por los algoritmos con bajo coste computacional para lograr que la operación de los procesos sea más eficiente y confiable. Esto es de particular relevancia en el caso de las computadoras de vuelo usadas en las aeronaves autónomas. La detección, el aislamiento y la identificación de fallas ayudan en estrategias de gestión para mejorar tanto el mantenimiento predictivo como la seguridad operativa. Este artículo combina una representación basada en componentes principales con estadística multivariada para detectar y aislar anomalías en procesos. El resultado es un algoritmo con bajo coste computacional, el cual se validó con mediciones experimentales en un turborreactor antes y después de años de funcionamiento. Los resultados muestran que el algoritmo desarrollado es capaz de determinar con éxito los componentes de ensuciamiento en el turborreactor.

Palabras Clave: Detección de fallas, Análisis de componentes principales, estadística multivariada, fusión de sensores, monitoreo de procesos.

1. INTRODUCCIÓN

Una falla puede definirse como cualquier cambio negativo en el estado técnico de una pieza o componente de una máquina que le impide realizar su función satisfactoriamente. En la industria, las fallas causan pérdidas en tiempos de operación de máquinas y gastos adicionales por el mantenimiento emergente [1,2]. Por esto, el mantenimiento es reconocido como una parte integral de la vida de una máquina y puede mejorar la productividad a lo largo de la operación si se detectan a tiempo y reparan las fallas [3,4,5].

Existen tres tipos de estrategias de mantenimiento: correctivo, preventivo y predictivo. El mantenimiento correctivo se refiere a la reparación de una máquina cuando ya ha ocurrido la falla, mientras que el preventivo se define en periodos de tiempo y está programado dentro de la cadena de operación. En años recientes se ha postulado la idea de mantenimiento predictivo para mejorar la respuesta ante los cambios en el desempeño de la maquinaria. Su implementación, sin embargo, depende de la capacidad de conocer el estado o condición de la máquina a partir de su revisión rutinaria o continua. Los controles de calidad actuales exigen sistemas de diagnóstico más precisos que permitan mejorar las rutinas de mantenimiento al basarse en la degradación de los componentes y no en tiempo de uso [1,5].

Para poder realizar un diagnóstico adecuado, es necesario determinar cuáles son los parámetros por evaluar. No obstante, dependen del equipo a analizar, las condiciones de operación y la resolución deseada de la predicción. Consecuente a esto, el Subcomité 5 del Comité Técnico 108 de la Organización internacional de Estándares (ISO/TC108/SC5) se ha enfocado en estándares para el monitoreo de la condición y diagnóstico de máquinas; publicando la ISO 13374-1:2003. Este estándar establece una guía general para la especificación de softwares relacionados con el procesamiento de datos, comunicación y presentación del monitoreo y diagnóstico de la condición de una máquina [6]. El resultado de este análisis es la predicción de una falla potencial o una pérdida de eficiencia a partir de la condición actual del equipo.

Lo anterior deja ver que, aunque existen elementos de estandarización en el manejo del diagnóstico y detección de fallas, cada proceso requiere de conocimiento experto específico. Este conocimiento por lo general se ha desarrollado a través de la experiencia en campo. En el caso de las aplicaciones de turbo maquinaria, existen dos objetivos de diagnóstico principales a considerar: la necesidad de una operación segura de los sistemas de propulsión

Sección	Modo de fallo	Efecto
Compresor	<ul style="list-style-type: none"> · Ensuciamiento · Erosión · Corrosión 	<ul style="list-style-type: none"> · Reducción de la razón de presiones · Pérdida de la eficiencia isentrópica · Reducción de la capacidad de flujo · Incremento en la temperatura de descarga
	<ul style="list-style-type: none"> · Incremento del espacio entre las paredes 	<ul style="list-style-type: none"> · Reducción de la razón de presiones · Pérdida de la eficiencia isentrópica · Reducción de la capacidad de flujo · Fuga del aire comprimido
Cámara de combustión	<ul style="list-style-type: none"> · Corrosión o agrietamiento 	<ul style="list-style-type: none"> · Pérdida en la eficiencia de la combustión
	<ul style="list-style-type: none"> · Obstrucción de los inyectores 	<ul style="list-style-type: none"> · Variaciones en el flujo de combustible · Combustión incompleta · Pérdida de eficiencia en la combustión
Turbina	<ul style="list-style-type: none"> · Incremento del espacio entre la punta del álabe y la pared 	<ul style="list-style-type: none"> · Reducción de la razón de presiones · Incremento de la capacidad de flujo · Pérdida de eficiencia isentrópica
	<ul style="list-style-type: none"> · Erosión · Ensuciamiento · Corrosión 	<ul style="list-style-type: none"> · Reducción de la razón de presiones · Incremento de la capacidad de flujo · Pérdida de eficiencia isentrópica · Reducción en los márgenes de desplome

Tabla 1. Tipos comunes de fallas y su efecto en el desempeño de los componentes [8,9,10,11].

en aeronaves y la capacidad de asegurar la estabilidad del suministro eléctrico del cual dependen muchas instituciones. En estas aplicaciones, se ha estimado que el 70% de las fallas se omiten debido a que las herramientas implementadas para detectar, identificar y diagnosticar fallas son ineficientes [7]. Los mecanismos de falla más comunes en el generador de gas se presentan en la Tabla 1 acorde a [8,9,10,11]. Esta tabla muestra que, en general, es posible asociar comportamientos específicos, observables a través de mediciones, con modos de fallo particulares. Es decir, es posible diagnosticar fallas a nivel de los componentes e incluso determinar la causa raíz de las fallas evaluando la relación entre las propiedades termodinámicas del gas en diferentes puntos del conducto de flujo. La Tabla 1 representa el resumen del conocimiento experto que es necesario para definir las variables a medir en cada caso. Por ejemplo, si se observa que en una aplicación particular de turborreactores existen fallas recurrentes debido al incremento del espacio entre paredes del compresor, entonces sería recomendable monitorear la razón de presiones entre el compresor y sus secciones adyacentes.

Debido a las observaciones anteriores, los algoritmos para detectar fallas basados en modelos termodinámicos han sido especialmente populares y exitosos en su aplicación [12,13]. Sin embargo, dificultan la integración de la información de sensores cuando no se tiene un modelo explícito. Como ejemplo de este fenómeno está el nivel de vibración, el cual es un importante indicador del estado de salud del motor del cual no se tiene un modelo explícito. Es por ello que se han desarrollado herramientas de diagnóstico basadas en datos. Entre las alternativas disponibles, destacan aquellas basadas en un Análisis de Componentes Principales (PCA por sus siglas en inglés). En [14] se integró esta herramienta para combinar un modelo de desempeño con la vibración de un turborreactor, incrementando en un 25% la precisión del diagnóstico. Sin embargo, este algoritmo resultante aún requería una estructura de modelo tentativa para operar correctamente. En otras aplicaciones, como la evaluación de daños estructurales y el monitoreo de sistemas de sistemas de energía en edificios, se han implementado con éxito herramientas de detección de fallas que no requieren de un modelo a priori [15,16]. En estos algoritmos se evaluó el ajuste de los datos respecto a un espacio de componentes principales desarrollado con una estructura sin fallas.

Con base en la discusión anterior se puede concluir que existe una necesidad de herramientas apropiadas para la detección de fallas en turbo maquinaria que no estén basadas explícitamente en modelos termodinámicos. En este sentido, es posible tratar al turborreactor simplemente como un sistema que relaciona diferentes variables de entrada y salida. Este tipo de sistemas requiere de herramientas de análisis multivariado para obtener buenos resultados. Por otro lado, la metodología de PCA es útil para analizar datos en un contexto de detección y diagnóstico de fallas, ya que permite identificar las variables que más influyen en el comportamiento normal o anormal del sistema, así como detectar y clasificar las anomalías que se presenten. Además, PCA facilita la interpretación de los datos reduciendo la dimensionalidad de los mismos, lo que contribuye a simplificar el entendimiento del problema y a proponer soluciones adecuadas [17]. Debido a estas características, en este artículo se propone utilizar PCA para fusionar y compactar información, de manera que se genere una huella digital característica del motor en condiciones nominales. Después, propone utilizar herramientas de estadística multivariada para cuantificar la distorsión de la relación entre las variables medidas conforme a la degradación del turborreactor. Para validar el algoritmo resultante, se utilizan datos reales de operación medidos en 2012 y en 2019, tiempo durante el cual se utilizó el turborreactor sin requerir mantenimiento correctivo, por lo cual sufrió de un nivel de degradación normal. Este algoritmo tiene bajo coste computacional, (requiere pocos recursos de cómputo, como memoria, procesador o ancho de banda, para funcionar de manera eficiente y efectiva) y permite identificar las anomalías en los datos causadas por el deterioro del turborreactor.

2. MÉTODOS DE ESTADÍSTICA MULTIVARIADA

Los equipos actuales están equipados con un gran número de sensores, los cuales miden diferentes variables en los procesos. Los métodos estadísticos multivariados para monitorear procesos son algoritmos que extraen información importante de conjuntos grandes de datos. Por lo cual, la principal característica de estos métodos es la capacidad de manejar grandes cantidades de datos altamente correlacionados de alta dimensionalidad y en condiciones de ruido. Esto generalmente consiste en [17,18]:

- Evaluar las condiciones de operación.
- Construir una base de datos con datos característicos.
- Generar un modelo en condiciones normales.
- Proyectar los nuevos datos con respecto a los normales.
- Juzgar si los nuevos datos son estadísticamente normales o anormales.
- Identificar las variables responsables de las anomalías.
- Determinar la causa raíz de la anomalía.

El punto de partida es el conjunto de los datos medidos $z(i \times j)$, de i mediciones y j variables. Debido a que las variables medidas poseen diferente significados físicos y escalas, es necesario normalizar los datos de manera que tengan una media de cero con varianza unitaria. Para ello, se calculan el promedio y la varianza muestrales μ_b y σ_b de los datos correspondientes al modelo base (proceso en condiciones nominales) [17]:

$$\mu_b = \frac{1}{n} \sum_{i=1}^n z_{i,j} \quad (1)$$

$$\sigma_b^2 = \frac{1}{n-1} \sum_{i=1}^n (z_{i,j} - \mu_b)^2 \quad (2)$$

Después, se procede a normalizar los datos mediante:

$$Z_{i,j} = \frac{z_{i,j} - \mu_b}{\sigma_b} \quad (3)$$

En este caso es importante notar que los parámetros μ_b y σ_b son estadísticos (muestrales) y no los parámetros poblacionales.

2.1 ANÁLISIS DE COMPONENTES PRINCIPALES

El objetivo de *PCA* es obtener una matriz de transformación $P(j \times j)$, que permita mapear los datos estandarizados a un espacio de componentes principales [18,19]. Los datos transformados $T_d(i \times j)$ están descritos por:

$$T_d = ZP \quad (4)$$

Este algoritmo busca un conjunto de ejes ortogonales que capturen la máxima varianza de los datos. Por tanto, con esta transformación se busca que la matriz de covarianza $C(j \times j)$ de los datos transformados sea diagonal:

$$C = \frac{1}{n-1} T_d^T T_d = \text{diagonal} \quad (5)$$

Esto implica que se transforma de un conjunto de variables correlacionadas entre sí, a un conjunto de variables sin correlación. Si se seleccionan solamente los componentes principales que proveen de la mayor varianza, es posible reducir la dimensionalidad de los datos a una nueva dimensión l utilizando una submatriz $P_l(j \times l)$ de P (tal que $l < j$) con la menor pérdida de información en términos estadísticos para ese número de variables. Al realizar este análisis generalmente se observa que el número de componentes que capturan la mayor varianza son en efecto menos que la dimensionalidad original de los datos [20]. Mientras que el resto de los componentes principales se asocian con incertidumbre en el proceso.

Una vez que se ha obtenido la matriz de transformación con los datos nominales, es posible monitorear cambios en el proceso a partir de nuevas mediciones. Por ejemplo, para cada nueva medición de $x_i(i \times j)$ se obtiene su normalización $X_i(i \times j)$ aplicando las ecuaciones (1) y (2). La proyección de la nueva medición en el

espacio reducido de componentes principales $\bar{T}(i \times l)$ estará dada por una transformación mediante una matriz de transformación reducida $P_l(j \times l)$ con la dimensión $l < j$ que mantenga los componentes principales más relevantes [19, 20]:

$$\bar{T} = X_i P_l \quad (6)$$

El producto $P_l P_l^T$ genera una matriz diagonal, la cual en un escenario ideal consiste en una matriz unitaria, caso en el que sería posible reconstruir a X_i a partir de \bar{T} sin pérdidas de información. En condiciones reales, existe una pérdida de información debido a la compresión y descompresión de los datos mediante la matriz de dimensiones reducidas. Al tratar de recuperar la medición original X_i a partir \bar{T} se obtiene [19,20]:

$$\bar{X}_i = \bar{T} P_l^T = X_i P_l P_l^T \quad (7)$$

De esta manera, la pérdida de información introducida por el uso de la matriz de transformación reducida está dada por $(X_i - \bar{X}_i)$.

2.2 DETECCIÓN DE ANOMALÍAS

Una vez obtenida la matriz de transformación P_l en el presente artículo se propone utilizar esta matriz como una huella digital representativa de la operación nominal del proceso. Entonces, al mapear los nuevos datos con el modelo *PCA* se podrán detectar e identificar condiciones anormales en el proceso mediante el uso de diferentes métricas. El monitoreo por medio de estadísticas se realiza para detectar fallas, mientras que las gráficas de contribución son empleadas para identificar el origen de las fallas [15,16,19]. Por lo tanto, se pueden aislar los siguientes fenómenos: las mediciones de un sensor alcanzan valores inusuales, múltiples sensores se desvían de las condiciones normales de operación o el proceso está operando bajo condiciones transitorias.

2.3 ESTADÍSTICA Q

En este caso, el modelo desarrollado por medio de *PCA* se utiliza para evaluar las nuevas mediciones. Las diferencias entre la proyección de los datos de regreso al espacio original y los datos estandarizados se denominan residuos, r_i , es decir:

$$r_i = X_i - \bar{X}_i \quad (8)$$

Combinando las ecuaciones (7) y (8) se puede calcular el residuo en términos de una nueva medición y la matriz de transformación:

$$r_i = X_i (I - P_l P_l^T) \quad (9)$$

Este residuo describe la contribución de cada variable medida al error en la reconstrucción de los datos después de la compresión realizada por el modelo de *PCA* [15,16,19]. Es decir, este residuo cuantifica la varianza de los nuevos datos que no fue capturada por el modelo de *PCA* obtenido con el proceso normal. La estadística Q muestra que tan bien se ajustan las nuevas mediciones a los componentes principales a partir del residuo. Este indicador es un escalar que representa la calidad del ajuste entre los datos y el modelo. La estadística Q muestra la homogeneidad en un grupo de muestras; es decir, la varianza causada por elementos estocásticos y aquella relacionada a un cambio en el comportamiento en el

conjunto de variables medidas. De manera específica la estadística Q se puede calcular a partir de los residuos como [15,16,19]:

$$Q = r_i r_i^T \quad (10)$$

2.4 ESTADÍSTICA HOTELLINGS

La estadística Hotellings (T^2) es una generalización de la estadística t de Student que se utiliza para verificar hipótesis sobre la media de una base de datos multivariada. Dentro del contexto de un análisis de PCA, la estadística T^2 pondera el tamaño de la proyección de un dato particular sobre los ejes principales calculados con PCA. Es decir, que tanto se aleja un dato particular de la media de la base de datos con la que se realizó el análisis PCA [15,16,19]. En el contexto de detección de fallas, la estadística T^2 permite cuantificar y detectar anomalías en nuevas mediciones; mientras más alto sea el valor de T^2 mayor es la probabilidad de una anomalía en el proceso. Esta estadística está definida como la suma del cuadrado de los datos transformados con varianza (λ) unitaria, es decir [15,16,19]:

$$T_i^2 = \sum_{i=1}^l \left(\frac{X_i P_i}{\lambda_i} \right)^2 \quad (11)$$

El uso conjunto de las estadísticas Q y T^2 permite determinar si un dato particular es anómalo en dos sentidos: 1) si no está representado por los componentes principales mayores (estadística Q) y 2) si se aleja de la media de los datos nominales (estadística T^2). La figura 1 muestra este concepto de manera gráfica. En esta figura se muestran datos en tres dimensiones con diferentes combinaciones de Q y T^2 en términos de dos componentes principales. En particular, un valor alto de Q implica que el dato no se puede reconstruir con los componentes principales reducidos, mientras que un valor T^2 implica lejanía del origen de los ejes principales. En general, es posible afirmar que una combinación de valores altos en las estadísticas Q y T^2 es indicativa de una anomalía.

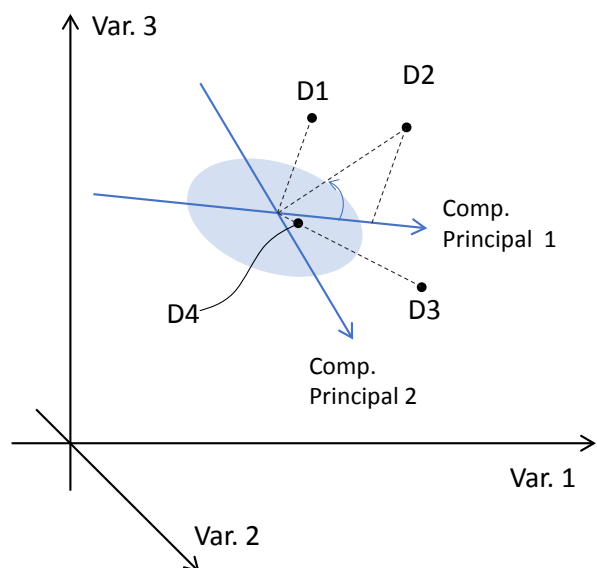


Figura 1. Proyección de datos específicos de tres dimensiones en un subespacio de dos componentes principales y su relación con las estadísticas Q y T^2 . En particular: D1 (Q : alto, T^2 : bajo), D2 (Q : alto, T^2 : alto), D3 (Q : bajo, T^2 : alto), D4 (Q : bajo, T^2 : bajo).

3. ALGORITMO DE DETECCIÓN DE FALLAS PROPUESTO

La combinación de las estadísticas Q y T^2 permite detectar un cambio en la operación de un proceso, que en este caso podría

representar una falla si se determina que el cambio no se puede atribuir a una modificación en las condiciones de operación del mismo [15,16,19]. Si las mediciones muestran un valor anormal de T^2 , pero un valor normal de Q , puede representar un cambio en la región de operación. En cambio, para identificar una falla que ya se ha detectado, la influencia de las variables sobre el residuo muestra aquellas mediciones que están generando una varianza anormal. El procesamiento de datos necesario para calcular este algoritmo se resume en la Figura 2.

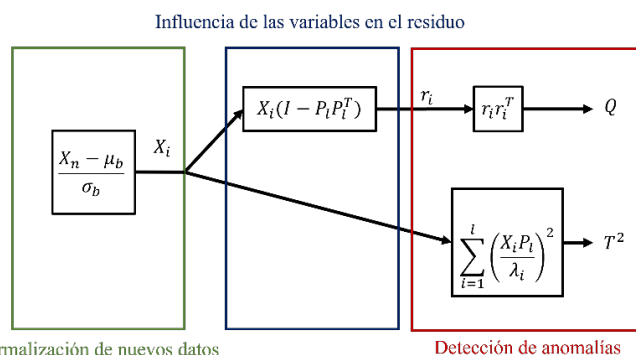


Figura 2 Operación del algoritmo generalizado para detección de fallas.

En resumen, el método generalizado para detección de fallas consiste en:

- A. Etapa de identificación de operación nominal.
 1. Estandarización de los datos. Es decir, calcular la base de datos Z .
 2. Desarrollar un modelo de PCA con base en los datos del proceso normal. Es decir, calcular la matriz P .
 3. Reducir la dimensión del espacio de PCA. Es decir, calcular la matriz P_i .
- B. Etapa de monitoreo y detección de fallas (Figura 2).
 4. Escalar los nuevos datos con la desviación estándar y el promedio obtenidos con los datos del proceso normal. Es decir, la base de datos X_i .
 5. Proyectar los nuevos datos al espacio de componentes principales y calcular los residuos r_i de cada variable.
 6. Calcular las estadísticas Q y T^2 para determinar si ha ocurrido una falla.
 7. Si ha ocurrido alguna falla, identificar la causa con la gráfica de contribución al residuo r_i correspondiente.

Finalmente, además de estas estadísticas existen otros elementos que se pueden obtener a partir de PCA que pueden servir para propósitos de detección y diagnóstico de fallas, como aquellos reportados en [19,20,22]. No obstante, para esta aplicación los resultados experimentales que se presentan a continuación muestran que estos parámetros son suficientes.

3.1 CASO DE ESTUDIO: DEGRADACIÓN DE UN TURBORREACTOR

Los turborreactores son la fuente primaria de empuje para aeronaves civiles, militares y comerciales. Estos mecanismos son similares en operación a las turbinas de gas. Ambos requieren de sistemas adecuados para detectar, identificar y aislar fallas.

El turborreactor utilizado como banco de pruebas es el modelo SR-30, manufacturado por Turbine Technologies. Este banco de pruebas, mostrado en la Figura 3a, ha sido utilizado con fines de docencia en la Universidad Autónoma de Nuevo León. Esta tiene una velocidad de rotación máxima de 87,000 RPM y empuje máximo de 178N. Así mismo, cuenta con un conjunto de sensores que

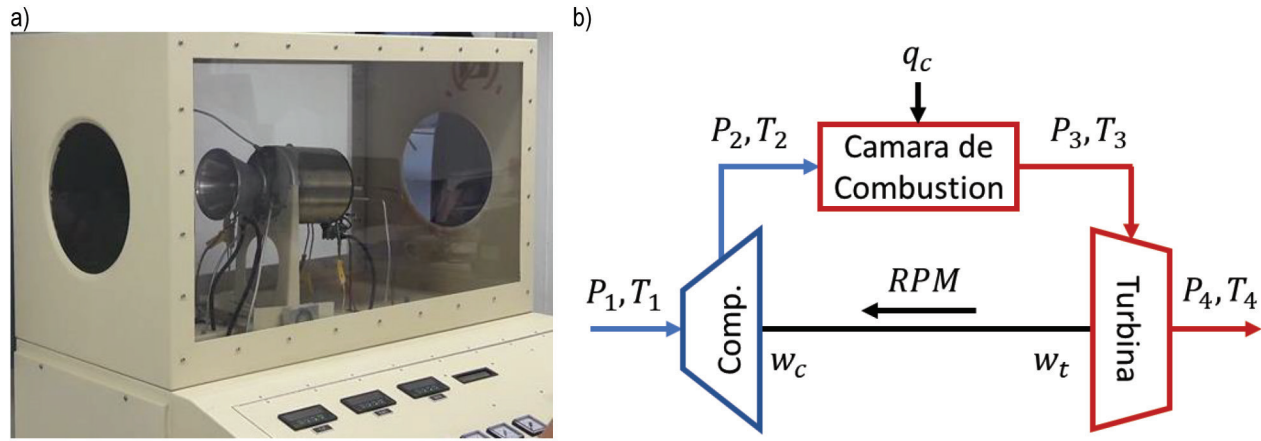


Figura 3. a) Banco de pruebas para el turborreactor SR-30. b) Diagrama de un turborreactor.

permiten realizar mediciones de empuje (T), velocidad de rotación (SS), consumo de combustible (FF) y propiedades del gas en cada estación.

En este proceso las variables normalmente medidas son presión y temperatura a la entrada del compresor (P_1, T_1), de la cámara de combustión (P_2, T_2), de la turbina (P_3, T_3) y de salida de la turbina (P_4, T_4). La Figura 3b muestra un esquema del funcionamiento del turborreactor.

El banco de pruebas ha sido usado previamente para adquirir datos con el fin de desarrollar y evaluar diferentes modelos matemáticos para estimar el desempeño de turborreactores [21] y se ha evaluado con éxito la aplicabilidad de diferentes técnicas de minado de datos probando compatibilidad con los experimentos [22]. En este caso, los datos adquiridos corresponden a la operación del proceso nominal en 2012, y aquellos obtenidos con el turborreactor degradado después de 7 años de operación. En total, se tomaron datos de 25 puntos de operación estacionarios del motor nominal alrededor del régimen dado por la velocidad de 60,000 RPM. Para el motor degradado, se tomaron datos en 42 puntos en condiciones de operación similares. Para el análisis, se proponen las siguientes variables de entrada al algoritmo de monitoreo, de acuerdo con la instrumentación disponible:

$$X_i = \left[\frac{P_2}{P_1}, \frac{P_3}{P_2}, \frac{P_4}{P_3}, SS, \frac{T_2}{T_1}, \frac{T_3}{T_2}, \frac{T_4}{T_3}, T \right] \quad (12)$$

Es decir, se consideran como entrada al algoritmo la ganancia/pérdida de temperatura y presión en cada etapa del turborreactor, así como la velocidad de rotación de la turbina y el empuje resultante.

3.1.1 RESULTADOS OBTENIDOS IMPLEMENTANDO ESTADÍSTICA MULTIVARIADA

Siguiendo la metodología detallada en la Sección 2.2 y al inicio de la Sección 3, se obtiene una representación en componentes principales utilizando los datos del motor nominal. Se observa en la Figura 4 que la transformación de los datos del motor degradado sobre el espacio de los dos primeros componentes principales no se ajusta igual que los datos con el motor nominal. El comportamiento del motor degradado muestra mayor varianza sobre los componentes y está alejado del promedio de los datos.

La diferencia entre el comportamiento del motor en ambas condiciones se muestra en la Figura 5 mediante las estadísticas Q y T^2 . Acorde con esta figura, los datos correspondientes al motor degradado se pueden interpretar como datos de un proceso con un mayor índice de falla.

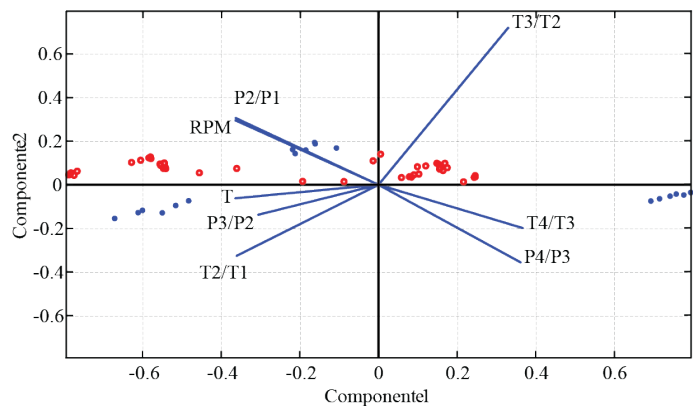


Figura 4. Transformación de los datos del motor nominal (azul) y del motor degradado (rojo) en el espacio de componentes principales. Las líneas azules representan la orientación de las variables originales.

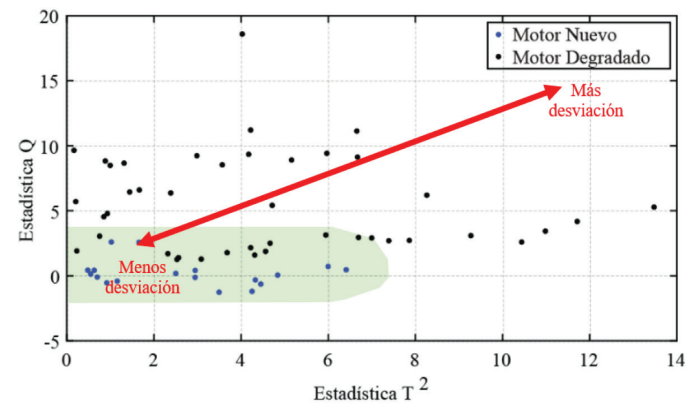


Figura 5. Estadística Q y Hotellings. Esta gráfica permite observar la degradación en la operación del turborreactor. Los datos que representan al motor degradado muestran mayor variación ajena al comportamiento capturado por los componentes principales. En este caso, las muestras deberían estar contenidas dentro de la región que comprende el comportamiento del motor nominal (verde).

Ya que se detectó la degradación del proceso, se procede a analizar la influencia de cada variable analizada sobre el residual. En la Figura 6, se promedian los valores obtenidos con las 42 muestras del motor degradado, con la desviación estándar respectiva para cada variable, lo cual resulta en: 0.1624 (P_2/P_1), 0.7500 (P_3/P_2), 0.1513 (P_4/P_3), 0.1719 (SS), 0.2037 (T_2/T_1), 0.2164 (T_3/T_2), 0.5366 (T_4/T_3) y 0.1498 (T). Se observa claramente que la presión en la cámara de combustión y la razón de temperaturas en la turbina presentan el comportamiento más alejado del nominal.

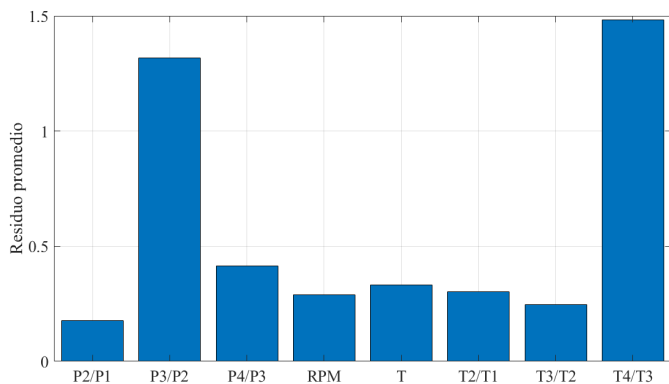


Figura 6. Promedio de la influencia de los residuos entre los datos y el espacio de componentes principales. Se observa que el mayor residuo se debe a la razón de presiones de la cámara de combustión y a la eficiencia isentrópica de la turbina.

Una ventaja de este análisis es que simplifica el problema del dimensionamiento y análisis de i mediciones de 10 variables (siendo en este caso miles de puntos de medición) a dos gráficas. Por medio de estas representaciones gráficas es posible detectar e identificar fallas de manera eficiente y efectiva, ya sea de manera manual (es decir, mediante la inspección humana) o a través un posterior algoritmo de clasificación (como una red neuronal).

3.1.2 VALIDACIÓN TERMODINÁMICA

Con el objetivo de validar los hallazgos anteriores, en esta sección se presenta un análisis del desempeño del turborreactor utilizando teoría de termodinámica. En esta etapa se comparan los indicadores más importantes del estado del turborreactor:

- Razón de presiones en cada etapa. Indica la capacidad de ejercer/extraer trabajo en el compresor y la turbina, y el nivel de pérdida de presión en la cámara de combustión.

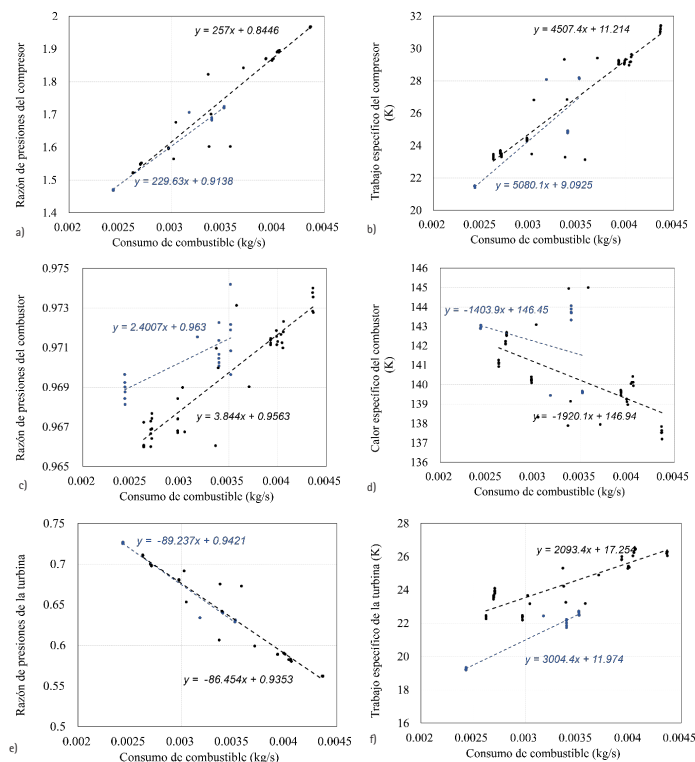


Figura 7. Análisis termodinámico del turborreactor en condiciones estacionarias variando flujo de combustible, nominal (azul) y degradado (negro). a) Razón de presiones del compresor. b) Trabajo específico del compresor. c) Razón de presiones de la cámara de combustión. d) Calor específico de la cámara de combustión. e) Razón de presiones de la turbina. f) Trabajo específico de la turbina.

$$\pi = \frac{P_s}{P_e} \quad (13)$$

- Trabajo específico. Indica la efectividad de la turbina y el compresor al ejercer trabajo sobre el fluido.

$$w_e = \frac{\dot{W}}{\dot{m}C_p} = T_s - T_e \quad (14)$$

- Calor específico de la combustión. Muestra que tan efectivamente se está transfiriendo la energía liberada por el proceso de combustión hacia el fluido. En esta etapa se considera que el flujo másico total es aproximadamente igual al flujo másico en la entrada. Es decir, $\dot{m}_0 + \dot{m}_f \approx \dot{m}$ por lo tanto el calor específico resulta:

$$q_c = \frac{\dot{Q}}{\dot{m}C_p} = T_s - T_e \quad (15)$$

La Figura 7 muestra los resultados del análisis sobre el compresor. En este caso, la tendencia en la razón de presiones (Fig. 7a) y el trabajo específico (Fig. 7b) es consistente con el turborreactor nominal y degradado. Lo cual indica que el compresor no tiene un nivel de degradación significativo. La calidad de las regresiones lineales se puede evaluar mediante su coeficiente de determinación r^2 y su estadística F. Para la Fig. 7a se tiene en el caso normal $r^2=0.97$ y $F=770$; y para el degradado $r^2=0.94$ y $F=717$. Para la Fig. 7b se tiene en el caso normal $r^2=0.78$ y $F=58$; y para el degradado $r^2=0.88$ y $F=296$. Estos resultados muestran un buen ajuste y suficiente significancia estadística, ya que para un nivel de significancia de 0.01 es necesario que $F > 6.359$.

En el caso de la cámara de combustión, las Figuras 7c (normal $r^2=0.56$, $F=20.6$ y degradado $r^2=0.86$, $F=264$) y 7d (normal $r^2=0.13$, $F=2.5$ y degradado $r^2=0.4$, $F=27$) muestran los resultados del análisis. En el caso del calor específico (Fig. 7d), el análisis muestra que existe poca relación entre las variables dependiente e independiente en ambos casos. Sin embargo, la razón de presiones (Fig. 7c) muestra una mayor relación entre las variables y una clara diferencia entre el 2012 y el 2019. Esto sugiere que la razón de presiones ha cambiado con el tiempo, lo cual se puede relacionar con un problema en la inyección del combustible.

Finalmente, el comportamiento termodinámico de la turbina se muestra en las Figuras 7e (normal $r^2=0.98$, $F=801$ y degradado $r^2=0.95$, $F=839$) y 7f (normal $r^2=0.96$, $F=477$ y degradado $r^2=0.78$, $F=142$). La tendencia en la razón de presiones muestra un comportamiento bastante consistente a lo largo del tiempo y un buen ajuste en la regresión. Sin embargo, el trabajo específico tiene una tendencia notablemente diferente, la turbina degradada extrae un trabajo específico mayor al fluido. Esto muestra que, para mantener la misma razón de presiones es necesario extraer más energía del flujo. Una de las causas probables son ensuciamiento de los álabes, que cambian la geometría efectiva del perfil.

4. DISCUSIÓN

Se analizó el comportamiento del turborreactor siguiendo el procedimiento propuesto en la Sección 3. Las estadísticas Q y T^2 resultantes (Fig. 5) indican una anomalía en los datos del motor degradado. El residuo de cada variable con respecto a la representación de componentes principales (Fig. 6) sugiere que la razón de presiones de la cámara de combustión y la razón de temperaturas de la turbina son la causa de las anomalías. Siguiendo la Tabla 1, finalmente se diagnostica que puede existir ensuciamiento en

los inyectores de combustible y ensuciamiento de los álabes de la turbina.

Por otro lado, se realizó un análisis termodinámico para validar los resultados obtenidos mediante el algoritmo desarrollado. En particular, las Figuras 7a y 7b muestran que las tendencias en el comportamiento de la razón de presiones y la eficiencia específica de la turbina han cambiado después de años de operación, lo cual corresponde con los hallazgos del método propuesto.

Si bien se muestra que ambos enfoques han llegado a resultados similares, se hace un énfasis en las propiedades importantes del algoritmo propuesto: *i)* permite combinar variables de diferente naturaleza física entre sí para generar una *huella digital* del modelo nominal y *ii)* puede ser implementado sin conocimiento de los fenómenos físicos subyacentes al proceso a analizar.

5. CONCLUSIONES Y TRABAJO FUTURO

Se presenta un algoritmo de detección e identificación de fallas aplicado para analizar el comportamiento de un turborreactor degradado por tiempo de operación. Este método está basado en el monitoreo en condiciones estacionarias, por lo cual está limitado a aplicaciones donde la correlación temporal no es un factor determinante para detectar la falla. No obstante, en el caso del turborreactor, se obtuvieron buenos resultados aún con esta limitante.

El algoritmo desarrollado se comprende de los siguientes pasos:

- 1) Generar una representación de componentes principales utilizando la información disponible del proceso en condiciones nominales.
- 2) Normalizar los nuevos datos medidos durante la operación del proceso.
- 3) Utilizar la representación en componentes principales del proceso nominal para evaluar cambios en el comportamiento a través de nuevos datos medidos.
- 4) Evaluar la diferencia entre los nuevos datos medidos y la información del proceso nominal utilizando estadística multivariada (estadística Q y T^2).
- 5) En caso de haber detectado una falla, identificar su origen analizando la influencia de las variables sobre el residuo.

Los resultados experimentales muestran que el algoritmo es capaz de detectar la degradación e identificar el componente específico en un turborreactor. En comparación con los algoritmos utilizados convencionalmente para detectar fallas en turborreactores, el presente provee de las siguientes ventajas: *i)* independencia de la naturaleza de las variables al permitir mezclar variables provenientes de diferentes mediciones físicas en un mismo sistema de diagnóstico, *ii)* bajo coste computacional lo cual permite su implementación en computadoras de vuelo con recursos limitados *iii)* puede implementarse sin conocimientos previos del proceso y *iv)* el desarrollo del modelo nominal es directo y no presenta dificultades en su generación. Sin embargo, una limitante importante es que se deben evaluar puntos de operación en condiciones estáticas cercanas ya que es posible que el proceso no mantenga sus propiedades en rangos de operación amplios ni condiciones transitorias. Por otro lado, el rango de operación se puede extender mediante el cálculo de un conjunto de modelos locales, lo cual puede representar un trabajo futuro.

Por otro lado, el método presentado en este artículo también puede ser aplicable a una gran variedad de procesos. Aunque formalmente sería necesario validar experimentalmente cada una de

estas aplicaciones, es posible afirmar que la aplicabilidad del método presentado estará sujeta a las siguientes condiciones:

- Acceso a información del proceso en condiciones nominales.
- Que las variables medidas tengan suficiente poder descriptivo.

El segundo punto se podría validar midiendo la calidad de un modelo basado en correlaciones lineales; por ejemplo, como se presenta en [22] para el caso del turborreactor. En este caso, se trata de verificar si es posible construir un modelo predictivo a partir de las mediciones disponibles utilizando una descomposición en componentes principales. En caso afirmativo (es decir, si se puede construir un modelo), la probabilidad de que el método de detección de fallas propuesto funcione es alta.

REFERENCIAS

- [1] Sanchez-Robles, J., Gomez De Leon-Hijes, F., Alarcon-Garcia, M., Martinez-Garcia, F., Belen-Rivera, E. (2021). Predictive maintenance for the efficient use of industrial agitators in distillers and reactors. DYN, 96(1). 17-21.
- [2] Hamza Sabir, Mohammed Ouassaid, Nabil Ngote, An experimental method for diagnostic of incipient broken rotor bar fault in induction machines. Heliyon, Vol. 8, Iss. 3, 2022
- [3] Selçuk ŞY, Ünal P, Albayrak Ö, Jomâa M. A Workflow for Synthetic Data Generation and Predictive Maintenance for Vibration Data. Information. 2021; 12(10):386.
- [4] Joerg Leukel, Julian González, Martin Riekert, Adoption of machine learning technology for failure prediction in industrial maintenance: A systematic review, Journal of Manufacturing Systems, Vol 61, 2021: 87-96
- [5] Bram de Jonge, Philip A. Scarf, A review on maintenance optimization, European Journal of Operational Research, Vol. 285, (2020): 805-824
- [6] ISO/TC 108/SC 5 (2003). ISO 13374-1:2003, Revision 2019, Condition monitoring and diagnostics of machine systems. Obtenido de: <https://www.iso.org/standard/21832.html>
- [7] Sarnecki JA, Witoś MI. Condition-based maintenance of turbojet engines based on compressor blade vibration measuring method. Transactions of the Institute of Fluid-Flow Machinery. 2003(114):133-46.
- [8] Hanachi H, Mechefske C, Liu J, Banerjee A, Chen Y. Performance-based gas turbine health monitoring, diagnostics, and prognostics: A survey. IEEE Transactions on Reliability. 2018 May 8;67(3):1340-63.
- [9] Tahan M, Tsoutsanis E, Muhammad M, Karim ZA. Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: A review. Applied energy. 2017 Jul 15;198:122-44.
- [10] Abdel-Hamid I. Mourad, Abdulla Almomani, Irfan Ahmad Sheikh, Ammar H. Elsheikh, Failure analysis of gas and wind turbine blades: A review, Engineering Failure Analysis, 2023(146)
- [11] Jingqi Li, Yulong Li, Micro gas turbine: Developments, applications, and key technologies on components, Propulsion and Power Research, 2023(12):1-43
- [12] Fentaye AD, Baheta AT, Gilani SI, Kyprianidis KG. A review on gas turbine gas-path diagnostics: State-of-the-art methods, challenges and opportunities. Aerospace. 2019 Jul;6(7):83.
- [13] Sangjo Kim, Ju Hyun Im, Myungho Kim, Junghoe Kim, You Il Kim, Diagnostics using a physics-based engine model in aero gas turbine engine verification tests, Aerospace Science and Technology, 2023(133)
- [14] Zaccaria V, Rahman M, Aslanidou I, Kyprianidis K. A review of information fusion methods for gas turbine diagnostics. Sustainability. 2019 Jan;11(22):6202.
- [15] Mujica LE, Rodellar J, Fernandez A, Güemes A. Q-statistic and T2-statistic PCA-based measures for damage assessment in structures. Structural Health Monitoring. 2011 Sep;10(5):539-53.
- [16] H. A. N. Søndergaard, H. R. Shaker and B. N. Jørgensen, Energy Systems Condition Monitoring: Dynamic Principal Component Analysis Application, 2022 IEEE 10th International Conference on Smart Energy Grid Engineering (SEGE), Canada, 2022: 81-87
- [17] Ge Z, Song Z. Multivariate statistical process control: Process monitoring methods and applications. Springer Science & Business Media; 2012 Nov 28.
- [18] Kai Zhang, Kaixiang Peng, Ruohui Chu, Jie Dong, Implementing Multivariate Statistics-Based Process Monitoring: A Comparison of Basic Data Modeling Approaches, Neurocomputing, 2018(209):172-184
- [19] D. Garcia-Alvarez, A. Bregon, B. Pulido, C.J. Alonso-Gonzalez, Integrating PCA and structural model decomposition to improve fault monitoring and diagnosis with varying operation points, Engineering Applications of Artificial Intelligence, 2023(122)
- [20] Klaus Nordhausen, Anne Ruiz-Gazen, On the usage of joint diagonalization in multivariate statistics, Journal of Multivariate Analysis, 2022(188)
- [21] Villarreal-Valderrama F, Liceaga-Castro E, Zambrano-Robledo P, Amézquita-Brooks L. Experimental Evaluation of Different Microturbojet EGT Modeling Approaches. Journal of Aerospace Engineering. 2021 Jan 1;34(1):04020087.
- [22] Villarreal-Valderrama F, Juárez-Pérez P, García-Pérez U, Amézquita-Brooks L. Applicability of correlational data-mining to small-scale turbojet performance map generation. International Journal of Turbo & Jet-Engines. 2021 Nov 23.