

# Revisión Sistemática: Anomalías en Series Temporales Univariantes

1<sup>st</sup> Jackeline Bermeo Pacheco  
*Universidad de Cuenca*  
*Maestría en Ciencia de Datos*  
Cuenca, Ecuador  
jackeline.bermeo@ucuenca.edu.ec

**Abstract**—Esta revisión sistemática describe el estado actual de la investigación relacionada con los métodos empleados para el análisis de anomalías en series temporales univariantes. Se abordan las siguientes dimensiones clave: detección, predicción, almacenamiento, comparación entre eventos, visualización y contextualización externa. El estudio examina si los enfoques identificados integran varias de estas dimensiones de forma conjunta o si se centran en dimensiones específicas. El proceso de revisión se llevó a cabo conforme a las directrices PRISMA 2020 y las guías metodológicas de Kitchenham, a partir de búsquedas realizadas en Google Scholar, arXiv, Semantic Scholar y EBSCO, con filtros por idioma, año de publicación (2018–2025) y disponibilidad de texto completo. Los resultados muestran que alrededor del 80% de los estudios revisados abordan la detección de anomalías, mientras que las demás dimensiones se encuentran representadas en proporciones variables. Este trabajo documenta las dimensiones tratadas en los enfoques existentes y proporciona una visión general de la literatura reciente sobre el tema.

**Index Terms**—detección de anomalías, predicción de anomalías, almacenamiento de anomalías, comparación de eventos anómalos, contextualización de anomalías, series temporales univariantes, revisión sistemática, framework.

## I. INTRODUCCIÓN

El análisis de anomalías en series temporales ha sido aplicado en una amplia variedad de dominios, incluyendo la climatología [1], la medicina [2], la predicción de epidemias [3], la industria manufacturera [4], y el monitoreo visual en sistemas aeroespaciales y financieros [5]. Esta diversidad evidencia la versatilidad de los enfoques desarrollados y su aplicabilidad en contextos operativos críticos.

En la literatura especializada se han identificado múltiples enfoques centrados predominantemente en la detección de anomalías. Sin embargo, otras funciones complementarias como el almacenamiento estructurado de eventos anómalos [6], la comparación entre eventos [7], la contextualización externa [8], lapredicción y la visualización han sido menos exploradas o abordadas de forma aislada.

La necesidad de incorporar contexto adicional ha sido señalada por Müller [8], quien propone su integración como

parte de un marco analítico más amplio, aunque sin implementación práctica. Por su parte, Blázquez-García et al. [7] subrayan que la ausencia de datos contextuales limita la interpretación de los eventos anómalos, sugiriendo su incorporación como línea futura de investigación.

Esta diversidad plantea interrogantes sobre si dichas dimensiones han sido abordadas conjuntamente o si los enfoques actuales continúan tratándolas de manera fragmentada. Diversos trabajos ilustran esta dispersión: Braei y Wagner [1] revisan métodos de detección en series univariantes; Wolny [9] propone una metodología multicriterio sin aprendizaje profundo; Boniol et al. [10] destacan limitaciones como la escalabilidad e interpretabilidad; Lai et al. [4] exploran la transferencia de conocimiento entre dominios; mientras que Natha et al. [11] presentan una revisión sistemática centrada en técnicas de detección.

Ante esta situación, el presente estudio plantea una revisión sistemática que permita analizar de forma estructurada las dimensiones de detección, predicción, almacenamiento, comparación, contextualización y visualización de anomalías en series temporales univariantes, así como su distribución metodológica y temática. Se busca identificar si estas dimensiones han sido integradas en propuestas completas o si, por el contrario, se abordan de manera aislada en investigaciones independientes.

### Preguntas de investigación:

- ¿Qué técnicas han sido propuestas para la detección, predicción, almacenamiento, comparación, contextualización y visualización de anomalías en series temporales univariantes?
- ¿Qué características presentan los enfoques revisados en relación con la cobertura de estas dimensiones funcionales?
- ¿Qué tipos de datos y técnicas se emplean comúnmente en los estudios seleccionados?

### Objetivos específicos:

- Identificar y recopilar literatura relevante publicada entre 2018 y 2025 sobre análisis de anomalías en series

temporales univariantes.

- Describir las dimensiones funcionales abordadas por cada estudio y destacar aquellas con menor representación o integración.

La revisión fue desarrollada siguiendo las directrices PRISMA 2020 y las recomendaciones metodológicas de Kitchenham [12], permitiendo garantizar la transparencia, trazabilidad y rigurosidad del proceso de selección y análisis.

## II. METODOLOGÍA

La metodología fue desarrollada siguiendo las directrices PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). El objetivo fue identificar, seleccionar y analizar estudios que abordaran la detección, predicción, almacenamiento, comparación y contextualización de anomalías en series temporales univariantes.

Se llevaron a cabo búsquedas en cinco bases de datos: Google Scholar, Semantic Scholar, arXiv, CORE y EBSCO, considerando publicaciones entre 2018 y 2025. Se aplicaron criterios de inclusión y exclusión en distintas etapas del proceso, incluyendo revisión por título, resumen y lectura completa. La Figura 1 presenta el diagrama de flujo PRISMA que resume este procedimiento.

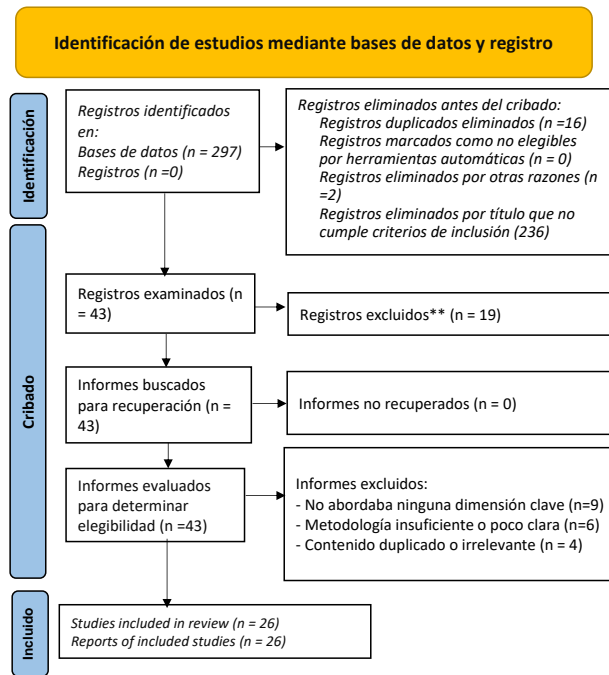


Fig. 1. Diagrama de flujo PRISMA 2020 del proceso de selección de estudios.

### A. Fuentes de búsqueda y estrategia

Las búsquedas se realizaron en Google Scholar, arXiv, CORE, Semantic Scholar y EBSCO. Inicialmente se aplicaron combinaciones amplias de términos como *"univariate time series"* y *"anomaly detection"*, lo que generó un alto volumen de resultados, especialmente en Google Scholar y

Semantic Scholar. Posteriormente, se refinaron las cadenas de búsqueda mediante operadores lógicos como AND y filtros avanzados como *allintitle*, además de incorporar términos específicos vinculados a las dimensiones analizadas (*framework, comparison, storage, contextualization*). Este refinamiento permitió enfocar la búsqueda en estudios con mayor probabilidad de relevancia.

También se evaluó la utilidad de cada repositorio consultado. CORE fue descartado debido a la escasa pertinencia de los documentos recuperados. En los casos en que no se disponía de acceso completo a estudios potencialmente útiles, se buscaron versiones alternativas en otras plataformas.

### B. Criterios de inclusión y exclusión

Se incluyeron estudios que cumplieran con los siguientes criterios:

- Abordaban una o más de las siguientes dimensiones: detección, predicción, almacenamiento, comparación o contextualización de anomalías en series temporales.
- Fueron publicados entre los años 2018 y 2025.
- Estaban redactados en idioma inglés y disponibles con acceso al texto completo.

Se excluyeron aquellos estudios que:

- No presentaban una descripción metodológica o resultados suficientes para el análisis.
- Correspondían a documentos de carácter divulgativo, como tutoriales, blogs o publicaciones sin revisión por pares.

### C. Proceso de selección

El proceso de análisis se desarrolló de manera progresiva y estructurada a lo largo de varias etapas, combinando revisión manual, filtrado sistemático y apoyo complementario mediante herramientas de inteligencia artificial.

En la fase inicial, la búsqueda se realizó explorando cada base digital de manera independiente. La primera selección se efectuó a partir del título de los documentos, con el objetivo de descartar aquellos no alineados con los objetivos del estudio. En los casos en que se identificaban estudios potencialmente relevantes pero no se disponía de acceso completo, se procedió a localizar versiones alternativas del texto completo en otros repositorios, ya sea en su versión publicada o en forma de preprint. Una vez garantizado el acceso al contenido, se llevó a cabo una depuración de duplicados. Posteriormente, se revisaron los resúmenes (abstracts), y finalmente se analizó el contenido completo de los estudios preseleccionados.

Durante la etapa final, se incorporó el uso de inteligencia artificial (ChatGPT) como herramienta auxiliar para facilitar la lectura y comprensión de artículos con redacción ambigua o estructuras metodológicas poco claras.

A lo largo del proceso se revisó un conjunto amplio de publicaciones provenientes de fuentes académicas. Si bien muchos estudios fueron descartados por no cumplir con los criterios de forma o fondo establecidos, aproximadamente 43 pasaron a la etapa de revisión completa por su pertinencia temática y metodológica. De estos, se seleccionaron 26 estudios que

abordan al menos una de las dimensiones del análisis de anomalías desde una perspectiva metodológica o conceptual significativa. Los 17 estudios restantes fueron excluidos por no presentar evidencia suficiente, carecer de implementación técnica o abordar únicamente contextos multivariantes sin posibilidad de extrapolación metodológica a series univariantes.

#### D. Extracción y análisis de datos

Para cada uno de los estudios incluidos en el análisis final, se elaboraron fichas individuales con información estructurada. Estas fichas recopilaban datos básicos como el título, autores, año de publicación, DOI, resumen y una síntesis de los puntos clave del enfoque propuesto.

Posteriormente, se definieron tres niveles de análisis para facilitar la comparación sistemática entre los estudios:

- **Características metodológicas generales:** se registró si el documento tenía una naturaleza teórica o práctica, si empleaba datos reales, validaba sus resultados en entornos reales, contaba con revisión por pares y ofrecía acceso al código fuente.
- **Dimensiones abordadas:** se codificó la presencia o ausencia de siete dimensiones: detección, predicción, almacenamiento, comparación, contextualización mediante datos externos, visualización y propuesta de framework.
- **Perfil del estudio:** se identificó el año del estudio y el tipo de estudio (revisión, conceptual, aplicado, etc.).

La información extraída fue incorporada en fichas informativas individuales para cada estudio incluido en la revisión. A partir de estas fichas, se elaboró un documento consolidado que se encuentra disponible como material complementario en el siguiente enlace: <https://github.com/jackelinebermeo/revisionSistemicaAnalisisAnomalias/blob/main/FichasRevisionSistemica%20Analisis%20Anomalias.pdf>

Esta estructura de codificación permitió establecer patrones, contrastes y vacíos entre los estudios revisados, proporcionando la base para los resultados presentados en la siguiente sección.

### III. RESULTADOS

Tras aplicar el protocolo de búsqueda, filtrado y selección, se identificaron 26 estudios que cumplieron con los criterios de inclusión. Los resultados del análisis sistemático se presentan a continuación mediante tablas de síntesis, organizadas según dimensiones metodológicas, temáticas y estructurales, con el fin de facilitar la comparación entre enfoques evaluados.

#### A. Evaluación de criterios metodológicos en los estudios seleccionados

Los estudios analizados fueron clasificados en dos grandes categorías metodológicas: *estudios teóricos o de revisión* (n = 12) y *estudios prácticos o experimentales* (n = 14). Esta distinción permite una evaluación diferenciada, ya que criterios como el uso de datos reales o la disponibilidad de código

no son aplicables ni esperables en investigaciones de carácter exclusivamente conceptual o bibliográfico.

En el caso de los estudios teóricos, el 100% cuenta con revisión por pares, pero ninguno realiza validaciones empíricas o pruebas con datos reales, tal como ocurre en Foorthuis [13], Chalapathy y Chawla [2], o Blázquez-García et al. [7]. Este grupo incluye revisiones sistemáticas, revisiones conceptuales y análisis de desafíos metodológicos, como en Müller [8] y Zamanzadeh et al. [14].

TABLA I  
EVALUACIÓN METODOLÓGICA DE ESTUDIOS PRÁCTICOS  
(NRO = 14)

Criterios	No	Parcial	Sí
Uso de datos reales	1	0	13
Validación en entornos reales	10	4	0
Revisión por pares	2	0	12
Código disponible	5	1	8

Fuente: elaboración propia.

La Tabla I resume la evaluación metodológica de los 14 estudios clasificados como prácticos, considerando cuatro criterios: uso de datos reales, validación en entornos reales, revisión por pares y disponibilidad de código. Cada criterio fue clasificado según tres niveles de cumplimiento: No, Parcial y Sí.

Los siguientes análisis se centran exclusivamente en los estudios de naturaleza práctica. En cuanto al uso de datos reales, se observó que 13 estudios emplearon conjuntos empíricos obtenidos de sensores, registros operativos o bases de datos reales. Por ejemplo, Carmona et al. [15] trabajaron con series temporales industriales como SMAP, SMD, MSL y KPI, mientras que Park et al. [16] utilizaron conjuntos como SWaT, UCR y NAB, enfocados en datos con ruido estructural. Alnegheimish et al. [6], por su parte, integraron señales multivariantes de radar y telemetría.

En contraste, la validación en entornos reales presenta una cobertura significativamente más limitada. No se identificaron estudios validados en contextos productivos; únicamente cuatro investigaciones incluyeron algún tipo de validación en entornos controlados o simulados, razón por la cual fueron clasificadas como "Parcial".

En cuanto a la revisión por pares, 12 de los 14 estudios fueron publicados en congresos o revistas académicas con procesos de revisión formal, mientras que dos casos carecen de esta garantía de calidad metodológica.

En relación a la disponibilidad de código, ocho estudios incluyeron repositorios accesibles y funcionales, uno proporcionó documentación o fragmentos parciales, y cinco no ofrecieron ningún tipo de implementación reproducible.

#### B. Características generales de los estudios

Abordando nuevamente los 26 estudios, estos muestran una diversidad significativa en cuanto a su orientación metodológica y nivel de aplicación. Para su caracterización se consideraron tres atributos principales: el *tipo de estudio* (como revisiones, análisis conceptual o estudios aplicados),

su *naturaleza* (teórica o práctica), y la *presencia de un framework*. La Tabla II presenta un resumen de esta clasificación, proporcionando una visión general del enfoque adoptado en cada caso y permitiendo identificar cuántos estudios incorporan marcos estructurados en su planteamiento.

TABLA II  
ESTUDIOS POR TIPO, NATURALEZA Y PRESENCIA DE FRAMEWORKS  
(NRO = 26)

Tipo de estudio	Naturaleza	Nro Total	FW
Análisis conceptual	Teórico	1	0
Estudio aplicado	Práctico	9	8
Estudio de investigación técnica	Práctico	1	1
Estudio técnico experimental	Práctico	1	1
Framework aplicado	Práctico	3	3
Revisión conceptual	Teórico	1	1
Revisión no sistemática	Teórico	3	0
Revisión sistemática	Teórico	7	0

Nota: "Nro" indica el número de estudios en cada categoría. "FW" representa cuántos de ellos incluyen un framework explícito.

Fuente: elaboración propia.

*Tipo de estudio:* Los estudios incluidos en la revisión se clasificaron según su tipo general, considerando su propósito y diseño metodológico. Se identificaron cuatro categorías principales: revisiones sistemáticas, revisiones no sistemáticas, revisiones o análisis conceptuales, y estudios prácticos (aplicados, técnicos o con frameworks operativos).

En total, doce estudios se clasifican como teóricos. Siete de ellos corresponden a revisiones sistemáticas que examinan el estado del arte en torno a la detección, predicción o caracterización de anomalías. Por ejemplo, Natha et al. [11] realizan una síntesis de técnicas de aprendizaje automático aplicadas a series temporales, mientras que Blázquez-García et al. [7] identifican brechas metodológicas en la literatura reciente. Tres estudios se agrupan como revisiones no sistemáticas, como la de Braei y Wagner [1], que comparan empíricamente múltiples enfoques de detección de anomalías sin seguir una metodología. La revisión conceptual de Foorthuis [13] propone una taxonomía estructurada de anomalías basada en criterios teóricos, y el análisis conceptual de Müller [8] desarrolla una arquitectura de referencia para el análisis y contextualización de anomalías en series temporales.

Los catorce estudios restantes se clasificaron como de tipo práctico. Nueve de ellos corresponden a estudios aplicados que evalúan técnicas de detección o predicción de anomalías sobre conjuntos de datos reales, utilizando métricas empíricas para validar su efectividad. Por ejemplo, Calikus et al. (2025) prueban modelos de deep learning sobre datos industriales, mientras que Geiger et al. (2020) evalúan la robustez de detectores ante distintos niveles de ruido en señales univariantes. Tres estudios presentan frameworks operativos con validación experimental, como el de Alnegheimish et al. [6], que integra componentes para visualización, detección y análisis, o el de Lai et al. (2023), que propone un sistema modular para detección contextual en IoT. Finalmente, dos estudios técnicos Wolny [9] y Doshi et al. [17] comparan enfoques mediante implementación directa en entornos simulados: el primero explora variaciones en métodos estadísticos de base, mientras

que el segundo introduce un modelo autoatentivo evaluado en datasets sintéticos y reales bajo condiciones controladas.

Esta clasificación permite distinguir entre investigaciones orientadas a la organización conceptual del conocimiento y aquellas centradas en el desarrollo, evaluación o implementación de soluciones prácticas.

*1) Distribución temporal de los estudios:* La Tabla III presenta la distribución por año de los estudios incluidos en esta revisión. Se identifican 17 estudios publicados entre 2022 y 2024. En particular, el año 2022 agrupa 7 estudios, mientras que 2023 y 2024 reúnen 3 y 7 estudios, respectivamente.

Entre las publicaciones previas a 2022, se encuentran los estudios de Braei et al. [1], Blázquez-García et al. [7], y Geiger et al. [18], que proponen distintos enfoques basados en reconstrucción para la detección de anomalías. Además, se incluyen investigaciones como la de Tealab [19], centrada en métodos híbridos, y el trabajo de Chalapathy y Chawla [2], que discute marcos de referencia para el aprendizaje profundo en este contexto. En conjunto, los años entre 2018 y 2021 concentran 7 estudios, proporcionando antecedentes metodológicos relevantes para el análisis actual.

TABLA III  
DISTRIBUCIÓN POR AÑO DE PUBLICACIÓN

Año	Total de estudios
2018	1
2019	1
2020	3
2021	2
2022	7
2023	3
2024	7
2025	2
<b>Total</b>	<b>26</b>

Fuente: elaboración propia.

### C. Cobertura de dimensiones funcionales

TABLA IV  
RESUMEN DE COBERTURA DE DIMENSIONES

Criterios	No	Parcial	Sí
Detección de anomalías	4	2	20
Predicción	9	13	4
Almacenamiento	23	2	1
Comparación	21	4	1
Considera contexto	17	9	0
Visualización	24	0	2

Fuente: elaboración propia.

El presente estudio evaluó seis dimensiones clave para caracterizar el grado de cobertura funcional en los enfoques de análisis de anomalías en series temporales univariantes: detección, predicción, almacenamiento, comparación, contextualización externa y visualización.

Los resultados muestran una amplia cobertura en la dimensión de detección, presente total en 20 de los 26 estudios, los cuales implementan de manera explícita mecanismos de detección de anomalías. Por ejemplo, Lai et al. [4], Carmona et al. [15], Alnegheimish et al. [6] y Doshi et al. [17] utilizan

modelos basados en aprendizaje profundo, como LSTM o Autoencoders. Otros estudios, como Foorthuis [13] o Blázquez-García et al. [7], exploran enfoques estadísticos o híbridos.

En contraste, la dimensión de predicción fue abordada de forma completa en solo cuatro estudios, mientras que trece la consideraron parcialmente. Por ejemplo, Tealab [19] y Alnegheimish et al. [6] discuten la incorporación de mecanismos de predicción como parte del proceso de análisis. En otros casos, como Chalapathy y Chawla [2] o Geiger et al. [18], se alude a su potencial integración sin implementaciones concretas.

La dimensión de almacenamiento fue abordada por Alnegheimish et al. [6] quienes implementan una base de datos documental (MongoDB) para registrar y gestionar eventos anómalos. Además, otros dos estudios (Müller [8] y Liu et al. [5]) hacen referencia a esta dimensión de forma parcial, ya sea desde una perspectiva arquitectónica general o como soporte técnico para la consulta y organización de resultados.

La comparación, entendida como el análisis de similitudes o recurrencias entre eventos anómalos dentro de una misma serie temporal, fue abordada por Liu et al. [5], que incorpora mecanismos explícitos para identificar y contrastar eventos anómalos mediante análisis. Cuatro trabajos adicionales fueron clasificados como “Parcial” por incluir referencias conceptuales o mecanismos limitados. Por ejemplo, Blázquez-García et al. [7] y Boniol y Palpanas [20] presentan esquemas de visualización que podrían facilitar la exploración manual de patrones repetitivos. Müller [8] y Wolny [9] discuten la comparación como componente relevante, aunque sin mecanismos formales o métricas específicas.

La contextualización externa, definida como la incorporación de información proveniente del entorno real (por ejemplo, datos climáticos, económicos o sanitarios) para enriquecer el análisis de anomalías, no fue implementada plenamente en ninguno de los estudios analizados. Tampoco se identificaron propuestas metodológicas que desarrollen su integración de manera operativa. No obstante, nueve estudios fueron clasificados como “Parcial” por incluir mecanismos auxiliares o datos generados en el entorno del experimento. Entre estos, Carmona et al. [15] utilizan etiquetas sintéticas asociadas a eventos, Natha et al. [11] emplean representaciones embebidas, y Ronghui Xu et al. [21] incorpora metadatos sobre condiciones experimentales (como configuración de sensores y temporalidad), lo cual proporciona cierto contexto operativo, aunque no derivado de fuentes externas reales.

En cuanto a la visualización, solo dos estudios fueron clasificados como “Sí” por incorporar entornos gráficos que facilitan la exploración detallada de eventos anómalos. Alnegheimish et al. [6] presentan una interfaz interactiva para el análisis visual, y Liu et al. [5] integran componentes visuales que permiten representar la evolución temporal y detectar patrones atípicos. En relación con esta dimensión, algunos estudios como Boniol et al. [10] y Leites et al. [22], destacan su importancia metodológica para mejorar la interpretabilidad y la trazabilidad del análisis.

## IV. DISCUSIÓN

### A. Síntesis crítica de hallazgos

Los resultados evidencian una alta concentración en la dimensión de detección, en contraste la predicción, el almacenamiento, la comparación, la contextualización externa y la visualización presentan una cobertura considerablemente menor. Esta distribución sugiere una orientación predominante hacia la identificación de anomalías, sin una articulación sistemática con mecanismos destinados a su análisis posterior, contextualización o gestión operativa.

En particular, la predicción se incorpora en varios trabajos como componente adicional, aunque en la mayoría de los casos no se implementa de forma integral. En relación con el almacenamiento, solo un estudio contempla su desarrollo estructurado, mientras que otros lo abordan de manera parcial, como parte del diseño general. La comparación y la contextualización externa se abordan en términos conceptuales en algunos estudios, pero no incluyen procedimientos técnicos ni integración de fuentes externas de datos. De forma similar, la visualización es limitada: cuando está presente, se restringe a funciones básicas de representación gráfica, sin componentes interactivos ni capacidades analíticas extendidas.

Esta distribución asimétrica en la cobertura funcional indica que los enfoques actuales tienden a abordar dimensiones específicas de manera aislada, sin configurar arquitecturas integradas que permitan el tratamiento completo de eventos anómalos. En conjunto, los hallazgos muestran una fragmentación metodológica. Aunque la detección ha sido ampliamente desarrollada, las demás dimensiones se abordan de forma parcial o están ausentes.

### B. Limitaciones metodológicas comunes

Desde una perspectiva metodológica, los estudios analizados presentan limitaciones en cuanto a su aplicabilidad, reproducibilidad y alcance. Aunque 13 de los estudios prácticos emplearon datos reales para sus experimentos, ninguno fue validado en entornos productivos. Solo uno de ellos, Lai et al. [4], reporta una validación en un entorno parcialmente controlado. En el caso de Park et al. [16], aunque se utilizan conjuntos de datos complejos con ruido estructural, no se detalla una validación contextual específica, por lo que no puede considerarse como validación ambiental plena.

En términos de transparencia y replicabilidad, 8 de los 14 estudios prácticos informan la disponibilidad pública de código o herramientas para reproducir los experimentos.

Algunos autores reconocen limitaciones estructurales en sus enfoques. Müller [8] y Wolny [9] advierten sobre la fragmentación funcional, al centrarse exclusivamente en dimensiones como la detección o la visualización, sin incorporar mecanismos complementarios como almacenamiento, contextualización o comparación. Por su parte, Alnegheimish et al. [6] señalan desafíos en la escalabilidad y en la capacidad de sus modelos para adaptarse a dominios distintos a los originalmente considerados.

Esta revisión identificó tres vacíos principales:

- **Enfoque exclusivo en detección:** 20 de los 26 estudios se enfocan en la detección de anomalías, sin articular dimensiones complementarias como predicción, comparación, almacenamiento o contextualización. Ejemplos representativos de esta orientación son: Xu et al. [23], cuyo modelo Anomaly Transformer emplea mecanismos de atención para detectar patrones anómalos, sin extender su funcionalidad a otras etapas del ciclo de análisis; Tuli et al. [24], con TranAD, una arquitectura GAN enfocada en detección no supervisada; y Park et al. [16], quien utiliza un enfoque jerárquico de aprendizaje profundo restringido a la etapa de detección. Aunque Doshi et al. [17] incorporan puntuaciones secuenciales como salida del modelo, no desarrollan mecanismos explícitos de comparación, contextualización o almacenamiento.
- **Escasa integración de dimensiones complementarias:** Ninguno de los estudios revisados implementa de forma completa funciones adicionales como la contextualización externa, la comparación, el almacenamiento o la visualización. Tres estudios Alnegheimish et al. [6], Müller [8], y Lai et al. [4] discuten explícitamente la importancia del contexto externo, aunque sin integrarlo operativamente. Otros siete trabajos, como los de Carmona et al. [15], Natha et al. [11] y Calikus et al. [25], incluyen mecanismos indirectos o datos auxiliares simulados. Alnegheimish et al. [6] y Wu Sun et al. [26] son los únicos que incorporan componentes de almacenamiento formal o herramientas visuales interactivas. Finalmente, aunque algunos trabajos como Müller [8] o Boniol et al. [10] abordan conceptualmente la comparación entre eventos, no presentan métodos técnicos ni métricas específicas que permitan su implementación sistemática.
- **Ausencia de enfoques integradores:** Ninguno de los estudios revisados contempla de forma simultánea e integrada las seis dimensiones funcionales analizadas (detección, predicción, comparación, almacenamiento, contextualización externa y visualización). Esta fragmentación metodológica limita la trazabilidad de los eventos anómalos y reduce el potencial de los sistemas para ofrecer análisis completos, auditables y adaptables. La falta de marcos que integren múltiples funciones refuerza la necesidad de avanzar hacia propuestas holísticas en el tratamiento de anomalías en series temporales univariantes.

Cabe señalar que esta revisión fue realizada por una única persona, lo que puede introducir limitaciones en términos de sesgos en la selección, interpretación o clasificación de los estudios. Aunque se aplicaron criterios sistemáticos y se documentaron las decisiones adoptadas, no puede descartarse la influencia de apreciaciones individuales en algunos juicios analíticos.

Aunque la mayoría de los estudios analizados se concentra en una dimensión específica, se identificaron esfuerzos intermedios que apuntan hacia una integración parcial de funciones. Alnegheimish et al. [6] destaca por proponer una arquitectura modular que incluye visualización, almacenamiento estructurado y mecanismos de detección, además de una interfaz interactiva que favorece la interpretación de resultados. De forma complementaria, Wu y Sun [26] implementan un sistema de metadatos asociado a cada anomalía detectada, lo que permite su trazabilidad temporal, aunque sin vincular estos datos a un contexto operativo externo.

Braei y Wagner [1] proponen un enfoque estadístico para la detección de anomalías basado en patrones de recurrencia, que permite comparar subsecuencias dentro de una misma serie temporal. Este análisis intra-serie contribuye a una mejor comprensión de comportamientos repetitivos y podría ser útil en investigaciones orientadas en el desarrollo de mecanismos comparativos en sistemas de alerta temprana, aunque dicha aplicación no se aborda explícitamente en el estudio.

En relación con la contextualización externa, se identificó una ausencia generalizada de integración operativa. Por esta razón, se aplicaron criterios flexibles en la codificación. Fueron clasificados como “Parcial” aquellos estudios que, sin incorporar directamente variables del entorno, utilizan etiquetas sintéticas, embeddings derivados de la serie o referencias explícitas al valor del contexto. Carmona et al. [15] y Xu et al. [21] representan este grupo: el primero mediante la simulación de escenarios etiquetados, y el segundo mediante el uso de metadatos auxiliares relacionados con el proceso de captura.

En cuanto a la propuesta de frameworks estructurados, se identificaron siete estudios con arquitecturas organizadas que incluyen al menos dos dimensiones clave. Lai et al. [4], por ejemplo, plantea un sistema de adaptación contextual para mejorar la transferencia entre dominios, mientras que Park et al. [16] combina transformaciones visuales con aprendizaje profundo jerárquico. Sin embargo, en todos los casos estas estructuras se enfocan en una combinación reducida de funciones, sin abordar dimensiones como almacenamiento persistente o comparación.

Desde el plano teórico, algunas revisiones ofrecen modelos clasificatorios útiles para comprender la evolución del campo. Chalapathy y Chawla [2] organizan los enfoques de deep learning según su arquitectura y técnica de entrenamiento. Zamanzadeh et al. [14] amplían esta taxonomía incluyendo modelos recientes y señalando retos técnicos como la dependencia de hiperparámetros, la dificultad de generalización y la baja capacidad explicativa de ciertas redes neuronales.

Además, ciertos trabajos, aunque centrados en datos multivariantes, presentan hallazgos metodológicamente transferibles. Suboh et al. [27] subrayan la necesidad de métricas robustas para evaluar la detección de anomalías en entornos de alta dimensionalidad, un reto que también se manifiesta en escenarios univariantes. Boniol et al. [10], por su parte,

desarrollan una taxonomía centrada en el ciclo de análisis (desde la recolección hasta la interpretación), proporcionando un marco útil para evaluar la cobertura funcional de los sistemas, incluso en el caso de datos univariantes.

Estos aportes, aunque dispersos, ofrecen pistas valiosas para el diseño de enfoques más integrados, con una mejor articulación entre componentes funcionales y una mayor aplicabilidad en contextos reales.

#### *E. Implicaciones para futuras investigaciones*

Esta revisión evidenció una concentración funcional en la detección de anomalías, con escasa integración de dimensiones como almacenamiento, comparación o contextualización. Esto sugiere un campo aún fragmentado, donde las soluciones tienden a abordar tareas de forma aislada.

Futuras investigaciones podrían centrarse en diseñar marcos integrados que combinen múltiples funciones, permitiendo trazabilidad, interpretación y reutilización de eventos anómalos. En particular, se requiere avanzar en:

- **Comparación:** diseñar métodos que permitan comparar eventos anómalos dentro de la misma serie temporal, identificando patrones repetitivos, similitudes estructurales o recurrencias a lo largo del tiempo. Esto facilitaría el agrupamiento de anomalías con características comunes y el análisis de su evolución.
- **Contextualización externa:** incorporar datos del entorno (como clima, condiciones operativas o variables externas relevantes) para interpretar mejor las causas o consecuencias de las anomalías.
- **Almacenamiento estructurado:** implementar mecanismos para registrar y gestionar eventos anómalos de forma persistente, permitiendo su trazabilidad, auditoría y reutilización en análisis posteriores o sistemas predictivos.

Además, resulta necesario explorar marcos modulares y reutilizables que combinen técnicas estadísticas y de aprendizaje profundo, superando enfoques cerrados centrados solo en la detección. Esto permitiría desarrollar soluciones más completas, adaptables y útiles en dominios críticos como la industria, la medicina o el monitoreo ambiental.

### V. CONCLUSIONES

Esta revisión sistemática analizó 26 estudios publicados entre 2018 y 2025. Los resultados evidencian una cobertura desigual entre las distintas dimensiones analizadas. Si bien la detección de anomalías está presente en la mayoría de los enfoques, otras funciones como la predicción, el almacenamiento, la comparación, la contextualización externa y la visualización aparecen con menor frecuencia o se abordan de manera parcial.

Existe una brecha en la escasa conexión con entornos reales, no se identificó un estudio que incluya la validación operativa en contextos productivos, y solo algunos evalúan sus propuestas en escenarios parcialmente controlados. La mayoría de modelos se prueban exclusivamente sobre datasets públicos o sintéticos, lo cual reduce su capacidad de generalización y su utilidad en situaciones reales donde las condiciones son más complejas e inciertas.

Otro hallazgo crítico es la discrepancia entre lo que se declara y lo que se implementa. Varias dimensiones como la comparación de anomalías o la contextualización externa son mencionadas como relevantes, pero no se abordan con mecanismos técnicos concretos.

Pese a estas limitaciones, algunos estudios muestran avances parciales hacia arquitecturas más integradas. Casos como los de Alnegheimish et al., Liu et al. o Lai et al. proponen marcos que incorporan múltiples funciones, aunque centrados mayormente en visualización y detección. Asimismo, estudios como el de Braei y Wagner exploran mecanismos de comparación a través de análisis estadísticos, y Carmona et al. implementan estrategias indirectas de contextualización mediante simulaciones etiquetadas.

Con base en estos hallazgos, se identifican implicaciones importantes para futuras investigaciones:

- Se requiere avanzar hacia marcos de análisis integradores, capaces de enlazar detección con comparación, almacenamiento, contextualización y visualización, bajo una lógica funcional coherente.
- Es fundamental fortalecer la validación empírica, mediante pruebas en entornos reales o simulaciones representativas del contexto operativo.
- La comparación de eventos anómalos dentro de una misma serie debe formalizarse mediante métricas o estructuras que permitan analizar recurrencias o patrones persistentes.
- La contextualización debe superar su condición declarativa y convertirse en un componente activo, incorporando variables externas o condiciones ambientales que permitan entender mejor las causas o implicaciones de los eventos detectados.
- Finalmente, se debe fomentar la ciencia abierta y la reproducibilidad, mediante la publicación de datos, código y documentación técnica que permita evaluar y extender los enfoques existentes.

En conjunto, esta revisión aporta una visión crítica y estructurada del estado actual del campo. No solo documenta los avances en la detección de anomalías, sino que visibiliza las carencias en cuanto a integración funcional, validación contextual y articulación metodológica. Este panorama puede servir como hoja de ruta para orientar la investigación futura hacia soluciones más robustas, explicables y alineadas con los desafíos reales que plantea el análisis de series temporales en entornos dinámicos y sensibles.

*Respecto a las preguntas de investigación, esta revisión permitió identificar las técnicas más utilizadas, evidenciar las limitaciones actuales en cobertura funcional y operativa, y caracterizar el perfil metodológico de los estudios. Los resultados ofrecen una base empírica clara que puede guiar el desarrollo de sistemas más completos y efectivos.*

### REFERENCES

- [1] M. Braei and S. Wagner, "Anomaly Detection in Univariate Time-series: A Survey on the State-of-the-Art," Apr. 2020, arXiv:2004.00433 [cs]. [Online]. Available: <http://arxiv.org/abs/2004.00433>

- [2] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," *arXiv preprint arXiv:1901.03407*, 2019, arXiv:1901.03407 [cs.LG]. [Online]. Available: <https://arxiv.org/abs/1901.03407>
- [3] "A systematic literature review of time series methods applied to epidemic prediction," *Informatics in Medicine Unlocked*, vol. 50, p. 101571, 2024. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2352914824001278>
- [4] K.-H. Lai, L. Wang, H. Chen, K. Zhou, F. Wang, H. Yang, and X. Hu, "Context-aware Domain Adaptation for Time Series Anomaly Detection," Apr. 2023, arXiv:2304.07453 [cs]. [Online]. Available: <http://arxiv.org/abs/2304.07453>
- [5] D. Liu, S. Alnegheimish, A. ZYTEK, and K. Veeramachaneni, "MTV: Visual Analytics for Detecting, Investigating, and Annotating Anomalies in Multivariate Time Series," Dec. 2021, arXiv:2112.05734 [cs]. [Online]. Available: <http://arxiv.org/abs/2112.05734>
- [6] S. Alnegheimish, D. Liu, C. Sala, L. Berti-Equille, and K. Veeramachaneni, "Sintel: A Machine Learning Framework to Extract Insights from Signals," in *Proceedings of the 2022 International Conference on Management of Data*, Jun. 2022, pp. 1855–1865, arXiv:2204.09108 [cs]. [Online]. Available: <http://arxiv.org/abs/2204.09108>
- [7] A. Blázquez-García, A. Conde, U. Mori, and J. A. Lozano, "A review on outlier/anomaly detection in time series data," Feb. 2020, arXiv:2002.04236 [cs]. [Online]. Available: <http://arxiv.org/abs/2002.04236>
- [8] A. Mueller, "Open Challenges in Time Series Anomaly Detection: An Industry Perspective," Feb. 2025, arXiv:2502.05392 [cs]. [Online]. Available: <http://arxiv.org/abs/2502.05392>
- [9] Silesian University of Technology, Faculty of Organization and Management, Department of Economics and Computer Science and M. Wolny, "ANOMALY DETECTION IN UNIVARIATE TIME SERIES USING A MULTI-CRITERIA APPROACH," *Scientific Papers of Silesian University of Technology. Organization and Management Series*, vol. 2024, no. 213, pp. 665–680, 2024. [Online]. Available: <https://managementpapers.polsl.pl/wp-content/uploads/2025/02/213-Wolny.pdf>
- [10] P. Boniol, Q. Liu, M. Huang, T. Palpanas, and J. Paparrizos, "Dive into Time-Series Anomaly Detection: A Decade Review," Dec. 2024, arXiv:2412.20512 [cs]. [Online]. Available: <http://arxiv.org/abs/2412.20512>
- [11] S. Natha, "A Systematic Review of Anomaly detection using Machine and Deep Learning Techniques," *Quaid-e-Awam University Research Journal of Engineering, Science & Technology*, vol. 20, no. 1, pp. 83–94, Jun. 2022. [Online]. Available: <http://publications.quest.edu.pk/ojs/index.php/qjrj/article/view/66>
- [12] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Tech. Rep., 2007, version 2.3, July 2007. [Online]. Available: <https://www.researchgate.net/publication/302924724/Guidelines-for-performing-Systematic-Literature-Reviews-in-Software-Engineering>
- [13] R. Foorthuis, "On the nature and types of anomalies: a review of deviations in data," *International Journal of Data Science and Analytics*, vol. 12, no. 4, pp. 297–331, Oct. 2021. [Online]. Available: <https://link.springer.com/10.1007/s41060-021-00265-1>
- [14] Z. Z. Darban, G. I. Webb, S. Pan, C. C. Aggarwal, and M. Salehi, "Deep learning for time series anomaly detection: A survey," *ACM Computing Surveys*, vol. 56, no. 5, pp. 1–42, 2023, revisión sistemática de técnicas de aprendizaje profundo aplicadas a series temporales. [Online]. Available: <https://doi.org/10.1145/3691338>
- [15] C. U. Carmona, F.-X. Aubet, V. Flunkert, and J. Gasthaus, "Neural Contextual Anomaly Detection for Time Series," in *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*. Vienna, Austria: International Joint Conferences on Artificial Intelligence Organization, Jul. 2022, pp. 2843–2851. [Online]. Available: <https://www.ijcai.org/proceedings/2022/394>
- [16] H. Park and H. Jang, "Enhancing Time Series Anomaly Detection: A Knowledge Distillation Approach with Image Transformation," *Sensors*, vol. 24, no. 24, p. 8169, Dec. 2024. [Online]. Available: <https://www.mdpi.com/1424-8220/24/24/8169>
- [17] K. Doshi, S. Abudalou, and Y. Yilmaz, "Tisat: Time series anomaly transformer," *arXiv preprint arXiv:2203.05167*, 2022, arXiv:2203.05167 [cs.LG]. [Online]. Available: <https://arxiv.org/abs/2203.05167>
- [18] A. Geiger, D. Liu, S. Alnegheimish, A. Cuesta-Infante, and K. Veeramachaneni, "TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks," Nov. 2020, arXiv:2009.07769 [cs]. [Online]. Available: <http://arxiv.org/abs/2009.07769>
- [19] A. Tealab, "Time series forecasting using artificial neural networks methodologies: A systematic review," *Future Computing and Informatics Journal*, vol. 3, no. 2, pp. 334–340, Dec. 2018. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2314728817300715>
- [20] P. Boniol, J. Paparrizos, and T. Palpanas, "New Trends in Time Series Anomaly Detection."
- [21] R. Xu, H. Miao, S. Wang, P. S. Yu, and J. Wang, "PeFAD: A Parameter-Efficient Federated Framework for Time Series Anomaly Detection," Jul. 2024, arXiv:2406.02318 [cs]. [Online]. Available: <http://arxiv.org/abs/2406.02318>
- [22] J. Leites, V. Cerqueira, and C. Soares, "Lag Selection for Univariate Time Series Forecasting using Deep Learning: An Empirical Study," May 2024, arXiv:2405.11237 [stat]. [Online]. Available: <http://arxiv.org/abs/2405.11237>
- [23] J. Xu, H. Wu, J. Wang, and M. Long, "Anomaly Transformer: Time Series Anomaly Detection with Association Discrepancy," Jun. 2022, arXiv:2110.02642 [cs]. [Online]. Available: <http://arxiv.org/abs/2110.02642>
- [24] S. Tuli, G. Casale, and N. R. Jennings, "TranAD: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data," May 2022, arXiv:2201.07284 [cs]. [Online]. Available: <http://arxiv.org/abs/2201.07284>
- [25] E. Calikus, S. Nowaczyk, and O. Dikmen, "Context discovery for anomaly detection," *International Journal of Data Science and Analytics*, vol. 19, no. 1, pp. 99–113, Jan. 2025. [Online]. Available: <https://link.springer.com/10.1007/s41060-024-00586-x>
- [26] W. Sun, H. Li, Q. Liang, X. Zou, M. Chen, and Y. Wang, "On data efficiency of univariate time series anomaly detection models," *Journal of Big Data*, vol. 11, no. 1, p. 83, Jun. 2024. [Online]. Available: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-024-00940-7>
- [27] S. Suboh, I. A. Aziz, S. M. Shaharudin, S. A. Ismail, and H. Mahdin, "A Systematic Review of Anomaly Detection within High Dimensional and Multivariate Data," *JOIV : International Journal on Informatics Visualization*, vol. 7, no. 1, p. 122, Feb. 2023. [Online]. Available: <http://joiv.org/index.php/joiv/article/view/1297>