Universidad Politécnica de Madrid

Escuela Técnica Superior de Ingenieros Informáticos

Grado en Ingeniería Informática

Trabajo Fin de Grado

Monitorización y Análisis de Datos Operativos en Infraestructuras de Red Cuántica

Autor: Miguel López Ferreiro
Tutor: Alberto Juan Sebastián Lombraña

Madrid, Enero 2026

Este Trabajo Fin de Grado se ha depositado en la ETSI Informáticos de la Universidad Politécnica de Madrid para su defensa.

Trabajo Fin de Grado
Grado en Ingeniería Informática

Título: Monitorización y Análisis de Datos Operativos en Infraestructuras de Red Cuántica

Enero 2026

Autor: Miguel López Ferreiro
Tutor: Alberto Juan Sebastián Lombraña
Departamento de Lenguajes y Sistemas Informáticos
Escuela Técnica Superior de Ingenieros Informáticos
Universidad Politécnica de Madrid

Resumen

La Distribución Cuántica de Clave (Quantum Key Distribution, QKD) se presenta como una tecnología fundamental para garantizar la seguridad de las comunicaciones frente a las amenazas derivadas del avance de la computación cuántica. Aunque la seguridad de la QKD se apoya en principios físicos bien establecidos, su despliegue en infraestructuras reales plantea nuevos retos de carácter operativo, especialmente en lo relativo a la monitorización continua, la interpretación de métricas cuánticas y la detección temprana de degradaciones del sistema.

Este Trabajo Fin de Grado aborda el problema de la monitorización operativa de infraestructuras QKD desde una perspectiva ingenieril, centrada en el análisis de datos operativos reales procedentes de sistemas comerciales desplegados en red. El objetivo principal es diseñar y evaluar un pipeline reproducible de análisis y detección de anomalías capaz de transformar métricas cuánticas y clásicas (QBER, tasa de clave segura, entre otras) en información accionable para la operación de enlaces QKD, en ausencia de datos etiquetados de fallos o ataques.

Para ello, se analizan datasets reales de distintos fabricantes y se desarrolla un flujo de trabajo basado en la limpieza de datos, la ingeniería de características temporales y multivariantes, y el uso de técnicas de aprendizaje automático no supervisado, concretamente Isolation Forest y Autoencoders. Estos modelos se emplean como herramientas de apoyo al análisis operativo, priorizando la interpretabilidad, la robustez frente al ruido y la reducción de falsos positivos.

La evaluación del sistema se lleva a cabo mediante demostradores operativos controlados, diseñados para reproducir escenarios físicamente plausibles como degradaciones del canal cuántico, saturaciones del plano clásico, deriva lenta y perturbaciones estocásticas. Los resultados muestran que el sistema propuesto es capaz de detectar anomalías estructurales relevantes y de distinguirlas de variaciones benignas del sistema, evitando alarmas injustificadas en condiciones de operación normal.

El trabajo concluye que las técnicas de aprendizaje automático no supervisado, combinadas con conocimiento del dominio y un análisis cuidadoso de los datos, constituyen una herramienta viable para apoyar la monitorización operativa de infraestructuras QKD reales, y sientan las bases para futuras extensiones orientadas al mantenimiento predictivo y a su integración en arquitecturas de control de red cuántica.

Abstract

Quantum Key Distribution (QKD) is increasingly regarded as a key technology to ensure the long-term security of communications against the threats posed by quantum computing. While the security of QKD is grounded in well-established physical principles, its deployment in real-world infrastructures introduces relevant operational challenges, particularly with respect to continuous monitoring, the interpretation of quantum metrics, and the early detection of system degradations.

This Bachelor's Thesis addresses the problem of operational monitoring of QKD infrastructures from an engineering perspective, with a focus on the analysis of real operational data obtained from commercial QKD systems deployed in networked environments. The main objective is to design and evaluate a reproducible monitoring and anomaly detection pipeline capable of transforming quantum and classical operational metrics (such as QBER and secure key rate) into actionable information, in the absence of labeled data describing failures or attacks.

To achieve this goal, real datasets from different manufacturers are analyzed (in this Thesis QTI and Thoshiba machines are analyzed), and a data processing pipeline is developed based on data cleaning, temporal and multivariate feature engineering, and the application of unsupervised machine learning techniques, specifically Isolation Forest and Autoencoders. These models are used as decision-support tools, with an emphasis on interpretability, robustness to noise, and the reduction of false positives.

The system is evaluated using controlled operational demonstrators designed to reproduce physically plausible scenarios, including quantum channel degradation, classical plane saturation, slow drift effects, and stochastic instabilities. The results indicate that the proposed approach is capable of detecting relevant structural anomalies while avoiding unjustified alarms under normal operating conditions.

The thesis concludes that unsupervised learning techniques, when combined with domain knowledge and careful data analysis, constitute a viable solution for supporting the operational monitoring of real QKD infrastructures, and provide a solid foundation for future work on predictive maintenance and integration into quantum network control architectures.

Tabla de contenidos

1. Introducción

1.1. Contexto y motivación

La seguridad de las comunicaciones es actualmente un elemento central en el diseño, despliegue y operación de las infraestructuras digitales modernas. Sistemas tan diversos como las redes de telecomunicaciones, las infraestructuras financieras, los servicios gubernamentales o los entornos industriales críticos dependen de mecanismos criptográficos para garantizar la confidencialidad, la integridad y la autenticidad de la información intercambiada. En este contexto, la fiabilidad de los mecanismos de protección no constituye únicamente una cuestión teórica, sino un requisito operativo imprescindible para el funcionamiento continuo de esos servicios esenciales y para mantener la confianza en las infraestructuras digitales que los sustentan.

En el paradigma clásico, la mayoría de los sistemas criptográficos actualmente desplegados se apoyan en problemas matemáticos cuya seguridad se basa en la dificultad computacional de resolverlos en un tiempo razonable. Este enfoque ha demostrado su eficacia durante décadas y ha permitido el desarrollo de protocolos ampliamente utilizados, como las técnicas de intercambio y acuerdo de clave RSA, Diffie-Hellman o los esquemas basados en curvas elípticas. Sin embargo, la seguridad que proporcionan estos mecanismos es de naturaleza condicional, ya que depende del estado de la tecnología, de la capacidad de cálculo disponible y de la inexistencia de algoritmos más eficientes que los actualmente conocidos.

El avance progresivo de la computación cuántica introduce una amenaza estructural para este modelo. Algoritmos cuánticos bien conocidos, como el algoritmo de Shor, permiten resolver de forma eficiente problemas matemáticos que resultan intratables para los ordenadores actuales, llamados "clásicos". Aunque los ordenadores cuánticos actuales aún no disponen de la capacidad necesaria para comprometer sistemas criptográficos desplegados en producción, existe un amplio consenso en la comunidad científica y tecnológica en que este escenario es plausible a medio y largo plazo. Además, es habitual usar el argumento de que ya es posible guardar información protegida sensible que pueda ser desvelada en el futuro, lo que pone en riesgo también la información sensible en la actualidad. Esta previsión obliga a replantear las bases de la seguridad de las comunicaciones con suficiente antelación, especialmente en aplicaciones donde la confidencialidad de la información debe preservarse durante periodos prolongados de tiempo.

Como respuesta a esta amenaza, la criptografía post-cuántica (*post-quantum cryptography*, PQC) propone nuevos algoritmos diseñados específicamente para resistir ataques realizados con ordenadores cuánticos. No obstante, estas soluciones siguen basándose en supuestos computacionales cuya solidez a largo plazo aún debe validarse mediante un uso continuado y un análisis exhaustivo. En este sentido, la PQC puede entenderse como una evolución del modelo clásico, pero no como un cambio de paradigma en lo relativo a los fundamentos últimos de la seguridad.

Frente a este enfoque, la distribución cuántica de clave (*quantum key distri-*

bution, QKD) [1] plantea una alternativa conceptualmente distinta. En lugar de apoyarse en la dificultad de un problema matemático, la QKD permite establecer claves criptográficas con garantías de seguridad fundamentadas directamente en principios físicos, concretamente en las leyes de la mecánica cuántica. Esta característica convierte a la QKD en una tecnología especialmente atractiva para aplicaciones en las que la seguridad a largo plazo es crítica y donde la dependencia de supuestos computacionales se considera un riesgo relevante.

1.2. Fundamentos físicos y métricas operativas en QKD

La QKD se fundamenta en propiedades esenciales de la mecánica cuántica, como el principio de incertidumbre y el teorema de no clonación. Estas leyes imponen límites físicos a cualquier intento de medir o copiar un estado cuántico sin modificarlo. Como consecuencia directa, cualquier intento de interceptación durante la transmisión de información cuántica introduce perturbaciones inevitables en el sistema. Desde un punto de vista operativo, dichas perturbaciones se reflejan en métricas observables del enlace, siendo la más relevante la tasa de error de bits cuánticos (*quantum bit error rate*, QBER).

Un incremento anómalo del QBER puede ser indicativo de la presencia de ruido excesivo, desalineaciones ópticas, problemas técnicos o interferencias externas, y permite a las partes legítimas identificar situaciones potencialmente inseguras y descartar claves comprometidas antes de su utilización. Junto al QBER, otras métricas como la tasa de clave segura (*secure key rate*, SKR) resultan igualmente críticas, ya que determinan la utilidad práctica del enlace y su capacidad para soportar aplicaciones criptográficas reales de manera sostenida en el tiempo.

A lo largo de las últimas décadas, la QKD ha experimentado una evolución notable. Los primeros experimentos, restringidos a demostraciones en laboratorio y enlaces de corta distancia, han dado paso de forma progresiva a sistemas comerciales y despliegues en entornos reales. Esta evolución ha propiciado el desarrollo de las redes de distribución cuántica de clave (*quantum key distribution networks*, QKDN), que integran múltiples enlaces QKD, nodos intermedios de confianza y sistemas clásicos de gestión de claves con el objetivo de ampliar el alcance geográfico de esta tecnología y hacerla viable desde un punto de vista operativo.

A nivel europeo, uno de los esfuerzos más relevantes en este ámbito es la iniciativa European Quantum Communication Infrastructure (EuroQCI) [2], cuyo objetivo es el despliegue de una infraestructura cuántica paneuropea que combine enlaces terrestres de fibra óptica con comunicaciones satelitales. EuroQCI aspira a proporcionar servicios de comunicación altamente seguros a instituciones públicas, organismos gubernamentales y operadores estratégicos, constituyendo un pilar fundamental de la estrategia europea de soberanía tecnológica y ciberseguridad. Es, además, un esfuerzo para desplegar una primera generación de redes de comunicación cuántica basada en sistemas comerciales, lo que sienta las bases de un futuro ”internet cuántico”.

En este contexto, España participa activamente a través del ecosistema MadQ-

CI [3], que conecta distintos nodos cuánticos distribuidos en la Comunidad de Madrid. Este despliegue representa un entorno real de operación que permite evaluar no solo la viabilidad tecnológica de la QKD, sino también los retos asociados a su explotación continuada en escenarios reales y heterogéneos. La Figura 1 muestra la topología de las diversas infraestructuras de comunicación cuántica (QCI) que integran el ecosistema MadQCI; en azul, los nodos de la Universidad Politécnica de Madrid. Destaca el laboratorio del CeDInt, ubicado en el campus de Montegancedo, que alberga sistemas QKD reales, tanto experimentales como comerciales, y actúa como entorno de pruebas para proyectos nacionales y europeos.

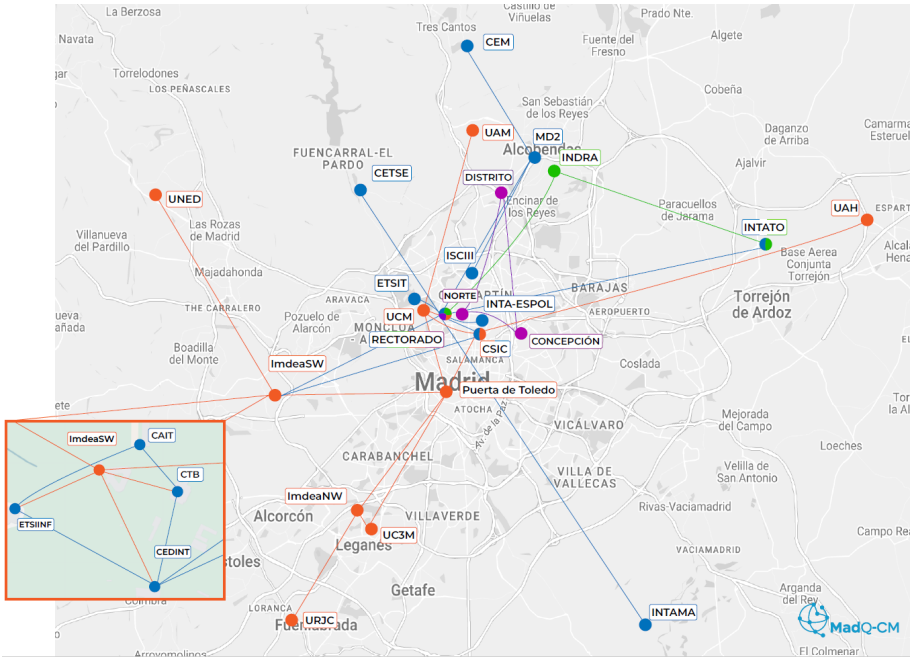


Figura 1: Esquema lógico de alto nivel del ecosistema MadQCI, destacando la ubicación del nodo UPM-CeDInt. En azul, los nodos que administra la UPM, mientras que en rojo se encuentran los nodos que despliega REDIMadrid. En morado y verde se destacan los nodos de los actores industriales Telefónica Innovación Digital e Indra respectivamente.

La transición de la QKD desde el laboratorio hacia infraestructuras operativas reales introduce retos adicionales que van más allá del diseño físico de los dispositivos [4]. Los sistemas QKD son especialmente sensibles a factores físicos y ambientales como la atenuación óptica del canal, las variaciones de temperatura, el ruido de los detectores o pequeñas desalineaciones en los componentes ópticos. Estas perturbaciones afectan de forma directa a métricas críticas como el QBER y la SKR, y pueden manifestarse de manera progresiva y difícil de identificar mediante mecanismos simples de supervisión.

Desde un punto de vista operativo, muchos despliegues actuales de QKD se apoyan en sistemas de gestión de claves (*key management systems*, KMS) y platafor-

mas de monitorización que emplean umbrales estáticos, definidos manualmente y codificados de forma rígida. Este enfoque resulta limitado en infraestructuras reales, ya que no es capaz de adaptarse a fenómenos inevitables como la deriva térmica natural de la fibra, el envejecimiento de los componentes o los cambios graduales en el entorno operativo. Como consecuencia, estos sistemas pueden generar falsas alarmas ante fluctuaciones normales o, por el contrario, pasar por alto degradaciones lentas que comprometen el rendimiento y la disponibilidad del enlace a medio plazo.

Además, las infraestructuras QKDN generan grandes volúmenes de datos operativos en forma de series temporales de métricas cuánticas y parámetros asociados al plano clásico de gestión de claves. La disponibilidad de estos datos no implica automáticamente la obtención de información operativa útil. Sin herramientas de análisis adecuadas, la monitorización queda limitada a una supervisión reactiva, poco eficaz en sistemas altamente estables donde las degradaciones relevantes pueden ser sutiles, multivariantes y dependientes del contexto.

Este Trabajo Fin de Grado surge precisamente de esta necesidad: analizar y explotar datos operativos reales de infraestructuras QKD desde un enfoque claramente ingenieril, orientado a la monitorización continua y a la toma de decisiones operativas informadas. El objetivo no es únicamente detectar valores extremos, sino transformar métricas técnicas de bajo nivel en información interpretable que permita evaluar el estado de salud del enlace cuántico, identificar patrones anómalos y anticipar situaciones de riesgo antes de que deriven en fallos críticos.

Para abordar este problema, se plantea el desarrollo de un *pipeline* de análisis reproducible basado en herramientas de propósito general, que combine análisis estadístico, ingeniería de características y técnicas de aprendizaje automático no supervisado. Estas técnicas no se emplean como un fin en sí mismas, sino como apoyo al análisis de sistemas complejos en un contexto en el que no existen etiquetas explícitas de fallos o ataques reales. El enfoque adoptado busca sentar las bases de un sistema de monitorización operativa alineado con las necesidades reales de infraestructuras como MadQCI, aportando rigor metodológico, interpretabilidad y escalabilidad, y contribuyendo a la transición de la QKD hacia despliegues robustos y mantenibles a largo plazo.

1.3. Objetivos del trabajo

El objetivo general de este Trabajo Fin de Grado es analizar datos operativos reales procedentes de infraestructuras de distribución cuántica de clave con el propósito de extraer conocimiento útil que apoye su monitorización, operación y toma de decisiones en entornos reales.

Aunque el título del trabajo hace referencia a infraestructuras de red cuántica, el alcance se centra de forma deliberada en la monitorización a nivel de enlace QKD (*link-level monitoring*), considerado como la unidad fundamental e indivisible de la red. Este enfoque responde a una motivación técnica clara: sin enlaces QKD que operen de forma correcta y estable, no es posible construir servicios de

red fiables, con independencia de los mecanismos de enrutamiento, retransmisión o gestión de claves de nivel superior. Por este motivo, el trabajo no aborda aspectos como el enrutamiento de claves, el *key forwarding* entre nodos de confianza o la optimización global de la red, sino que se focaliza en el análisis en profundidad del estado operativo de los enlaces individuales que constituyen la base de cualquier QKDN.

A diferencia de enfoques puramente teóricos o basados en simulación, este trabajo se centra en el análisis de datos reales generados por sistemas QKD en funcionamiento, afrontando problemas prácticos asociados a la heterogeneidad de las fuentes, la ausencia de etiquetas y la necesidad de interpretar métricas cuánticas desde una perspectiva operativa.

De forma más concreta, los objetivos específicos del trabajo son los siguientes:

- Analizar el contexto operativo de las infraestructuras QKD actuales, identificando las métricas cuánticas y los parámetros más relevantes para evaluar el estado de un enlace desde un punto de vista práctico.
- Inspeccionar y caracterizar conjuntos de datos operativos reales procedentes de sistemas QKD experimentales y comerciales, identificando patrones de estabilidad, variabilidad y comportamiento anómalo.
- Diseñar e implementar procesos de limpieza, normalización y preprocesamiento que permitan integrar fuentes de datos heterogéneas en un formato coherente, trazable y reproducible.
- Aplicar técnicas de análisis estadístico y exploratorio para estudiar la evolución temporal de las métricas y las relaciones multivariantes entre ellas.
- Definir indicadores operativos (*key performance indicators*, KPI) que sinteticen el estado de salud del enlace cuántico y faciliten su interpretación por parte de un operador.
- Explorar el uso de técnicas de aprendizaje automático no supervisado, como Isolation Forest y autoencoders, como herramientas de apoyo para la detección de estados operativos anómalos en ausencia de datos etiquetados.
- Evaluar el comportamiento del sistema propuesto mediante demostradores operativos físicamente plausibles, analizando tanto su utilidad como sus limitaciones.
- Extraer conclusiones sobre la viabilidad de un flujo de trabajo de monitorización basado en herramientas de propósito general aplicado a infraestructuras de red cuántica reales.

2. Estado del arte y trabajos previos

Este capítulo presenta una revisión crítica y contextualizada del estado del arte relevante para el desarrollo de este Trabajo Fin de Grado. El objetivo no es únicamente describir los conceptos fundamentales asociados a la distribución cuán-

tica de clave, sino situarlos en su contexto tecnológico actual, analizar su evolución desde los entornos de laboratorio hasta los despliegues operativos reales e identificar, desde una perspectiva claramente ingenieril, los principales retos y carencias de los enfoques de monitorización existentes.

En particular, se pone el énfasis en aquellos aspectos más relevantes desde el punto de vista operativo, ya que este trabajo no persigue contribuir al diseño de nuevos protocolos cuánticos ni a la demostración formal de su seguridad, sino al análisis y la explotación de los datos que generan estos sistemas cuando operan de forma continuada en infraestructuras reales. En este sentido, la revisión del estado del arte se orienta a comprender qué métricas se consideran críticas, cómo se monitorizan en la práctica y qué limitaciones presentan los enfoques actuales al enfrentarse a datos reales, ruidosos y no etiquetados.

En primer lugar, se introducen los fundamentos de las redes de distribución cuántica de clave, prestando especial atención a la interpretación física y operativa de sus métricas principales. A continuación, se revisan los principales esfuerzos de estandarización relacionados con la operación de infraestructuras QKD, destacando el papel que desempeñan la observabilidad, el control y la gestión de la información operativa. Posteriormente, se analizan las herramientas genéricas empleadas para el análisis de datos en redes de comunicación clásicas y su posible aplicabilidad al contexto cuántico. Finalmente, se revisan de forma crítica los enfoques basados en aprendizaje automático para la detección de anomalías y las arquitecturas de monitorización propuestas para redes cuánticas, identificando los vacíos existentes que motivan la propuesta desarrollada en este trabajo.

2.1. Fundamentos de las redes de distribución cuántica de clave

La distribución cuántica de clave es una técnica de comunicación segura que permite a dos partes legítimas, tradicionalmente denominadas Alice y Bob, establecer una clave criptográfica compartida con garantías de seguridad basadas en principios físicos. A diferencia de la criptografía clásica, cuya seguridad se puede apoyar en supuestos computacionales, la QKD se fundamenta directamente en las leyes de la mecánica cuántica, lo que supone un cambio conceptual profundo en la forma de entender la seguridad de las comunicaciones.

En protocolos QKD ampliamente estudiados, como BB84, la información se transmite mediante estados cuánticos preparados en bases no ortogonales, es decir, ambiguas para un tercero que trate de averiguar los símbolos enviados. Estos estados no pueden ser medidos ni copiados sin introducir perturbaciones detectables, no por una limitación tecnológica, sino como consecuencia directa de principios fundamentales como el teorema de no clonación y el principio de incertidumbre. Como resultado, cualquier intento de intercepción por parte de un tercero, habitualmente denominado Eva, se traduce en un aumento observable del ruido del canal, lo que permite a las partes legítimas detectar la posible presencia de un ataque y descartar las claves comprometidas antes de su utilización. La información sobre la generación y detección de los símbolos con ciertas bases que poseen Alice y Bob les permiten a ellos, y sólo ellos, re-

solver la ambigüedad y generar el material clave. La Figura 2 representa un par QKD como el explicado.

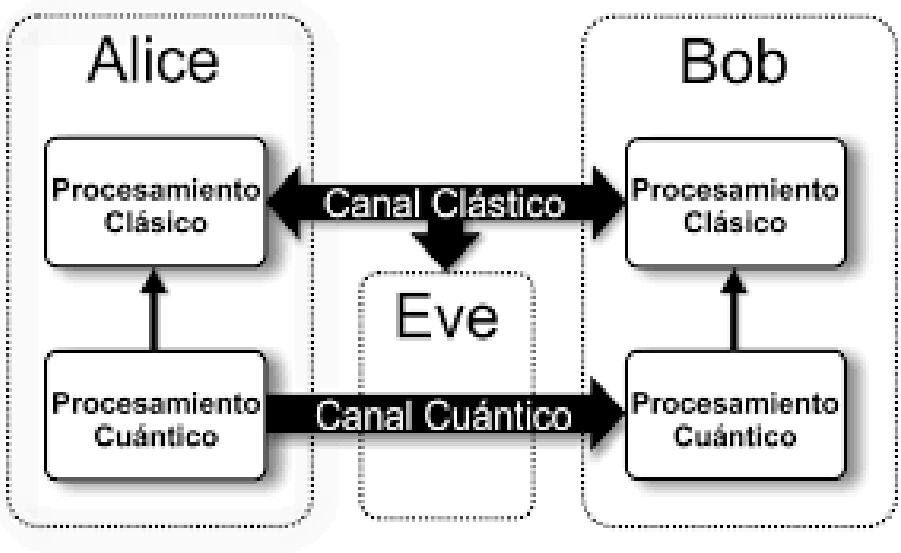


Figura 2: Representación esquemática de un enlace QKD punto a punto, mostrando los canales cuántico y clásico, y la posición teórica de un adversario (Eva). Nótese que el canal clásico puede ser público, no confidencial, pero las partes se tienen que poder autenticar, lo que habitualmente se hace con clave precompartida.

Los primeros sistemas QKD se desarrollaron como enlaces punto a punto, con demostraciones experimentales en laboratorio y alcances muy limitados. Estas restricciones se debían principalmente a las pérdidas ópticas de la fibra, a la eficiencia de los detectores y a la baja tasa de repetición de los sistemas de generación de fotones disponibles en ese momento. Aunque estos sistemas fueron clave para validar el paradigma QKD, a gran escala en entornos operativos no han sido viables.

Para superar estas limitaciones y permitir el despliegue de la tecnología en escenarios reales, surgieron las redes de distribución cuántica de clave. Estas redes integran múltiples enlaces QKD, nodos intermedios de confianza y sistemas clásicos de gestión de claves con el objetivo de ampliar el alcance geográfico de la tecnología y facilitar su integración con infraestructuras de comunicación existentes. Este enfoque ha hecho posible el despliegue de redes cuánticas en entornos urbanos y metropolitanos, pero también ha introducido una capa adicional de complejidad desde el punto de vista de la operación y la gestión.

Desde una perspectiva práctica, la seguridad y el rendimiento global de una QKDN no dependen únicamente del comportamiento del canal cuántico. El plano clásico, encargado de tareas como la corrección de errores, la amplificación de privacidad, la autenticación, la sincronización y la distribución de claves a los usuarios finales, desempeña un papel igualmente crítico. Como consecuencia,

la operación de una red QKD real debe entenderse como la de un sistema híbrido cuántico/clásico, en el que fallos o degradaciones en cualquiera de los dos planos pueden afectar de forma directa tanto a la disponibilidad como a la seguridad del servicio.

Por ello hay que imaginárnoslo como dos planos que, aun trabajando juntos, son muy diferentes y no se pueden estudiar juntos. Pero sí se puede entender cómo interoperan entre ellos y cómo las interferencias o cambios en uno afectan al otro.

Métricas cuánticas operativas

El estado operativo de un enlace QKD se evalúa fundamentalmente a partir de un conjunto reducido de métricas cuánticas que condensan el comportamiento físico del sistema. Entre ellas, las más relevantes desde el punto de vista de la monitorización son la tasa de error de bits cuánticos y la tasa de clave segura.

El QBER mide la fracción de bits erróneos detectados durante la fase de intercambio cuántico y puede definirse formalmente como:

$$\mathrm{QBER} = \frac{N_{\mathrm{err}}}{N_{\mathrm{tot}}},$$

donde N_{err} representa el número de detecciones erróneas y N_{tot} el número total de detecciones válidas consideradas en el protocolo. Esta métrica proporciona una medida directa del nivel de ruido efectivo del canal cuántico y de la calidad de la transmisión cuántica. Así como la calidad del mismo canal.

Desde un punto de vista físico, un aumento del QBER puede deberse a múltiples causas que no implican necesariamente la existencia de un ataque activo (de hecho por eso el tutor eligió la palabra anomalías). Entre las más habituales se encuentran las desalineaciones de bases de generación y detección, el ruido oscuro de los detectores, las fluctuaciones térmicas en la fibra óptica, la deriva temporal de componentes ópticos o la presencia de interferencias externas. En el caso de un ataque de intercepción y reenvío, la intervención de Eva introduce errores adicionales como consecuencia de la imposibilidad de medir estados cuánticos sin perturbarlos, lo que también se refleja en un incremento del QBER. Por ello, y como explicaremos más adelante, siempre tiene que haber un experto en este tipo de sistemas para, al saltar la alarma de detección de anomalía, revisar el motivo de esta.

Desde un punto de vista operativo, un QBER elevado compromete la capacidad del sistema para extraer una clave segura. En protocolos QKD prácticos, si el QBER supera un umbral crítico determinado por el propio protocolo y por los parámetros de post-procesado, la sesión debe abortarse, ya que no es posible garantizar la seguridad de la clave generada. Este comportamiento establece una frontera clara entre operación normal y operación insegura, pero no aporta información detallada sobre el origen de la degradación ni sobre su evolución temporal. En algunos sistemas cuánticos que no forman parte del alcance de este trabajo, como los QKD de variables continuas, usan otra estimación de

parámetros, como el ruido de exceso. Son, en todo caso, equivalentes desde el punto de vista de su evaluación y, en presencia de un parámetro anómalo, se descarta la clave generada.

Por su parte, la SKR representa la velocidad neta a la que se obtiene la clave criptográfica final tras completar los procesos de postprocesado clásico, como la corrección de errores y la amplificación de privacidad. Esta métrica determina la utilidad práctica del enlace, ya que una SKR excesivamente baja limita la capacidad del sistema para soportar aplicaciones criptográficas reales, incluso cuando el enlace es formalmente seguro desde un punto de vista teórico.

Ambas métricas están estrechamente relacionadas. En condiciones normales de operación, un incremento del QBER suele traducirse en una reducción de la SKR, ya que el sistema debe descartar una mayor fracción de bits durante el post-procesado. Sin embargo, en infraestructuras reales esta relación no siempre es directa ni lineal, debido a la intervención de factores adicionales asociados al plano clásico, como cuellos de botella en el procesamiento, latencias de comunicación o limitaciones del sistema de gestión de claves. Desde el punto de vista de la monitorización, resulta por tanto insuficiente analizar estas métricas de forma aislada, siendo necesario estudiar su evolución temporal, su coherencia mutua y su comportamiento conjunto. En este trabajo de fin de grado se trabajará siempre con estas dos métricas juntas.

2.2. Estandarización y aspectos operacionales de QKD

El despliegue de la QKD a escala de red ha puesto de manifiesto la necesidad de definir arquitecturas comunes, interfaces estandarizadas y procedimientos operativos interoperables. En la práctica, muchas infraestructuras QKD actuales se basan en el uso de nodos de confianza, en los que las claves se descifran y se vuelven a cifrar para ser reenviadas a lo largo de la red, permitiendo extender el alcance geográfico de la tecnología utilizando infraestructuras clásicas existentes.

Esta arquitectura, la más usado a día de hoy, introduce nuevos puntos de fallo de naturaleza no cuántica y desplaza parte de la carga de seguridad hacia la correcta operación de los sistemas clásicos. La operación segura de una QKDN no depende únicamente de la física del enlace cuántico, sino también de la gestión, monitorización y coordinación de los sistemas clásicos asociados.

Desde un punto de vista operativo, esta realidad refuerza la necesidad de disponer de mecanismos de gestión y monitorización que permitan evaluar el estado del sistema de forma continua y detectar degradaciones antes de que se traduzcan en interrupciones del servicio o en situaciones potencialmente inseguras.

Esfuerzos de estandarización (ITU-T y ETSI)

Los principales esfuerzos de estandarización en el ámbito de las redes QKD han sido liderados por organismos internacionales como ITU-T y ETSI, que han abordado la tecnología desde una perspectiva funcional y de integración en redes de telecomunicaciones existentes.

El ITU-T ha desarrollado un marco funcional para las redes de distribución cuántica de clave, recogido principalmente en la serie de recomendaciones ITU-T Y.3800. Estas recomendaciones definen los distintos bloques funcionales de una QKDN, incluyendo los módulos de generación de clave, los sistemas de gestión de claves y los mecanismos de control y orquestación de la red. Un aspecto especialmente relevante de este marco es la necesidad explícita de disponer de información fiable y actualizada sobre el estado operativo de los enlaces QKD, tanto a nivel cuántico como clásico. La Figura 3 refleja esta arquitectura.

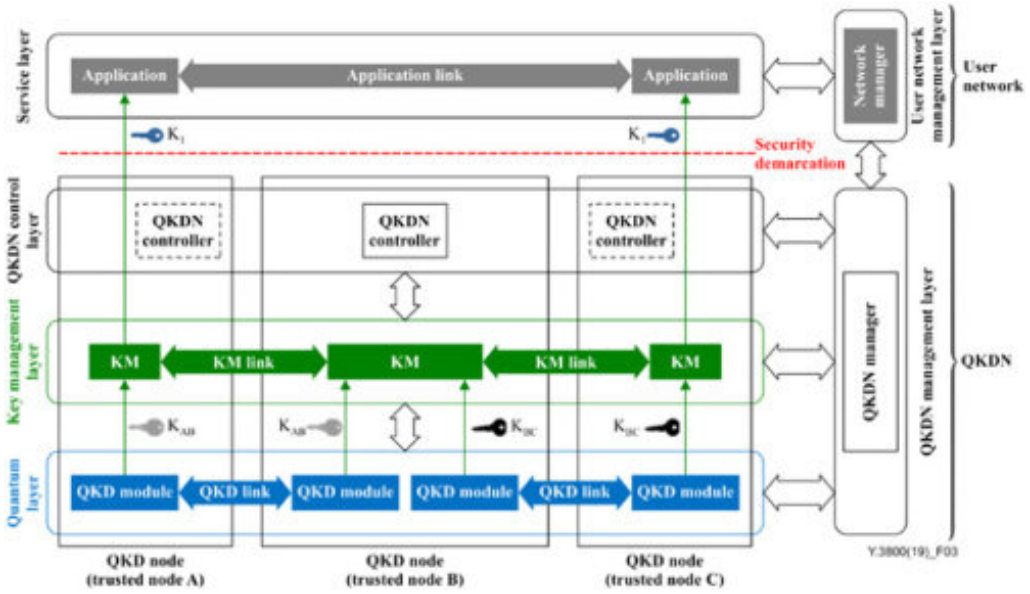


Figura 3: Arquitectura funcional de referencia para una red QKD según la recomendación ITU-T Y.3800. Se distinguen claramente la capa cuántica, que provee de material clave simétrico; la capa de gestión de claves, que usa cifrado incondicionalmente seguro para transportar la clave de los usuarios cifrándola con la clave generada por los QKD; las capas de control; que dotan a los sistemas QKD y KMS de las nociones de gestión de red que necesiten; y la de servicio, que consume el material clave transportado por la clave generada en los sucesivos saltos QKD.

Por otro lado, ETSI, a través del grupo ISG QKD, ha publicado numerosas especificaciones centradas en los requisitos de seguridad, las interfaces y la integración de sistemas QKD en infraestructuras de telecomunicaciones existentes. Estos documentos destacan de forma recurrente la importancia de la monitorización continua como un elemento esencial para la operación segura y eficiente de infraestructuras QKD reales, especialmente en entornos heterogéneos y multivendedor.

Tanto ETSI como ITU-T coinciden en que la monitorización no debe limitarse a la detección de fallos catastróficos, sino que debe proporcionar una visión continua del estado del sistema, capaz de identificar degradaciones progresivas, comportamientos anómalos y cambios de régimen operativo antes de que comprometan

la seguridad o la disponibilidad del servicio. De aquí surgió la idea de un doble algoritmo para esta detección de anomalías.

2.3. Herramientas genéricas para el análisis de datos de red

En redes de comunicación clásicas, la monitorización operativa se ha apoyado tradicionalmente en protocolos y herramientas basadas en umbrales fijos y alarmas predefinidas. Si bien este enfoque resulta eficaz para detectar fallos abruptos, presenta limitaciones importantes en sistemas complejos y altamente estables como las infraestructuras QKD, donde las anomalías relevantes pueden manifestarse de forma sutil, multivariante y dependiente del contexto operativo y temporal.

Lenguajes de programación como Python y R, a día de hoy, son los más usados en grandes empresas para el análisis de datos. Esto se debe a su facilidad para leer, modificar y ejecutar grandes cantidades de datos. Permiten implementar procesos personalizados de limpieza, análisis estadístico, visualización y modelado. Asímismo, permiten generar *pipelines* de trabajo, MLops, DEVops, etc.

Aunque el primer *script* fue con R, en este trabajo se ha optado por utilizar Python como lenguaje principal para el análisis operativo, apoyándose en bibliotecas ampliamente utilizadas en el ámbito de la ingeniería de datos, como pandas (la librería mas usada para el tratamiento de archivos tipo CSV), NumPy (herramienta clave para cálculos matemáticos) y matplotlib (para visualizaciones). También se ha usado herramientas como Pathlib para tener un entorno realista y reproducible, con archivos donde tenemos centralizadas variables constantes que serán usadas en diferentes *scripts*. Otras herramientas usadas fueron R o librerías como sklearn o torch.

Un enfoque así permitirá construir un *pipeline* reproducible, limpio y comprensible para cualquier experto que quiera usarlo en entornos de QKDN.

2.4. Detección de anomalías mediante aprendizaje automático en QKD

La detección de anomalías es un problema muy estudiado en el ámbito de la monitorización de sistemas complejos. En el contexto de las infraestructuras QKD, una anomalía puede corresponder tanto a un fallo técnico puntual, como a una degradación progresiva del canal o a un comportamiento operativo inesperado del plano clásico.

Una parte significativa de la literatura leída aborda este problema desde una perspectiva teórica o mediante simulaciones numéricas. Muchos trabajos se centran en la detección de ataques ideales o escenarios controlados, utilizando modelos simplificados del canal cuántico y datos sintéticos generados bajo supuestos bien definidos. Aunque estos estudios son valiosos para comprender las propiedades fundamentales de los protocolos, su aplicabilidad directa a infraestructuras reales es limitada.

Existe, por tanto, un vacío claro en la literatura en lo que respecta al tratamiento de datos operacionales reales, ruidosos y no etiquetados procedentes de despliegues QKD en producción. En estos escenarios, las anomalías relevantes no siempre corresponden a ataques teóricos bien definidos, sino a combinaciones complejas de factores físicos, ambientales y operativos. Es precisamente en este contexto donde el presente Trabajo Fin de Grado pretende aportar valor, centrándose en el análisis de datos reales y en problemas prácticos de monitorización.

Limitaciones de los métodos tradicionales

Los métodos tradicionales de monitorización basados en umbrales fijos presentan limitaciones importantes en sistemas QKD reales. La definición manual de umbrales adecuados resulta incompleta debido a la variabilidad natural del sistema y a la dependencia de múltiples factores físicos y ambientales, como la temperatura, la longitud del enlace o el estado de los componentes ópticos.

Además, estos enfoques suelen ser univariantes y no capturan relaciones estructurales entre métricas, lo que dificulta la detección de degradaciones progresivas que permanecen dentro de rangos aceptables en términos absolutos, pero que alteran la coherencia global del sistema y su comportamiento esperado a largo plazo. Cómo se dijo antes, este trabajo propone un estudio bivariable (QBER y SKR), las dos métricas clave en este tipo de sistemas.

Algoritmos de aprendizaje automático aplicados

Para superar estas limitaciones, diversos trabajos han explorado el uso de técnicas de aprendizaje automático, especialmente en su vertiente no supervisada. Algoritmos como Isolation Forest permiten identificar observaciones raras en espacios de alta dimensionalidad sin necesidad de datos etiquetados (no se tienen datos de este tipo en este trabajo), mientras que los *autoencoders* son capaces de capturar relaciones estructurales complejas entre variables mediante técnicas de aprendizaje profundo.

No obstante, la literatura coincide en que estas técnicas no deben emplearse como mecanismos autónomos de decisión, es decir nunca podría etiquetarse una alarma explícitamente como alerta de algún tipo.

En infraestructuras críticas como las QKDN, interpretar los resultados, la robustez frente a ruido y la reducción de falsos positivos son requisitos fundamentales. Por este motivo, los enfoques más prometedores combinan técnicas de aprendizaje automático con conocimiento del dominio y análisis operativo, evitando soluciones de tipo caja negra difícilmente justificables.

Arquitecturas de monitorización y capas de control

La monitorización de redes QKD debe integrarse dentro de arquitecturas de control y gestión que permitan reaccionar ante cambios en el estado operativo del sistema. En estas arquitecturas, los sistemas de gestión de claves desempeñan

un papel central, coordinando la generación, el almacenamiento y la distribución del material criptográfico entre los distintos componentes de la red.

Desde un punto de vista operativo, resulta esencial que estos sistemas dispongan de información fiable y actualizada sobre el estado de los enlaces QKD individuales. Aunque existen plataformas de monitorización ampliamente utilizadas en redes clásicas, en este trabajo se ha optado por un enfoque basado en el análisis directo de datos operativos y en visualizaciones generadas mediante Python, priorizando la interpretabilidad, la flexibilidad y el control completo sobre el proceso de análisis.

Este enfoque se alinea con el objetivo general del trabajo: desarrollar un *pipeline* de monitorización reproducible y flexible, capaz de evolucionar hacia integraciones más complejas en el futuro, pero centrado inicialmente en la comprensión profunda del comportamiento operativo de infraestructuras QKD reales.

3. Metodología, datasets y análisis exploratorio inicial

Este capítulo describe la metodología seguida para el análisis de los datos procedentes de infraestructuras reales de distribución de clave cuántica, así como los conjuntos de datos utilizados y los resultados obtenidos durante la fase inicial de análisis exploratorio. Esta etapa constituye un bloque central del trabajo, ya que actúa como nexo entre el marco teórico presentado en el estado del arte y la realidad operativa de sistemas QKD desplegados y en funcionamiento continuo.

El objetivo principal de este capítulo no es únicamente presentar resultados estadísticos preliminares, sino contextualizar los datos desde un punto de vista físico y operativo. Se busca comprender cómo se comportan los sistemas QKD reales, identificar qué patrones pueden considerarse normales y cuáles representan desviaciones relevantes, y sentar unas bases metodológicas sólidas sobre las que construir posteriormente el modelado avanzado y los mecanismos de detección de anomalías desarrollados en capítulos posteriores.

En infraestructuras QKD reales, el análisis exploratorio adquiere una importancia especial debido a varias características propias del dominio. En primer lugar, no existen etiquetas explícitas que indiquen de forma inequívoca la presencia de fallos o ataques, lo que impide aplicar directamente enfoques supervisados clásicos. Por otro lado los sistemas en general operan siendo altamente estables durante mucho tiempo, con variaciones pequeñas pero continuas asociadas a factores físicos inevitables. Comprender qué se considera comportamiento normal, qué tipo de variabilidad es esperable y qué desviaciones resultan operativamente relevantes constituye un paso imprescindible antes de aplicar cualquier técnica avanzada de aprendizaje automático.

3.1. Datasets operacionales y contextualización

El análisis realizado en este trabajo se apoya en tres conjuntos de datos operativos reales procedentes de sistemas QKD desplegados en el laboratorio del

edificio CeDInt de la Universidad Politécnica de Madrid, en el marco de la infraestructura MadQCI. Estos *datasets* corresponden a dos fabricantes distintos y a escenarios de operación claramente diferenciados, lo que permite estudiar comportamientos contrastados y evaluar la robustez de los enfoques propuestos en contextos heterogéneos.

Desde el punto de vista de adquisición, los datos llegan en forma de archivos CSV exportados directamente desde los sistemas de monitorización de los dispositivos QKD. En todos los casos se trata de volcados de logs operativos ya procesados por el fabricante o por el sistema de control del laboratorio, sin modificaciones manuales posteriores. Este formato, aunque sencillo, refleja fielmente una situación habitual en entornos reales, donde los datos disponibles para el análisis suelen proceder de exportaciones periódicas y no de flujos de telemetría diseñados específicamente con fines de investigación.

Los *datasets* permiten estudiar el comportamiento del sistema tanto en entornos controlados de laboratorio como en despliegues operativos prolongados en red, aportando una visión amplia y realista del funcionamiento de infraestructuras QKD actuales. Esta diversidad resulta especialmente valiosa para evitar conclusiones excesivamente dependientes de un único escenario y para poner de manifiesto las diferencias operativas entre plataformas con filosofías de diseño muy distintas.

Dataset QTI: parámetros físicos y operativos

El *dataset* QTI procede de un sistema QKD comercial desplegado y operado de forma estable en el entorno del laboratorio del CeDInt. Este sistema está basado en tecnología QKD de variables discretas de QTI, que implementa el protocolo BB84 con estados señuelo y codificación por división del tiempo (codificación *time-bin*). Representa una plataforma ampliamente utilizada en entornos experimentales, pilotos precomerciales y pruebas de interoperabilidad dentro del ecosistema QKD europeo.

Desde el punto de vista de adquisición, los datos se proporcionan en forma de archivos CSV con una estructura simple, en la que cada fila corresponde a una medición periódica del estado del enlace. La frecuencia de muestreo original es aproximadamente de una muestra cada 15 minutos, con ligeras variaciones debidas a la temporización interna del sistema, como puede observarse en las marcas temporales consecutivas.

Durante la inspección inicial se detectaron pequeñas irregularidades temporales, como diferencias de algunos minutos entre muestras consecutivas. No obstante, no se observaron pérdidas significativas de datos ni huecos prolongados que requirieran imputación explícita, por lo que se optó por trabajar con las marcas temporales originales, preservando la información real proporcionada por el sistema. Cabe decir que es el *dataset* más corto de los recibidos y que solo se consiguió un CSV con 87 muestras. Por lo que las pruebas del *pipeline* no se realizaron con esta máquina.

El sistema QTI analizado presenta un diseño orientado prioritariamente a la

robustez, la estabilidad y la fiabilidad del enlace cuántico, incluso en condiciones no ideales. Este enfoque se refleja tanto en la arquitectura del *hardware* como en los parámetros operativos configurados por el fabricante, priorizando tasas de error bajas, estabilidad temporal y repetibilidad del comportamiento frente a un rendimiento máximo en términos de generación de clave.

Desde un punto de vista metodológico, este tipo de sistema resulta especialmente interesante como referencia de comportamiento estable. Al operar en un entorno controlado, con perturbaciones externas limitadas, permite caracterizar con claridad el régimen normal del sistema y analizar cómo se comportan las métricas cuánticas cuando el enlace funciona de forma óptima durante largos periodos de tiempo.

El *dataset* son interesantes las siguientes variables:

- Marca temporal de cada medición, con resolución suficiente para el análisis temporal.
- Tasa de error de bits cuánticos (QBER).
- Tasa de clave segura (SKR), expresada en bits por segundo.
- Pérdidas ópticas del canal (`channel_loss`), expresadas en decibelios.

En estos primeros análisis sorprendió la SKR tan baja que tenía esta máquina. Tras preguntarle al tutor, comentó que aunque este valor es significativamente inferior al observado en otros sistemas comerciales de última generación, debe interpretarse en el contexto de los objetivos de diseño del fabricante. Según la documentación técnica y la experiencia operativa acumulada en el laboratorio, estos sistemas priorizan la estabilidad, la tolerancia a perturbaciones del canal y la repetibilidad de los resultados, lo que los convierte en plataformas especialmente adecuadas para estudios de comportamiento y validación experimental.

Datasets Toshiba (W25 y W27): datos operacionales en red

Los *datasets* Toshiba corresponden a un sistema QKD comercial de alto rendimiento desplegado en un entorno operativo real dentro de la infraestructura MadQCI. En este caso, Toshiba lleva fabricando estos sistemas por una década, basados en el protocolo de codificación en fase T12, y tienen un desempeño de un orden de magnitud mayor que el de sus competidores. Se analizan dos campañas de operación diferenciadas, denominadas Toshiba-2024-W25 y Toshiba-2025-W27, que reflejan distintos periodos temporales, condiciones de operación y volúmenes de tráfico.

En este caso, los datos también se proporcionan en formato CSV, con una estructura más reducida que el QTI, incluyendo solamente parámetros de: `Time`, `QBER`, `SecureKeyR`
La frecuencia de muestreo es significativamente mayor, con una muestra por minuto, lo que refuerza la idea comentada por el tutor, estas máquinas son más potentes que las QTI.

En la primera inspección no se observaron intervalos de tiempo irregulares entre muestras, consistentes con el funcionamiento normal del sistema y con la políti-

ca de exportación de datos del fabricante. No se detectaron huecos prolongados que hicieran necesaria la imputación de valores.

A diferencia del sistema QTI, los sistemas QKD de Toshiba están diseñados con un objetivo claramente orientado al rendimiento, priorizando la maximización de la tasa de generación de clave segura. Este enfoque se apoya en tecnologías avanzadas de detección, electrónica de alta velocidad y optimización intensiva del post-procesado clásico, muchas de ellas propietarias y no divulgadas públicamente.

En los *datasets* analizados, la tasa de clave segura alcanza valores medios en torno a los 380 kb/s, aproximadamente dos órdenes de magnitud superiores a los observados en el sistema QTI. Esta diferencia no está relacionada con la configuración de red ni con la topología del enlace, sino que responde principalmente a decisiones de diseño del fabricante y al grado de optimización tecnológica del sistema.

Este comportamiento ha sido documentado también en literatura científica y en discusiones técnicas dentro de la comunidad QKD, donde se destaca que los sistemas Toshiba presentan un rendimiento significativamente superior al de otras plataformas comerciales, debido a una explotación más eficiente de principios fundamentales de electrónica.

Los *datasets* Toshiba incluyen:

- Marca temporal de cada medición.
- QBER.
- Tasa de clave segura (SKR), expresada en bits por segundo.

A diferencia del *dataset* QTI, no se dispone de información explícita sobre las pérdidas ópticas del canal, lo que condiciona el análisis multivariante posterior y refleja una limitación habitual en datos operativos reales proporcionados por fabricantes comerciales. La Figura 4 representa los datos en crudo de los dos *datasets*.

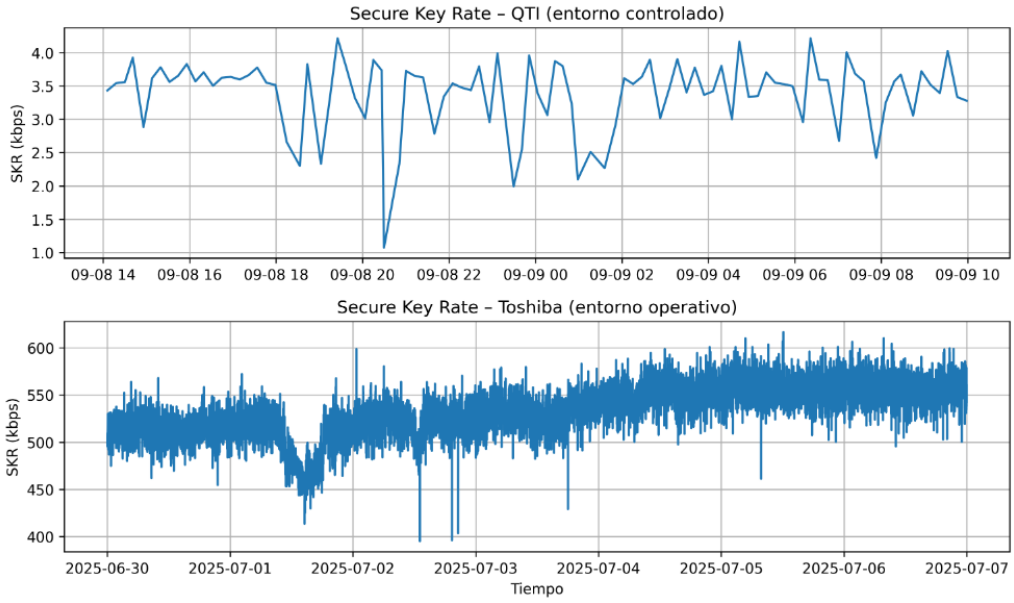


Figura 4: Comparativa visual de las series temporales de tasa de clave segura (SKR) para los sistemas QTI (superior) y Toshiba (inferior). Nótese la diferencia en órdenes de magnitud y en la variabilidad de la señal, así como la mayor tasa de muestreo del segundo.

3.2. Framework de análisis, preprocesamiento y estructuración temporal

Selección del entorno de análisis

La fase inicial de análisis exploratorio se realizó utilizando el lenguaje R, debido a su facilidad para el análisis estadístico y la visualización de series temporales. R proporciona estructuras de datos y bibliotecas (en especial dplyr) especializadas que facilitan un manejo preciso de los índices temporales y una exploración rápida de patrones de estabilidad y manejo.

Este análisis preliminar permitió obtener una primera comprensión del comportamiento general de los sistemas QKD, identificar rangos operativos habituales y detectar posibles problemas de calidad en los datos antes de abordar un *pipeline* más complejo orientado a producción.

Posteriormente, el procesamiento sistemático y el modelado avanzado se trasladaron a Python. Esta decisión responde a la necesidad de mayor flexibilidad algorítmica y control sobre el flujo de datos. A diferencia de herramientas de monitorización de tipo caja negra como Zabbix o Grafana, el uso de Python y bibliotecas como pandas nos permite aplicar operaciones matemáticas complejas sobre los datos, como ventanas deslizantes, derivadas temporales, retardos *lags* y combinaciones no lineales de métricas, que resultan difíciles o directamente inviables de implementar en plataformas estándar de monitorización.

Python facilita la integración directa con bibliotecas de aprendizaje automático

(cosa que R dificulta) y la automatización completa del flujo de trabajo, lo que resulta esencial para construir un *pipeline* reproducible y extensible, alineado con los objetivos de este Trabajo Fin de Grado.

Homogeneización de variables temporales y métricas

Uno de los principales retos iniciales fue la heterogeneidad de formatos entre *datasets*. Los distintos fabricantes utilizan convenciones diferentes tanto para las marcas temporales como para la denominación, estructura y unidades de las métricas proporcionadas.

Para resolver este problema se diseñó un proceso sistemático de homogeneización que incluyó:

- Conversión de todas las marcas temporales a un formato común (`timestamp`), garantizando coherencia temporal.
- Normalización de los nombres de métricas a un esquema coherente (`qber`, `skr`, `loss`).
- Verificación explícita de las unidades de la tasa de clave segura, confirmando que todos los valores están expresados en bits por segundo.

Adicionalmente, se abordó explícitamente el problema de las diferencias de escala entre sistemas. Mientras que el sistema QTI opera con valores de SKR del orden de kilobits por segundo, los Toshiba alcanzan centenas de kilobits por segundo. Esta disparidad de escalas plantea un problema directo para cualquier análisis multivariante o modelo de aprendizaje automático, ya que las variables de mayor magnitud numérica tienden a dominar el comportamiento del modelo.

Inicialmente se exploró el uso de técnicas de escalado como StandardScaler y MinMaxScaler para normalizar conjuntamente los datos. Sin embargo, este análisis puso de manifiesto una cuestión metodológica relevante: las diferencias de escala no son únicamente numéricas, sino que reflejan regímenes operativos y filosofías de diseño radicalmente distintas. Como resultado, y tras discutirlo con el tutor, se decidió realizar análisis independientes para cada tipo de máquina, evitando mezclar plataformas heterogéneas en un único modelo y preservando la interpretabilidad física de los resultados.

3.3. Análisis estadístico descriptivo y caracterización de estabilidad

Estabilidad operacional

Como primer paso del análisis exploratorio se calcularon estadísticas descriptivas básicas para las métricas QBER y SKR, incluyendo la media, la desviación estándar, la asimetría y la curtosis. Estas métricas permiten obtener una primera caracterización cuantitativa del comportamiento del sistema y comparar de forma objetiva la estabilidad relativa de las distintas plataformas.

El sistema QTI presenta valores medios de QBER más bajos (como ya se ha explicado anteriormente) y una desviación estándar significativamente menor que los sistemas Toshiba, lo que confirma su carácter altamente estable y su operación en un entorno controlado. En contraste, los *datasets* de Toshiba muestran una mayor variabilidad temporal, consistente con un entorno operativo real.

Las distribuciones de QBER y SKR en los sistemas Toshiba presentan valores de curtosis elevados, indicando distribuciones leptocúrticas con presencia de eventos más extremos. Este comportamiento es característico de sistemas que operan cercanos a su régimen óptimo, pero que experimentan caídas abruptas de rendimiento en situaciones puntuales, como reconfiguraciones internas, saturaciones del plano clásico o perturbaciones externas. La Figura 5 ejemplifica estos efectos.

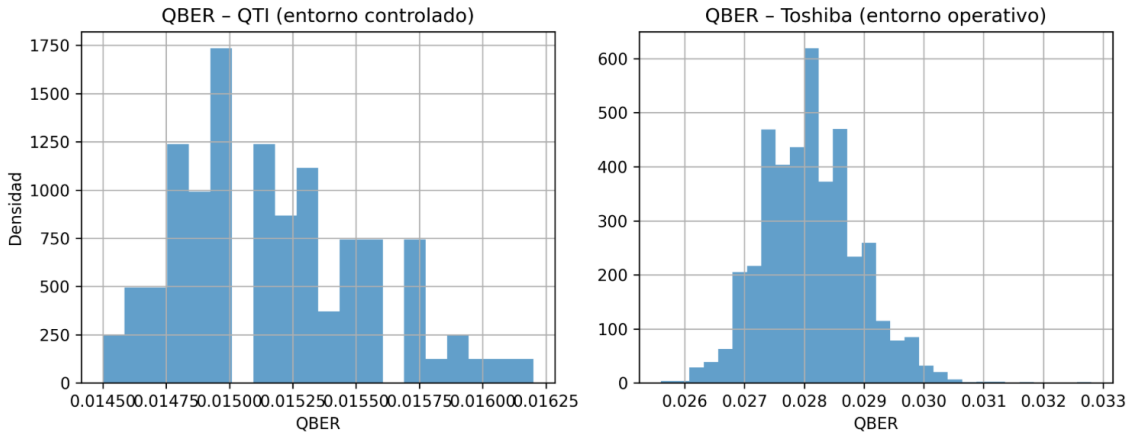


Figura 5: Análisis comparativo de la distribución de densidad del QBER. A la izquierda, el sistema QTI muestra una variabilidad mínima propia de un entorno controlado. A la derecha, el sistema Toshiba en producción presenta una distribución leptocúrtica con mayor media y dispersión, reflejando la dinámica de una red real.

3.4. Detección preliminar de valores atípicos

Como parte del análisis exploratorio, se realizó una detección inicial de valores atípicos mediante el cálculo del Z-score para las métricas QBER y SKR. Se consideraron atípicas aquellas observaciones que superaban tres desviaciones estándar respecto a la media histórica de cada sistema.

Este análisis permitió identificar un conjunto reducido de eventos extremos, concentrados principalmente en el *dataset* Toshiba-2025. Dado el carácter no gaussiano de las distribuciones observadas, esta detección debe interpretarse como una aproximación preliminar y no como un mecanismo definitivo de detección de anomalías.

No obstante, los eventos identificados corresponden a situaciones operativas relevantes, como interrupciones en la generación de clave o degradaciones severas del enlace, y resultaron especialmente útiles como referencia inicial para

el diseño, calibración y validación de los modelos de detección de anomalías desarrollados en fases posteriores del trabajo.

4. Desarrollo de la Propuesta: Modelado Avanzado y Detección de Anomalías

Este capítulo describe en profundidad el diseño, la implementación y la justificación técnica del sistema propuesto para la monitorización avanzada de infraestructuras de distribución cuántica de clave mediante técnicas de aprendizaje automático no supervisado. Este capítulo constituye el núcleo técnico del Trabajo Fin de Grado y es donde se materializa el enfoque ingenieril adoptado: transformar datos operativos reales, ruidosos y no etiquetados en información accionable para la operación de enlaces QKD.

A diferencia de enfoques puramente estadísticos o basados en umbrales fijos (cosa que se ha descartado anteriormente), la propuesta desarrollada en este trabajo se apoya en un *pipeline* modular, reproducible y explícitamente diseñado para capturar tanto anomalías puntuales como alteraciones estructurales del comportamiento operativo del sistema. El objetivo no es únicamente detectar valores extremos de una métrica concreta, sino identificar incoherencias físicas y funcionales en el comportamiento conjunto del sistema, incluso cuando las métricas individuales permanecen dentro de rangos considerados aceptables por el fabricante.

A continuación, se explicará desde los requisitos al *pipeline* en sí (preprocesamiento, ingeniería de características, detección de anomalías con IF y AE y la fusión y análisis de los modelos):

4.1. Requisitos funcionales y técnicos de la monitorización QKD

Las infraestructuras QKD presentan características operativas que las diferencian de forma clara de otros sistemas de telecomunicaciones clásicos. En particular, los enlaces QKD suelen operar durante largos periodos en regímenes altamente estables, mientras que las perturbaciones relevantes pueden manifestarse de forma sutil, progresiva o fuertemente dependiente del contexto operativo y ambiental (cuyos datos no se tienen en este trabajo).

Esta combinación de alta estabilidad y baja frecuencia de eventos anómalos plantea un reto significativo para la monitorización. Los mecanismos basados en umbrales simples tienden a ser demasiado rígidos: o bien se configuran de forma conservadora, generando un elevado número de falsas alarmas ante fluctuaciones naturales, o bien se ajustan de forma laxa y no detectan degradaciones progresivas que pueden comprometer el rendimiento o la disponibilidad del enlace a medio plazo.

Desde un punto de vista funcional, el sistema de monitorización propuesto debe cumplir los siguientes requisitos fundamentales:

- operar en ausencia de etiquetas, ya que no se dispone de ejemplos clasificados de fallos o ataques reales en infraestructuras QKD operativas;
- adaptarse a plataformas QKD heterogéneas, con fabricantes, tecnologías y regímenes operativos distintos;
- detectar tanto eventos abruptos, como caídas súbitas de rendimiento, como degradaciones suaves y progresivas del sistema;
- minimizar la generación de falsos positivos en entornos altamente estables; y
- proporcionar resultados interpretables desde un punto de vista físico y operativo, facilitando la toma de decisiones por parte de operadores humanos expertos en este tipo de entornos y máquinas.

Desde el punto de vista técnico, el sistema debe ser modular, reproducible y escalable, y compatible con su integración en *pipelines* automatizados. Estos requisitos descartan soluciones cerradas o excesivamente dependientes de herramientas propietarias, y motivaron el uso de Python como lenguaje principal de desarrollo, apoyándose en bibliotecas ampliamente utilizadas en ingeniería de datos y aprendizaje automático.

4.2. Preprocesamiento y normalización de datos

El *pipeline* comienza con un módulo de preprocesamiento (`preprocesamiento.py`) diseñado específicamente para tratar *datasets* operativos QKD heterogéneos de forma robusta y automatizada. Esta fase es crítica, ya que los modelos posteriores son especialmente sensibles a inconsistencias en las unidades, escalas, valores ausentes o estructuras de los datos.

El módulo de preprocesamiento carga los archivos CSV disponibles y estandariza los nombres de las columnas. A continuación, detecta automáticamente el tipo de *dataset* en función de su estructura interna y del identificador de fuente, evitando la introducción de reglas específicas por fabricante y favoreciendo la generalización del *pipeline*.

El preprocesamiento incluye las siguientes tareas principales:

- conversión segura de las marcas temporales a un formato común (`timestamp`).
- limpieza explícita de valores infinitos y no numéricos.
- eliminación de columnas con un porcentaje excesivo de valores ausentes (umbral del 40 %).
- eliminación de filas con valores no finitos restantes.
- inclusión de un identificador explícito de la fuente de datos `source`.

En el caso de los *datasets* QTI se conserva la información relativa a las pérdidas ópticas del canal, mientras que en los *datasets* Toshiba esta métrica se introduce como valor nulo (ya que no existe en esos *datasets*, cabe recordar que los

datasets de Toshiba solo contienen QBER Y SKR). Esta decisión evita introducir suposiciones artificiales y preserva la trazabilidad del origen real de los datos.

4.3. Ingeniería de características para aprendizaje automático

Las métricas QKD originales (QBER, SKR y pérdidas ópticas) describen únicamente el estado instantáneo del sistema y, por sí solas, no contienen información explícita sobre su dinámica temporal ni sobre relaciones estructurales entre variables. Sin embargo, la detección de anomalías en sistemas QKD depende en gran medida de cómo evolucionan estas métricas en el tiempo y de su coherencia mutua.

Dado que los modelos de aprendizaje automático no supervisado dependen de forma crítica de la representación del espacio de datos, se diseñó una fase específica y detallada de ingeniería de características. El objetivo no es aumentar artificialmente la dimensionalidad, sino dotar al modelo de memoria temporal, contexto operativo y relaciones físicas relevantes.

Categoría	**Variables Generadas**	**Justificación Físico-Operativa**
Contexto Temporal	Hora, Día semana	Captura ciclos térmicos y patrones de uso humano.
Memoria (Lags)	$t-1, t-3, t-5$ (min)	Modela la inercia térmica y operativa del sistema.
Dinámica	Δ SKR, Δ QBER	Detecta cambios abruptos frente a deriva lenta.
Estabilidad	Media/Var móvil (15 min)	Caracteriza el régimen de ruido local.

Cuadro 1: Resumen de las características sintéticas generadas durante la fase de ingeniería de características para dotar de contexto físico a los modelos no supervisados.

Características temporales y de contexto

Se incorporaron variables que describen el contexto temporal de cada medición, como la hora del día, el día de la semana, el día del año, un indicador binario de periodo nocturno y una segmentación temporal discreta del día. Estas variables permiten capturar patrones cíclicos asociados a variaciones ambientales, ciclos térmicos de la fibra o rutinas de operación del sistema, sin introducir conocimiento externo explícito. Datos que, *a priori* los *datasets* no nos dan explícitamente pero que sí se han generado de forma artificial para enriquecer la comprensión de los mismos.

Retardos temporales y memoria del sistema

Para cada métrica principal se generaron retardos temporales explícitos (*lags*) correspondientes a instantes anteriores ($t-1, t-2, \ldots, t-n$). Esta estrategia convierte el problema original, esencialmente instantáneo, en un problema con memoria explícita, permitiendo al modelo comparar el estado actual del sistema

con su historia reciente. Es, fundamentalmente, un análisis de la derivada de la serie temporal, lo que permite analizar su tendencia más o menos acentuada con el tiempo.

Desde un punto de vista operativo, esta memoria resulta fundamental para detectar cambios de régimen, rupturas de estabilidad o transiciones anómalas que no se manifiestan como valores extremos en un único instante, sino como desviaciones persistentes respecto al comportamiento histórico inmediato.

Derivadas discretas y cambios relativos

Específicamente, se aproximó la primera derivada en el tiempo de las métricas principales mediante diferencias finitas discretas. En particular, se calculó la derivada de la SKR como:

$$\frac{d\,\mathrm{SKR}}{dt} \approx \mathrm{SKR}(t) - \mathrm{SKR}(t-1).$$

Esta característica permite capturar la velocidad de cambio del sistema y detectar caídas abruptas de rendimiento incluso cuando el valor absoluto de la SKR sigue siendo formalmente aceptable según los umbrales del fabricante. De forma análoga, se calcularon variaciones relativas porcentuales, especialmente útiles en sistemas altamente estables donde los cambios absolutos son pequeños. Esta característica es fácilmente correlacionable con otras provenientes de otros sistemas de red cuántica, para deducir relaciones de causalidad entre ellas.

Ventanas móviles multiescala

Se calcularon estadísticas locales (media, desviación estándar, mínimo, máximo y mediana) sobre ventanas móviles de distintos tamaños. Este enfoque multiescala permite analizar la estabilidad del sistema a corto, medio y largo plazo, y detectar periodos de ruido anómalo, inestabilidad sostenida o comportamiento errático que no es visible a una única escala temporal.

Interacciones físicas entre métricas

Se introdujeron relaciones explícitas entre QBER, SKR y pérdidas ópticas, como cocientes y productos. Estas características permiten detectar incoherencias físicas, por ejemplo, situaciones en las que la relación histórica entre QBER y SKR se rompe sin una causa aparente, lo que constituye una señal de interés desde un punto de vista operativo.

Normalización global

Finalmente, se calcularon puntuaciones Z globales para cada métrica principal, proporcionando una referencia común respecto al comportamiento histórico completo del sistema. Aunque como hemos explicado anteriormente no es una métrica conclusiva, estas variables facilitan la detección de desviaciones globales y permiten comparar la severidad relativa de anomalías en distintos periodos.

4.4. Implementación de modelos de detección de anomalías no supervisados

Sobre el espacio de características enriquecido detallado en la sección anterior, se implementaron dos modelos complementarios de detección de anomalías: Isolation Forest y un *autoencoder* profundo. Ambos modelos se entrenan de forma independiente para cada fuente de datos, respetando las diferencias operativas entre plataformas.

El objetivo de estas técnicas es obtener una información más procesada de la que ofrecen las métricas anteriores, más cruda y, por ello, menos sujeta a la interpretación automatizada de anomalías.

Isolation Forest: configuración, entrenamiento y limitaciones

Isolation Forest se utiliza como detector de anomalías de carácter estadístico. El modelo se implementa mediante la biblioteca `scikit-learn` y se entrena siguiendo un enfoque de "one-class learning", aprendiendo el comportamiento normal directamente a partir de los propios datos operativos.

Se utilizaron 600 árboles (`n_estimators` = 600) para mejorar la estabilidad estadística del modelo, y un valor de contaminación conservador del $0{,}3\,\%$ (`contamination` = 0,003). Esta elección refleja el conocimiento del dominio: los sistemas QKD operan mayoritariamente en régimen estable y se espera que las anomalías reales sean poco frecuentes.

Antes del entrenamiento, las características numéricas se escalan mediante StandardScaler, garantizando que ninguna métrica domine el proceso de partición por su magnitud numérica. Asimismo, se aplica un periodo de *warm-up* inicial descartando las primeras muestras, evitando efectos transitorios de arranque (muchas de las pruebas iniciales fallaban, sobre todo con *autoencoders*, por esto mismo).

Isolation Forest resulta especialmente eficaz para detectar anomalías puntuales y desviaciones locales. No obstante, presenta limitaciones conocidas: su sensibilidad a degradaciones suaves y prolongadas es reducida, ya que estas pueden integrarse progresivamente en el espacio de normalidad aprendido.

Autoencoder: arquitectura, entrenamiento y limitaciones

Para complementar el análisis estadístico, se implementó un *autoencoder* denso utilizando PyTorch. Se diseñó un *autoencoder* profundo y simétrico. El codificador reduce progresivamente la dimensionalidad mediante capas densas con activaciones ReLU hasta un cuello de botella (*bottleneck*) de dimensión $K = 16$, forzando a la red a aprender las correlaciones más relevantes entre variables.

El decodificador reconstruye la entrada original a partir del espacio latente. La función de pérdida utilizada fue el error cuadrático medio (MSE por su siglas en inglés), y el entrenamiento se realizó con el optimizador Adam, utilizando una tasa de aprendizaje de 10^{-3}, durante 25 épocas y con lotes de tamaño 128. La Figura 6 muestra su arquitectura.

El *autoencoder* es especialmente sensible a anomalías estructurales y rupturas multivariantes prolongadas. Sin embargo, requiere un volumen suficiente de datos para generalizar correctamente y puede absorber degradaciones lentas como comportamiento normal si estas dominan el conjunto de entrenamiento.

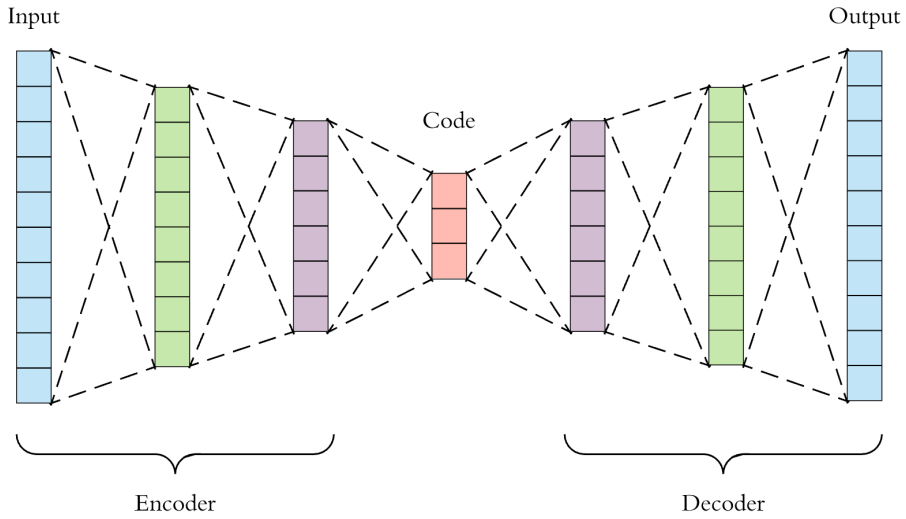


Figura 6: Arquitectura conceptual de un *autoencoder* profundo simétrico. La reducción de dimensionalidad en la capa central (*Code*) fuerza al modelo a aprender las estructuras latentes más relevantes de los datos de entrada, descartando el ruido no correlacionado.

Consideraciones sobre el uso de umbrales estáticos y recalibración del modelo

En el sistema propuesto, tanto el *autoencoder* como el modelo de Isolation Forest utilizan criterios de decisión basados en umbrales fijados *a priori* a partir del propio conjunto de datos. En el caso del *autoencoder*, una muestra se considera anómala cuando su error de reconstrucción supera un umbral definido como un percentil alto (en este trabajo, el percentil 99,5 %) de la distribución de errores observada durante la fase de entrenamiento. De forma análoga, en Isolation Forest la fracción esperada de anomalías se controla mediante el parámetro `contamination`, fijado de manera conservadora para cada fuente de datos.

Este enfoque resulta razonable en un contexto de análisis exploratorio y validación inicial, ya que permite establecer un criterio claro, reproducible y fácil de interpretar para la detección de desviaciones respecto al comportamiento normal aprendido por el modelo. No obstante, el uso de umbrales estáticos introduce limitaciones importantes cuando se considera el despliegue del sistema en infraestructuras QKD reales operando de forma continuada a lo largo del tiempo.

Las infraestructuras de red cuántica presentan fenómenos de evolución lenta y acumulativa, como la deriva estacional del canal óptico, el envejecimiento progresivo de componentes, ajustes periódicos de calibración o cambios en las

condiciones ambientales. Estos efectos pueden modificar gradualmente la distribución de las métricas operativas (por ejemplo, QBER o SKR) sin que exista un evento abrupto claramente identificable. En este escenario, un umbral fijo definido a partir de datos históricos puede volverse progresivamente inadecuado: o bien generará un aumento de falsos positivos al desplazarse la distribución normal, o bien perderá sensibilidad frente a degradaciones relevantes que queden integradas en el nuevo régimen operativo.

Este problema no es exclusivo del *autoencoder*, sino que afecta también al modelo de Isolation Forest. Aunque este último no utiliza un umbral explícito sobre una métrica física, el parámetro `contamination` actúa textitde facto como un umbral estadístico que determina qué fracción del comportamiento observado se considera anómala. Si la distribución subyacente del sistema cambia de forma significativa, un valor fijo de `contamination` puede dejar de reflejar adecuadamente la realidad operativa del enlace.

En el presente trabajo, esta limitación se asume de forma explícita. El objetivo no es proponer un sistema completamente autónomo y auto-adaptativo, sino analizar hasta qué punto técnicas no supervisadas entrenadas sobre datos reales son capaces de capturar el comportamiento normal y detectar desviaciones coherentes desde un punto de vista físico-operativo en ventanas temporales acotadas. En este contexto, el uso de umbrales estáticos permite aislar el comportamiento intrínseco de los modelos y facilita la interpretación de los resultados obtenidos en los demostradores experimentales.

No obstante, la recalibración periódica de los modelos y de sus umbrales constituye una línea de trabajo futura imprescindible para un despliegue real. Entre las estrategias posibles se incluyen el reentrenamiento regular sobre ventanas deslizantes de datos recientes, el uso de umbrales adaptativos basados en estadísticos robustos o la combinación con técnicas específicas de detección de cambio de régimen y deriva a largo plazo. Estas extensiones permitirían complementar la detección de anomalías abruptas abordada en este trabajo con mecanismos orientados al mantenimiento predictivo y a la adaptación continua del sistema.

Consideraciones sobre coste computacional y modo de operación

Aunque el objetivo principal de este trabajo es el análisis operativo y la detección de anomalías en infraestructuras QKD, resulta imprescindible discutir explícitamente el coste computacional del *pipeline* propuesto y su adecuación a distintos modos de operación. En entornos reales, la viabilidad de un sistema de monitorización no depende únicamente de su capacidad de detección, sino también de los recursos necesarios para ejecutarlo y del retardo introducido entre la observación de los datos y la generación de alertas.

El *pipeline* desarrollado en este trabajo sigue un enfoque claramente *batch-oriented*. El procesamiento completo de una ventana de datos —incluyendo carga, preprocesado, ingeniería de características, inferencia con Isolation Forest y *autoencoder*, y generación de salidas— tiene un tiempo de ejecución del orden

de un minuto en el entorno de pruebas utilizado. Este valor incluye tanto la inferencia de los modelos como el procesamiento previo necesario para garantizar coherencia y robustez en los datos de entrada.

Este tiempo de ejecución descarta explícitamente un uso en escenarios de tiempo real estricto o control en lazo cerrado. No obstante, esta limitación es coherente con el contexto operativo de las infraestructuras QKD actuales. En estos sistemas, las métricas clave (como QBER o SKR) se evalúan típicamente sobre ventanas temporales agregadas, asociadas a ciclos de reconciliación y amplificación de privacidad, y no a nivel de eventos individuales de alta frecuencia. En consecuencia, las decisiones operativas y de mantenimiento no requieren latencias del orden de milisegundos, sino una visión robusta y contextualizada del estado del enlace.

Desde este punto de vista, un tiempo de procesamiento del orden de decenas de segundos o minutos resulta compatible con un esquema de monitorización "*near real-time*", orientado a la supervisión continua, al soporte a operadores humanos y a la detección temprana de degradaciones relevantes. Además, el diseño modular del *pipeline* permitiría, en un entorno de producción, optimizar o paralelizar determinadas etapas, así como separar el procesamiento pesado en procesos asíncronos o programados.

Por tanto, el sistema propuesto no se plantea como un mecanismo de reacción inmediata, sino como una herramienta de análisis operativo y diagnóstico, alineada con los tiempos y necesidades reales de explotación de redes QKD. La reducción de latencia y la adaptación a escenarios de *streaming* constituirían una extensión natural del trabajo, pero quedan fuera del alcance de este Trabajo Fin de Grado.

4.5. Fusión de modelos y lógica de decisión operativa

Dado que ambos modelos presentan fortalezas y debilidades claramente complementarias, se diseñó un mecanismo explícito de fusión de resultados. Para cada observación se generan dos indicadores binarios: uno procedente de Isolation Forest y otro del *autoencoder*.

En el caso del *autoencoder*, se define un umbral de anomalía basado en el percentil $99{,}5$ del error de reconstrucción, asumiendo que solo una fracción muy pequeña de las observaciones corresponde a comportamientos anómalos severos.

Las salidas se combinan mediante una lógica de decisión explícita:

- **NONE**: ningún modelo detecta anomalía.
- **IF_ONLY**: solo Isolation Forest detecta anomalía.
- **AE_ONLY**: solo el *autoencoder* detecta anomalía.
- **BOTH**: ambos modelos detectan anomalía simultáneamente.

La categoría **BOTH** se interpreta como una anomalía de alta confianza y constituye el criterio de "alarma confirmada". Esta estrategia reduce la generación de falsos positivos críticos y proporciona una visión más matizada del estado del enlace cuántico, alineada con las necesidades reales de operación.

Las implicaciones prácticas de esta fusión, así como el análisis detallado de los eventos detectados, se presentan y discuten en los capítulos de resultados y validación experimental.

5. Evaluación, discusión y resultados operacionales

En este capítulo se evalúa el comportamiento del sistema de monitorización propuesto desde un punto de vista explícitamente operativo, alineado con las condiciones reales de funcionamiento de infraestructuras de distribución cuántica de clave. A diferencia de los escenarios clásicos de aprendizaje supervisado, habituales en muchos problemas académicos de detección de anomalías, en infraestructuras QKD reales no se dispone de etiquetas fiables que permitan calcular métricas estándar como la exactitud, el AUC o el F1-score frente a una verdad terreno conocida.

Esta ausencia de etiquetas no constituye una limitación puntual del presente trabajo, sino una característica estructural del dominio de aplicación. En sistemas QKD operacionales, los eventos anómalos relevantes son poco frecuentes, no siempre tienen una causa única bien definida y, en muchos casos, no pueden reproducirse de forma controlada sobre la infraestructura física por motivos de seguridad, coste y disponibilidad del servicio. En consecuencia, la evaluación del sistema debe apoyarse en criterios distintos a los utilizados en problemas de clasificación supervisada tradicionales.

Por este motivo, la validación se realiza mediante **demostradores operativos controlados**, diseñados específicamente para reproducir patrones anómalos físicamente plausibles sobre datos reales procedentes de sistemas QKD comerciales. Este enfoque permite evaluar el sistema en condiciones realistas, respetando las restricciones del dominio y evitando introducir supuestos artificiales que no se cumplirían en un despliegue operativo.

En concreto, este capítulo analiza si el sistema es capaz de:

- detectar perturbaciones relevantes del sistema que afectarían a la seguridad o a la disponibilidad de la clave.
- distinguir entre distintos tipos de anomalías, tanto de origen cuántico como asociadas al plano clásico.
- responder de forma coherente con el modelo físico-operativo de QKD.
- evitar sobrerreacciones ante ruido transitorio o inestabilidad no estructural.

5.1. Consideraciones sobre métricas de rendimiento clásicas

En problemas de detección de anomalías no supervisada, y especialmente en el contexto de infraestructuras críticas como las QKD, el uso de métricas clásicas de clasificación (*accuracy*, AUC, *precision*, *recall* o F1-score) presenta limitaciones fundamentales que deben ser explícitamente reconocidas.

Estas métricas presuponen la existencia de una verdad terreno objetiva y completamente etiquetada que permita distinguir de forma inequívoca entre estados normales y anómalos. En escenarios operativos reales de QKD, esta condición no se cumple por razones inherentes al propio dominio:

- no existen bases de datos públicas ni privadas con ataques reales o fallos críticos etiquetados de forma sistemática.
- muchas anomalías operativas no tienen una causa única claramente delimitada, sino que emergen de la interacción compleja entre factores físicos, ambientales y de control.
- la frontera entre comportamiento normal y anómalo no es fija, sino dependiente del contexto temporal, del régimen operativo y de las políticas de explotación del sistema.

Desde una perspectiva estrictamente ingenieril, forzar el uso de métricas clásicas en este contexto conduciría a una falsa sensación de rigor cuantitativo, ocultando en realidad la ausencia de una referencia objetiva. Por este motivo, el criterio de éxito adoptado en este trabajo no es la optimización de una métrica de clasificación abstracta, sino la coherencia semántica y operativa del sistema: es decir, si las anomalías detectadas corresponden a situaciones que un operador humano con experiencia en QKD consideraría razonablemente problemáticas o dignas de atención.

Este enfoque de evaluación, aunque menos habitual en entornos académicos puramente algorítmicos, se alinea de forma mucho más fiel con las necesidades reales de operación y mantenimiento de infraestructuras QKD, donde el objetivo no es clasificar cada muestra de forma aislada, sino apoyar la toma de decisiones informadas y reducir riesgos operativos.

5.2. Comparativa conceptual de métodos de detección

El sistema desarrollado permite comparar tres enfoques conceptualmente distintos para la detección de anomalías, cada uno con supuestos, fortalezas y limitaciones claramente diferenciados:

- **Z-score**: detección estadística univariante basada en desviaciones respecto a la media histórica, con el *script* de R.
- **Isolation Forest**: detección de rarezas estadísticas en espacios de alta dimensionalidad mediante particiones aleatorias.
- ***Autoencoder***: detección de incoherencias estructurales a partir del error de reconstrucción multivariante.

El análisis exploratorio inicial confirmó que el enfoque basado en Z-score resulta insuficiente en sistemas QKD reales. Las distribuciones de las métricas presentan colas pesadas, asimetrías, no estacionariedad y cambios de régimen que violan los supuestos implícitos de normalidad. Como consecuencia, este método tiende a generar falsos positivos en situaciones normales o, alternativamente, a ignorar degradaciones relevantes si los umbrales se relajan en exceso.

Isolation Forest mejora de forma significativa esta capacidad al incorporar contexto multivariante y no depender de supuestos paramétricos explícitos. Su fortaleza principal reside en la detección de anomalías puntuales y estadísticamente raras. No obstante, presenta limitaciones frente a degradaciones suaves y prolongadas que se integran progresivamente en el espacio de normalidad aprendido.

Por su parte, el *autoencoder* permite detectar anomalías estructurales que no se manifiestan como valores extremos individuales, sino como combinaciones incoherentes de métricas. La combinación de ambos enfoques permite cubrir un espectro más amplio de comportamientos anómalos, incrementando la robustez del sistema y reduciendo falsos positivos críticos.

5.3. Validación experimental mediante demostradores operativos

Motivación y defensa metodológica

Desde un punto de vista ideal, la validación de un sistema de monitorización QKD debería realizarse mediante experimentos controlados sobre infraestructura física real, introduciendo perturbaciones como variaciones térmicas, desalineaciones ópticas, cambios en la atenuación del canal o congestión inducida del plano clásico.

No obstante, por limitaciones temporales, logísticas y de disponibilidad del laboratorio durante el desarrollo de este Trabajo Fin de Grado, no fue posible llevar a cabo este tipo de pruebas físicas controladas. Por temas de tiempo con los profesionales del laboratorio, se decidió hacer estas demostraciones de forma artificial. Siempre basándose en posibles anomalías reales y cómo afectarían a este tipo de sistemas. El tutor y yo estuvimos debatiendo cómo hacer este tipo de demostraciones que, en un principio, iban a ser desde un plano clásico (con la librería Scapy de Python) simular estas anomalías y luego convertirlas al plano cuántico. Finalmente se decidió mediante *scripts* de Python (`src/python/demostradores/`) usando la librería Pandas manipular nuestros propios *datasets*.

Se diseñaron estos demostradores artificiales basados en la inyección controlada de perturbaciones sobre *datasets* los reales procedentes de sistemas QKD comerciales. Estas perturbaciones no son aleatorias ni arbitrarias, sino que modelan fenomenología física y operativa bien documentada en la literatura y observada en despliegues reales. Aunque la inyección sea sintética, la firma de datos resultante es realista.

5.4. Demostrador I: Saturación del plano clásico y degradación de la disponibilidad de clave

Tipo de anomalía esperable desde un punto de vista físico-operativo

En una infraestructura QKD operativa, la disponibilidad efectiva de clave segura no depende únicamente del canal cuántico, sino también de forma crítica del correcto funcionamiento del plano clásico. Componentes como el sistema de gestión de claves, los mecanismos de sincronización, los procesos de post-procesado o la capacidad de intercambio de información clásica pueden convertirse en cuellos de botella operativos.

Desde un punto de vista físico-operativo, una degradación severa del plano clásico puede manifestarse como una caída abrupta de la tasa de clave segura (SKR), incluso cuando las métricas cuánticas fundamentales, como el QBER, permanecen dentro de rangos normales. Este tipo de situación no corresponde a una perturbación directa del canal cuántico, sino a una incoherencia funcional entre ambos planos del sistema híbrido cuántico-clásico.

Por tanto, la firma esperable de este tipo de anomalía es una ruptura clara de la relación estructural entre QBER y SKR: el sistema continúa generando detecciones cuánticas con un nivel de error compatible con operación normal, pero la disponibilidad final de clave se degrada de forma abrupta debido a limitaciones o saturaciones en el procesamiento clásico.

Escenario analizado y datos utilizados

Para evaluar la capacidad del sistema de monitorización propuesto frente a este tipo de comportamiento, se analizó el *dataset* operativo correspondiente al sistema Toshiba-2025-W27. Este *dataset* refleja el funcionamiento de un sistema QKD de alto rendimiento en un entorno real de red, con valores de SKR del orden de cientos de kilobits por segundo y un QBER relativamente estable a lo largo del tiempo.

Sobre este conjunto de datos se introdujo una perturbación controlada y localizada temporalmente, consistente en una degradación severa de la tasa de clave segura acompañada de una variación moderada del QBER. El objetivo no es reproducir un evento concreto, sino generar una situación operativamente plausible que modele una saturación o mal funcionamiento del plano clásico, manteniendo el canal cuántico razonablemente estable.

Este escenario permite evaluar si el *pipeline* es capaz de identificar una anomalía estructural relevante sin apoyarse en valores extremos aislados y sin necesidad de etiquetas explícitas.

Respuesta del autoencoder: detección de incoherencia estructural

La Figura 7 muestra la evolución temporal del error de reconstrucción del *autoencoder* entrenado sobre el comportamiento normal del sistema Toshiba-2025-W27. Durante la mayor parte del periodo analizado, el error de reconstrucción

permanece bajo y estable, lo que indica que el modelo es capaz de reconstruir correctamente los patrones multivariantes habituales del sistema.

Sin embargo, en el intervalo correspondiente a la perturbación introducida, se observa un incremento abrupto y muy significativo del error de reconstrucción, superando ampliamente el umbral definido en el percentil 99 %. Este comportamiento indica que el estado del sistema durante ese intervalo no es compatible con la estructura aprendida por el *autoencoder*, y por tanto constituye una anomalía multivariante de alta severidad.

Cabe destacar que esta detección no se basa en una única métrica, sino en la incoherencia conjunta entre varias variables y sus relaciones temporales, lo que refuerza la robustez del resultado.

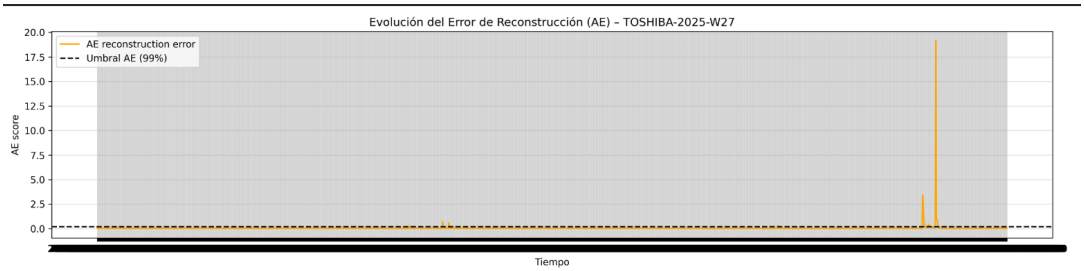


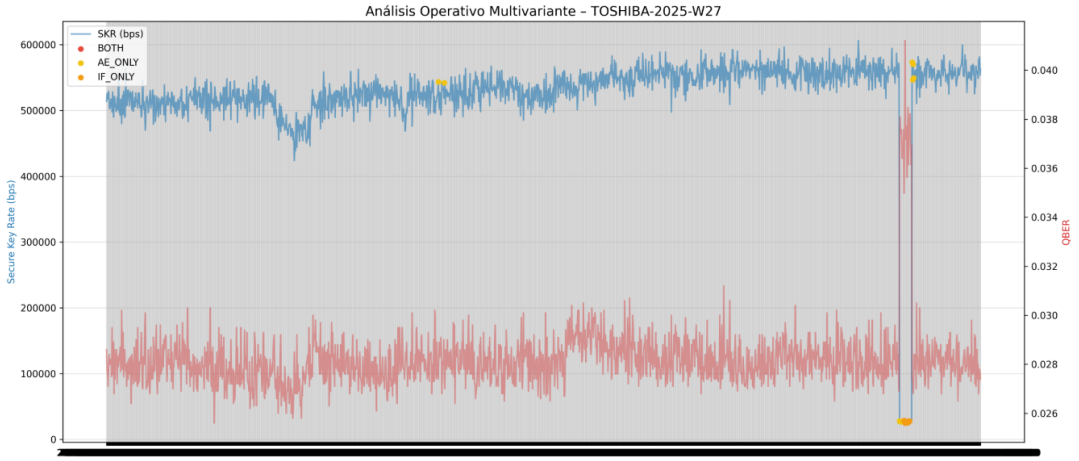
Figura 7: Evolución temporal del error de reconstrucción del *autoencoder* para el *dataset* Toshiba-2025-W27. El incremento abrupto del error durante el intervalo perturbado indica una ruptura clara del patrón operativo normal del sistema.

Análisis multivariante conjunto QBER/SKR

La Figura 5.4 muestra la evolución temporal conjunta de la SKR y el QBER, junto con las detecciones generadas por los distintos modelos. Durante el funcionamiento normal, ambas métricas presentan una relación estable: pequeñas fluctuaciones del QBER se traducen en variaciones suaves de la SKR, coherentes con la operación normal del sistema.

En el intervalo anómalo, esta relación se rompe de forma abrupta. La SKR cae varios órdenes de magnitud, alcanzando valores cercanos a cero, mientras que el QBER no experimenta un incremento proporcional ni sostenido. Esta falta de correlación es incompatible con una degradación del canal cuántico y apunta claramente a un problema de naturaleza funcional en el plano clásico.

El sistema de monitorización clasifica este intervalo como una anomalía de alta confianza (BOTH), al ser detectada simultáneamente por Isolation Forest y por el *autoencoder*. Esta coincidencia refuerza la interpretación operativa del evento como una situación crítica que requiere atención inmediata.



Espacio de fase QBER/SKR

El análisis se completa con la representación en el espacio de fase QBER/SKR, mostrada en la Figura 8. En este espacio, el comportamiento normal del sistema se concentra en una región compacta caracterizada por valores elevados de SKR y un rango acotado de QBER.

Las observaciones correspondientes al intervalo perturbado aparecen claramente separadas de esta región, desplazándose hacia una zona de SKR muy reducida sin un incremento significativo del QBER. Esta separación geométrica refuerza la interpretación de que el sistema se encuentra en un régimen operativo distinto y no compatible con el funcionamiento normal.

Desde un punto de vista operativo, esta visualización resulta especialmente útil, ya que permite identificar de forma inmediata incoherencias funcionales que no serían evidentes mediante el análisis univariante de cada métrica por separado.

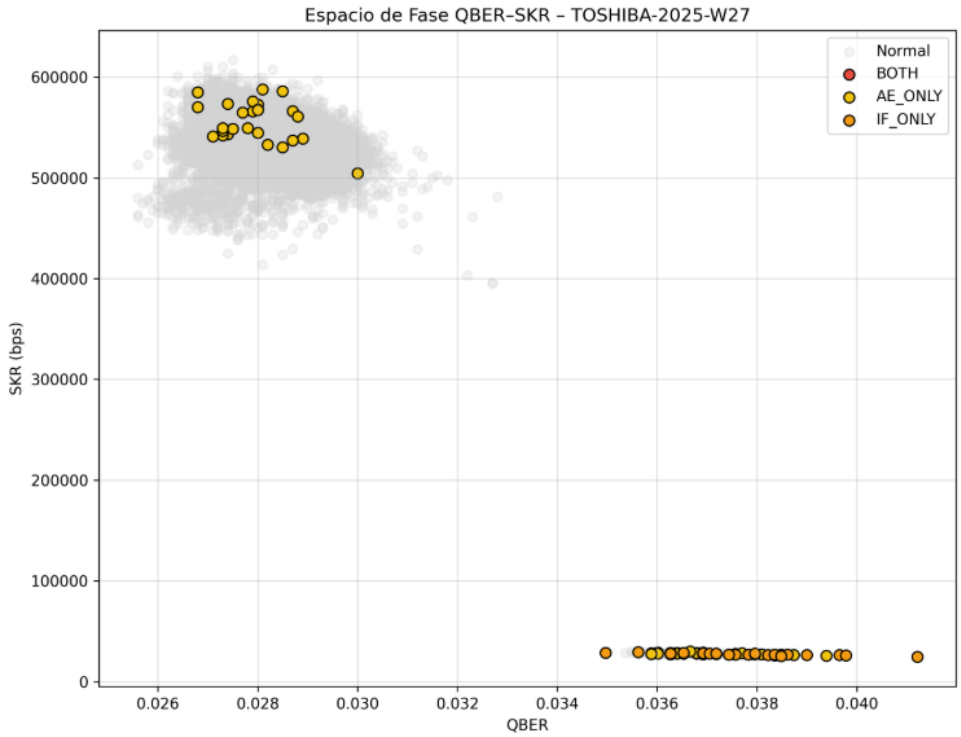


Figura 8: Espacio de fase QBER/SKR para el sistema Toshiba-2025-W27. Las observaciones anómalas se separan claramente del clúster de operación normal, mostrando una caída severa de la SKR sin degradación cuántica proporcional.

Discusión operativa

Este demostrador pone de manifiesto una de las principales fortalezas del sistema propuesto: su capacidad para detectar anomalías funcionales del plano clásico que se manifiestan como perturbaciones directas del canal cuántico. El *pipeline* no reacciona ante fluctuaciones normales del QBER, pero identifica de forma consistente situaciones en las que la disponibilidad de clave se degrada de manera incompatible con el comportamiento físico esperado.

5.5. Demostrador II: Degradación progresiva del canal por deriva óptica

Contexto físico-operativo

En infraestructuras QKD reales, no todas las degradaciones relevantes se manifiestan como eventos abruptos o discontinuidades claras en las métricas operativas. Un caso especialmente habitual corresponde a procesos de degradación lenta y progresiva del canal cuántico, compatibles con fenómenos como la deriva térmica de la fibra óptica, el envejecimiento gradual de componentes fotónicos o pequeñas desalineaciones acumulativas en el sistema óptico.

Este tipo de degradación se caracteriza por variaciones suaves y continuas de las métricas cuánticas, que permanecen dentro de rangos operativos aceptables

durante largos periodos de tiempo. Desde un punto de vista físico, estos procesos no introducen perturbaciones bruscas ni rupturas funcionales, sino que desplazan lentamente el régimen operativo del enlace, modificando de forma progresiva la relación entre parámetros como la tasa de error cuántico y la tasa de clave segura.

Desde el punto de vista de la operación, este escenario representa una situación potencialmente problemática a medio o largo plazo, pero que no requiere una respuesta inmediata. En la práctica, este tipo de degradaciones se gestionan mediante tareas de mantenimiento programado, recalibración del sistema o ajustes preventivos, y no mediante alarmas críticas en tiempo real. Por este motivo, constituye un caso límite especialmente interesante para evaluar no solo la sensibilidad del sistema propuesto, sino también su capacidad para no generar alertas innecesarias.

Descripción del escenario analizado

Para reproducir este comportamiento de forma controlada, se aplicó una modificación progresiva y continua sobre un subconjunto temporal del *dataset* `Toshiba-2025-W27`. En una ventana de operación acotada, se introdujo una degradación gradual consistente en:

- una reducción progresiva de la tasa de clave segura, alcanzando una caída acumulada del orden del 40 %.
- un incremento suave y sostenido del QBER, con un aumento acumulado de magnitud comparable.

Es importante subrayar que esta modificación no introduce discontinuidades abruptas ni valores extremos aislados. Todas las observaciones permanecen dentro de rangos físicamente plausibles y compatibles con el funcionamiento normal de un sistema QKD en producción. Desde un punto de vista estadístico, no se generan *outliers* evidentes en un instante concreto, sino una transición gradual entre dos estados operativos cercanos.

Este diseño permite evaluar explícitamente la respuesta del *pipeline* ante un fenómeno de deriva lenta, diferenciándolo de los escenarios de anomalías abruptas analizados en otros demostradores.

Respuesta del autoencoder y análisis del error de reconstrucción

La Figura 9 muestra la evolución temporal del error de reconstrucción del *autoencoder* durante el periodo analizado. Se observa que, a lo largo de la fase de degradación progresiva, el error de reconstrucción permanece mayoritariamente por debajo del umbral definido (percentil 99,5), sin generar detecciones persistentes de alta severidad (el pico que se ve en la figura se trata del error al reconstruir a estados normales tras la degradación, no a la propia degradación en sí).

Este comportamiento es coherente con el diseño y los supuestos del modelo. El *autoencoder* ha sido entrenado para aprender la estructura latente del com-

portamiento normal del sistema, y una transición lenta y continua puede ser absorbida progresivamente como parte de dicho comportamiento normal, siempre que no introduzca incoherencias abruptas entre las métricas.

Este resultado confirma que el *autoencoder* es especialmente eficaz para detectar rupturas estructurales claras, pero no está diseñado para actuar como un detector primario de deriva lenta a largo plazo. Esta limitación no debe interpretarse como un fallo del modelo, sino como una consecuencia directa de su objetivo que es priorizar la detección de situaciones que requieren una acción operativa inmediata.

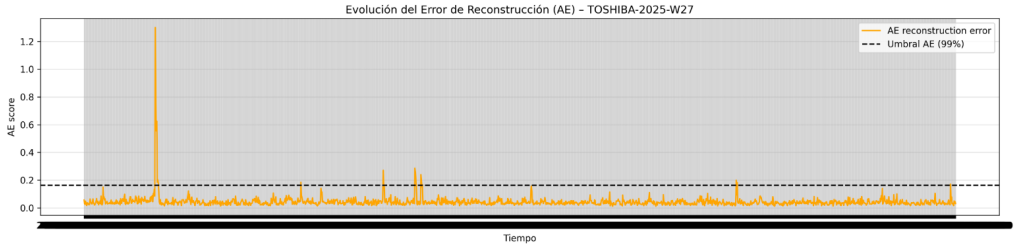


Figura 9: Evolución del error de reconstrucción del *autoencoder* durante un escenario de degradación progresiva. Cabe decir que el error que se ve es cuando se vuelven a estados “normales” .

Análisis multivariante temporal

La Figura 10 muestra la evolución conjunta de la SKR y el QBER a lo largo del tiempo, junto con las etiquetas de detección generadas por el sistema. Se observa una tendencia clara de disminución progresiva de la SKR acompañada de un incremento sostenido del QBER, sin discontinuidades abruptas ni episodios de inestabilidad extrema en un corto periodo de tiempo.

El *pipeline* no clasifica este patrón como una anomalía crítica persistente. Desde un punto de vista operativo, este comportamiento es coherente y deseable: en ausencia de una ruptura funcional clara o de la superación de umbrales de seguridad explícitos, el sistema evita generar alarmas críticas ante un fenómeno que no compromete de forma inmediata la disponibilidad ni la seguridad del enlace.

Este resultado pone de manifiesto una característica clave del sistema propuesto: su capacidad para distinguir entre anomalías de acción inmediata y procesos de degradación gradual que deben ser gestionados mediante mantenimiento y supervisión a largo plazo, y no mediante alertas operativas urgentes.

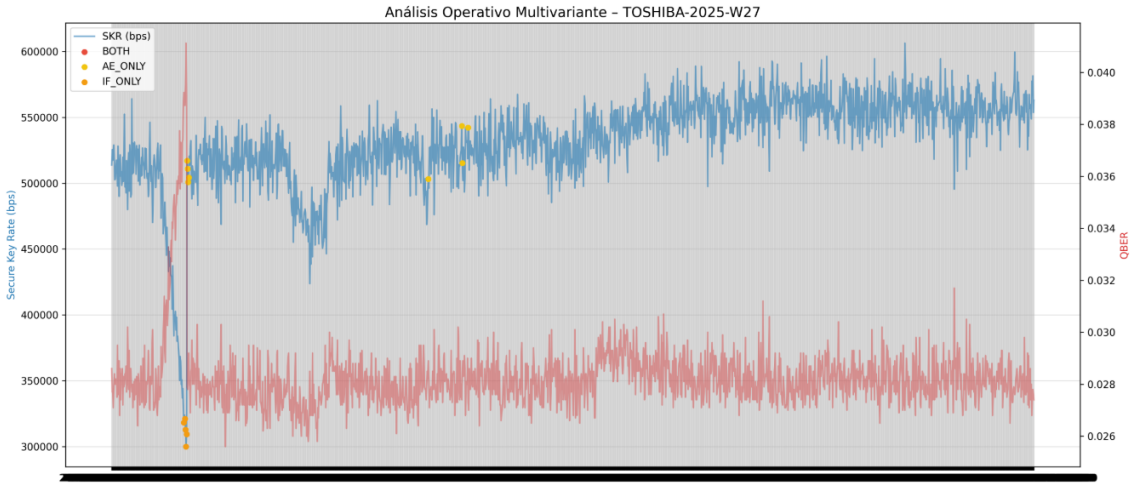


Figura 10: Evolución temporal conjunta de SKR y QBER durante un escenario de deriva lenta. El sistema mantiene una clasificación mayoritariamente normal, evitando falsas alarmas ante degradaciones progresivas hasta el final de esta.

Análisis en el espacio de fase QBER/SKR

El análisis en el espacio de fase QBER/SKR, mostrado en la Figura 11, proporciona una interpretación geométrica adicional del fenómeno. En este plano, el comportamiento normal del sistema ocupa una región densa bien definida, mientras que la degradación progresiva se manifiesta como un desplazamiento continuo dentro de dicha región, sin saltos abruptos hacia zonas claramente separadas.

Las observaciones correspondientes al periodo de deriva lenta siguen una trayectoria coherente con la relación física esperada entre QBER y SKR. Esta continuidad explica que no sean clasificadas sistemáticamente como anomalías de alta severidad. Desde un punto de vista físico-operativo, el sistema no está experimentando una disrupción funcional, sino una transición gradual entre estados operativos cercanos. Cabe destacar la diferencia en la figura entre los valores IFONLY y AEONLY.

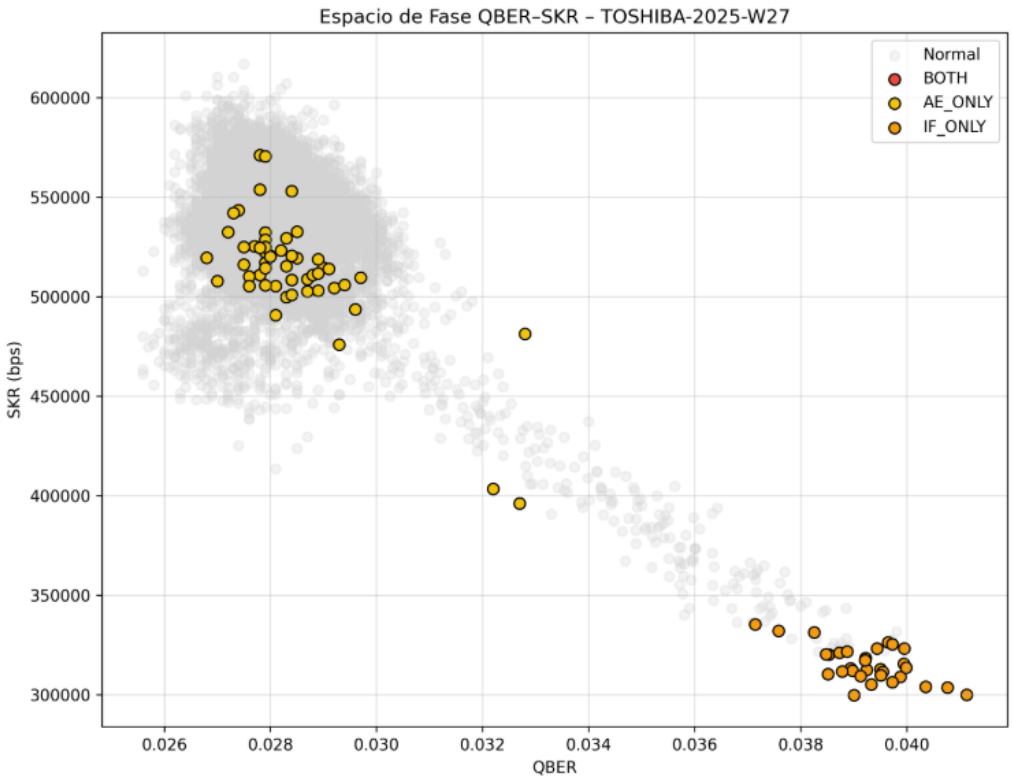


Figura 11: Espacio de fase QBER/SKR durante un escenario de degradación progresiva. El desplazamiento continuo dentro de la región normal explica la ausencia de detecciones abruptas.

Discusión y delimitación del ámbito de aplicación

Este demostrador pone de manifiesto una distinción clave entre la detección de anomalías y la detección de deriva. El sistema desarrollado en este trabajo está diseñado explícitamente para identificar eventos abruptos, rupturas estructurales e incoherencias multivariantes que requieren una intervención operativa inmediata. Desde este punto de vista, la ausencia de alarmas persistentes ante procesos de degradación lenta constituye un comportamiento coherente con los objetivos definidos para el sistema.

La detección de deriva a largo plazo responde a una problemática diferente, más vinculada al mantenimiento predictivo que a la gestión de incidentes operativos. Abordar este tipo de fenómenos exige mecanismos específicos, como el análisis de tendencias, el uso de modelos temporales explícitos o la aplicación de detectores de cambio de régimen, que operan a escalas temporales mayores y bajo criterios distintos a los empleados en la detección de anomalías abruptas.

En conjunto, este demostrador no pone de relieve una debilidad del *pipeline* propuesto, sino una delimitación adecuada de su ámbito de aplicación. El sistema muestra una elevada robustez frente a falsas alarmas y un comportamiento alineado con la práctica operativa real, lo que refuerza su utilidad como herra-

mienta de apoyo a la supervisión diaria de infraestructuras QKD.

5.6. Demostrador III: Perturbación directa del canal cuántico

Contexto físico-operativo

Una de las características fundamentales de la distribución cuántica de clave es que el canal cuántico es intrínsecamente sensible a cualquier perturbación que altere los estados transmitidos. Desde un punto de vista físico, cualquier proceso que introduzca ruido adicional en el canal se traduce inevitablemente en un incremento del QBER y, como consecuencia directa, en una reducción de la tasa de clave segura disponible tras el post-procesado clásico.

A diferencia de los escenarios analizados en demostradores anteriores, en este caso la anomalía se origina directamente en el canal cuántico y no en el plano clásico de control o gestión. Este tipo de perturbación modifica de forma explícita la relación física entre QBER y SKR, empujando al sistema hacia un régimen operativo degradado en el que la corrección de errores y la amplificación de privacidad consumen una fracción cada vez mayor de los bits intercambiados.

Desde el punto de vista operativo, este escenario representa una situación especialmente crítica, ya que afecta simultáneamente a la seguridad y a la disponibilidad de la clave, y suele manifestarse de forma coherente en múltiples métricas cuánticas.

Descripción del escenario analizado

Para reproducir este comportamiento, se analizó un periodo temporal acotado del *dataset* `Toshiba-2025-W27` en el que se introdujo una perturbación progresiva del canal cuántico. Durante este intervalo se observa:

- un incremento gradual y sostenido del QBER, desplazando al sistema fuera de su régimen nominal de operación.
- una reducción correlativa de la SKR, coherente con el aumento del coste asociado a la corrección de errores y a la amplificación de privacidad.

Es importante subrayar que este escenario no consiste en un único pico aislado, sino en una evolución temporal consistente, en la que ambas métricas se degradan de forma acoplada siguiendo la relación física esperable en protocolos QKD prácticos.

Respuesta del autoencoder

La Figura 12 muestra la evolución del error de reconstrucción del *autoencoder* durante el periodo analizado. En contraste con el escenario de deriva lenta presentado en el Demostrador II, en este caso el error de reconstrucción supera de forma clara y persistente el umbral definido, generando múltiples detecciones.

Este comportamiento indica que el *autoencoder* identifica una ruptura estructural en la relación multivariante entre las métricas. Aunque cada una de las

variables podría, de forma aislada, permanecer dentro de rangos técnicamente posibles, su combinación deja de ser coherente con el patrón aprendido durante el entrenamiento, lo que se refleja en un aumento significativo del error de reconstrucción.

Desde un punto de vista operativo, esta respuesta es especialmente relevante, ya que el modelo no reacciona únicamente a valores extremos puntuales, sino a la pérdida de coherencia física del sistema.

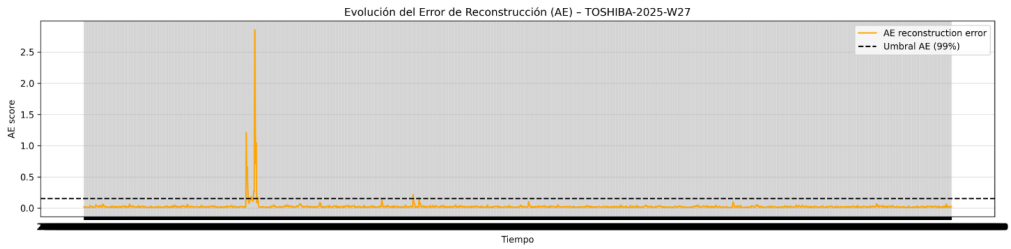


Figura 12: Evolución del error de reconstrucción del *autoencoder* durante una perturbación directa del canal cuántico. El aumento sostenido del error refleja una ruptura estructural clara en el comportamiento del sistema.

Análisis multivariante temporal

La evolución temporal conjunta de la SKR y el QBER se muestra en la Figura 13. Durante el periodo de perturbación se observa un incremento pronunciado del QBER acompañado de una caída significativa de la SKR, formando un patrón temporal consistente y claramente diferenciado del régimen normal de operación.

En este escenario, ambos modelos de detección —Isolation Forest y Autoencoder— identifican de forma consistente las observaciones correspondientes como anomalías de alta confianza (BOTH). Este resultado refleja que la degradación no solo es estadísticamente rara, sino también estructuralmente incoherente con el comportamiento normal del sistema.

Desde una perspectiva operativa, esta clasificación resulta coherente con la severidad del fenómeno: el sistema entra en un régimen en el que la generación de clave segura se ve fuertemente limitada y la estabilidad del enlace queda comprometida.

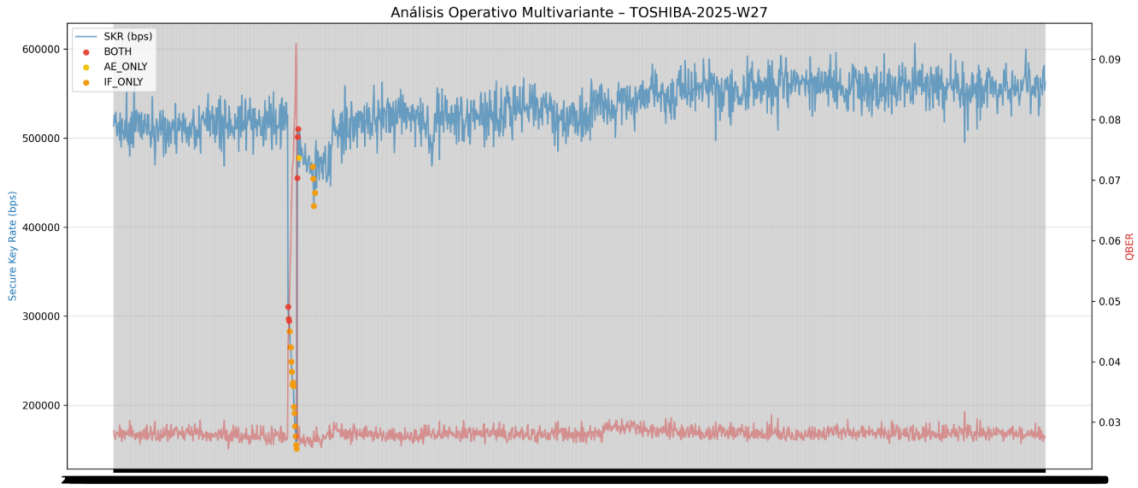


Figura 13: Evolución temporal conjunta de QBER y SKR durante una perturbación directa del canal cuántico. Ambos modelos detectan de forma consistente un estado operativo anómalo.

Análisis en el espacio de fase QBER/SKR

El análisis en el espacio de fase QBER/SKR, representado en la Figura 14, proporciona una interpretación geométrica clara del fenómeno. Mientras que el régimen normal de operación ocupa una región compacta y bien definida, las observaciones correspondientes al periodo de perturbación se desplazan progresivamente hacia una zona claramente separada del espacio.

Este desplazamiento sigue una trayectoria físicamente coherente: a medida que el QBER aumenta, la SKR disminuye de forma casi monótona, reflejando la relación fundamental entre ambas métricas en sistemas QKD prácticos. La separación clara entre ambas regiones explica por qué los modelos identifican estas observaciones como anomalías de alta confianza.

Desde el punto de vista del operador, esta representación facilita una interpretación inmediata del estado del sistema, permitiendo distinguir entre fluctuaciones normales y degradaciones severas del canal cuántico.

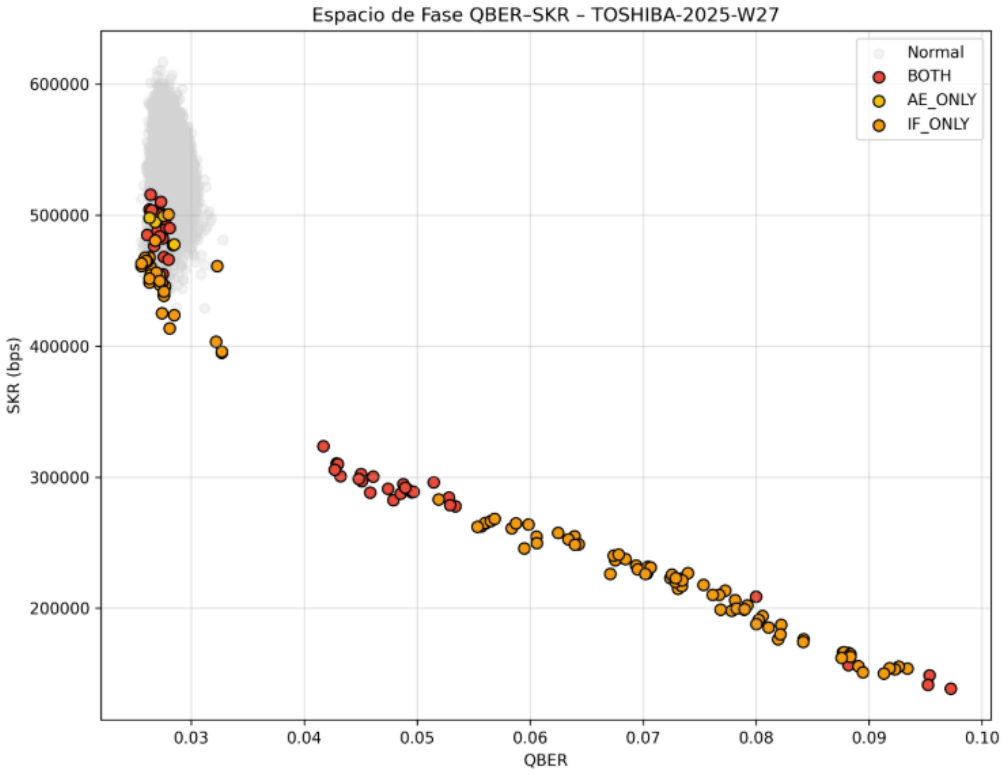


Figura 14: Espacio de fase QBER/SKR durante una perturbación del canal cuántico. La separación clara respecto al régimen normal refleja una degradación física significativa del enlace.

Discusión operativa

Este demostrador constituye uno de los casos más representativos desde el punto de vista físico y valida de forma clara la coherencia del sistema de monitorización propuesto con los principios fundamentales de la QKD. A diferencia de escenarios de deriva lenta o de perturbaciones puramente clásicas, la degradación directa del canal cuántico introduce una firma multivariante inequívoca que es capturada de forma consistente por ambos modelos.

El hecho de que la detección sea simultánea y persistente refuerza la utilidad del esquema de fusión propuesto, ya que permite discriminar situaciones realmente críticas de variaciones benignas del sistema. En este contexto, el sistema actúa como una herramienta eficaz para identificar estados operativos en los que la generación de clave segura deja de ser viable de forma sostenida.

En conjunto, los resultados obtenidos en este demostrador refuerzan la validez del enfoque adoptado y su alineación con el comportamiento físico esperado de infraestructuras QKD reales.

5.7. Demostrador IV: Inestabilidad estocástica del plano clásico

Contexto físico-operativo

En infraestructuras QKD reales, no todas las perturbaciones observadas en las métricas del sistema responden a fallos estructurales o degradaciones persistentes. Es relativamente habitual que, durante la operación normal, aparezcan episodios breves de inestabilidad asociados al plano clásico: picos de latencia, pérdidas puntuales de sincronización, congestión en red de control, o procesos internos automáticos (reajustes, reintentos, etc.).

Este tipo de comportamiento suele ser estocástico, con variaciones rápidas y poco correlacionadas entre sí. Desde un punto de vista operativo, lo importante aquí no es çazarçada pico, sino evitar que un sistema de monitorización se vuelva demasiado sensible y genere alertas sostenidas ante ruido transitorio que, en la práctica, no requiere intervención inmediata.

El objetivo de este demostrador es precisamente evaluar la robustez del *pipeline* frente a este tipo de inestabilidad no estructural y analizar si aparecen falsos positivos persistentes.

Escenario de validación y perturbación aplicada

Para reproducir un escenario de este estilo se introdujo una ventana temporal de perturbación (aprox. filas 8400 a 8600 en el *dataset* Toshiba-2025-W27) donde, en cada muestra, tanto la SKR como el QBER se modifican mediante factores aleatorios independientes:

- la SKR se multiplica por un factor uniforme en $[0{,}2,\ 0{,}9]$, generando caídas y recuperaciones bruscas
- el QBER se multiplica por un factor uniforme en $[1{,}0,\ 2{,}5]$, introduciendo picos intermitentes compatibles con desincronización o jitter del plano clásico.

En este demostrador queríamos capturar una tormenta de comportamiento irregular: muchas oscilaciones cortas, sin una deriva suave y sin una estructura temporal estable. En otras palabras, se busca tensionar el sistema con ruido fuerte pero esencialmente desordenado.

Resultados: respuesta del *autoencoder* (AE)

La Figura 15 muestra la evolución del error de reconstrucción del *autoencoder* junto con su umbral. Durante la ventana de inestabilidad aparecen picos aislados que superan el umbral en algunos instantes concretos. Esto es esperable: ciertas combinaciones instantáneas de métricas son raras respecto al régimen normal y el AE las reconstruye peor.

Sin embargo, lo relevante es que el error no se mantiene elevado de forma sostenida. El AE reacciona a instantes raros, pero no interpreta el episodio completo

como un cambio de régimen prolongado. Desde un punto de vista operativo, esto es una buena señal, porque evita convertir una tormenta transitoria en una alarma permanente.

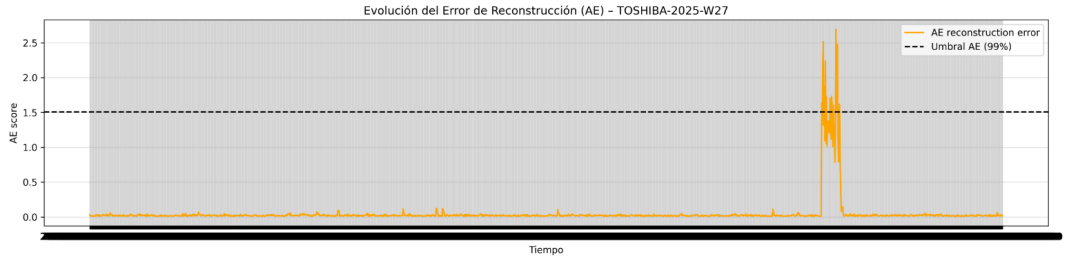


Figura 15: Evolución del error de reconstrucción del *autoencoder* (AE) para el Demostrador IV (inestabilidad estocástica). Se observan picos aislados que superan el umbral, sin elevación sostenida del error.

Resultados: análisis temporal multivariante y severidad

En la Figura 16 se representa la evolución conjunta de SKR y QBER, destacando las detecciones de los modelos y su fusión (IF_ONLY, AE_ONLY y BOTH). Se aprecia que las detecciones aparecen de forma dispersa dentro de la ventana, pero sin formar un bloque largo de anomalía confirmada.

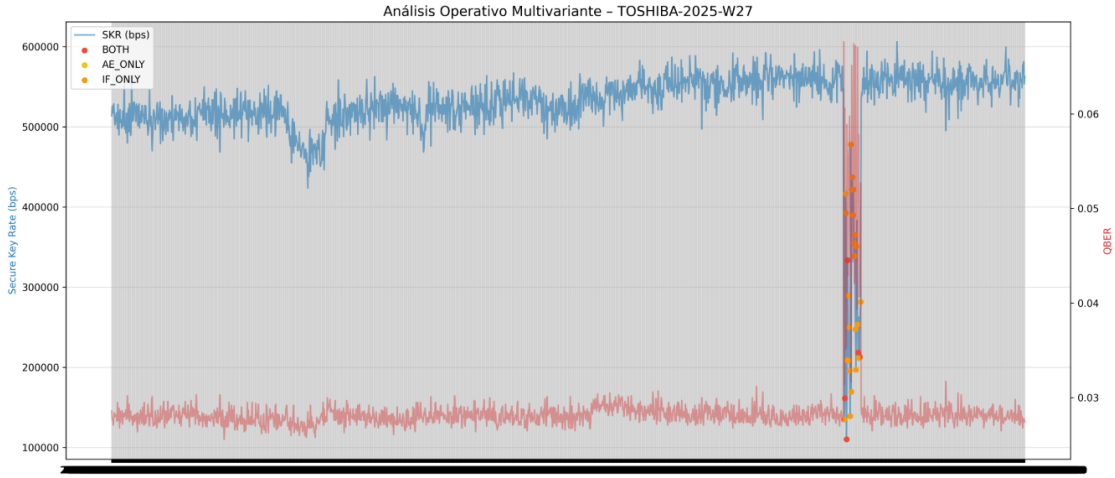


Figura 16: Análisis operativo multivariante (SKR y QBER) para el Demostrador IV.

Resultados: comportamiento en el espacio de fase QBER/SKR

La Figura 17 permite interpretar el episodio desde un punto de vista geométrico, observando dónde caen los puntos detectados en el plano QBER/SKR. Los puntos anómalos aparecen repartidos y, aunque algunos se alejan de la nube principal, no forman un clúster compacto ni una trayectoria suave (como ocurriría en una degradación progresiva del canal).

Esta dispersión es coherente con el tipo de perturbación aplicada: oscilaciones rápidas que rompen momentáneamente la coherencia entre métricas, pero sin establecer un nuevo régimen operativo estable.

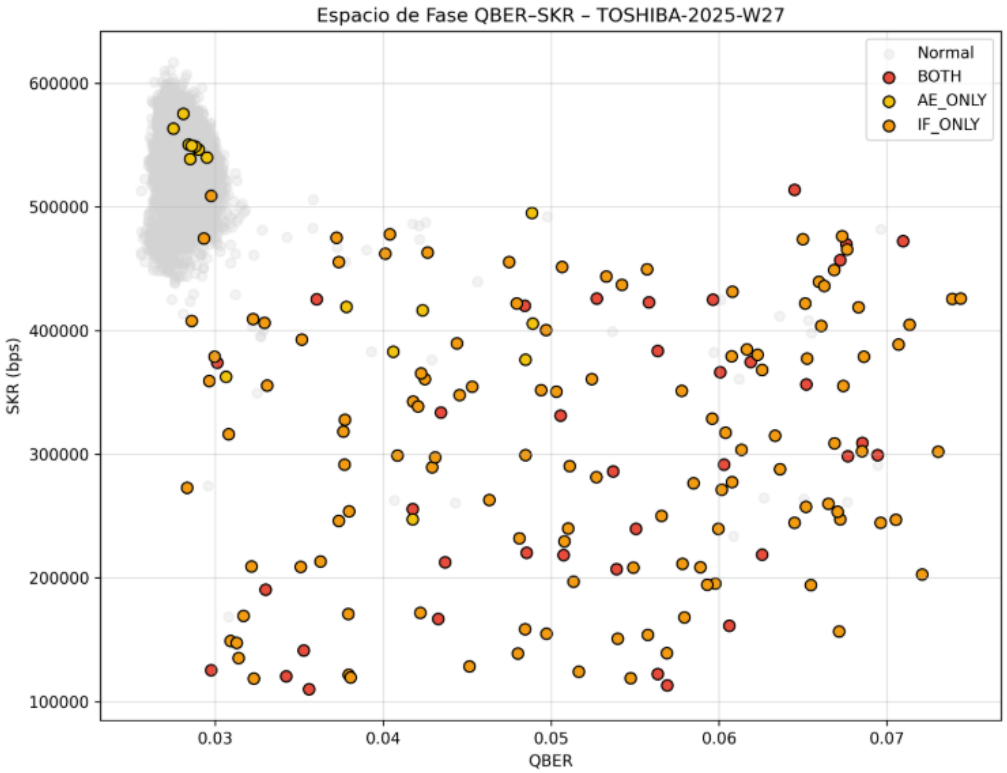


Figura 17: Espacio de fase QBER/SKR para el Demostrador IV. Los puntos detectados aparecen dispersos, reflejando una perturbación estocástica sin estructura estable ni cambio de régimen sostenido.

Discusión: falsos positivos y utilidad operativa

Este demostrador es especialmente útil para razonar sobre falsos positivos. En un entorno real, si el sistema generase alarmas sostenidas ante este tipo de ruido, se volvería poco operativo: el operador acabaría ignorando alertas y se perdería confianza en la herramienta.

El comportamiento observado indica que el *pipeline* es robusto frente a inestabilidad estocástica de alta frecuencia: puede marcar puntos puntuales como raros, pero evita declarar una anomalía persistente de alta severidad cuando no existe una estructura consistente en el tiempo. Los picos detectados pueden interpretarse como transitorios compatibles con ajustes internos, reintentos o inestabilidad técnica momentánea, sin que ello implique necesariamente un deterioro mantenido del enlace.

5.8. Conclusión global de la evaluación

En conjunto, los cuatro demostradores analizados muestran que el sistema propuesto distingue razonablemente entre: (i) eventos estructurales claros, (ii) degradaciones progresivas, y (iii) ruido operativo transitorio. El *pipeline* reacciona cuando existe una ruptura relevante de coherencia entre métricas, pero evita sobrerreacciones.

La ausencia de falsos positivos sostenidos en episodios de operación globalmente sana, junto con la detección consistente de patrones físicamente plausibles en los demostradores anteriores, refuerza la validez del enfoque y su aplicabilidad potencial como herramienta de apoyo para la operación de enlaces QKD reales.

6. Conclusiones y trabajo futuro

6.1. Conclusiones del proyecto y logro de objetivos

El objetivo principal de este Trabajo Fin de Grado ha sido el diseño, la implementación y la evaluación de un sistema de monitorización y detección de anomalías aplicable a infraestructuras reales de distribución cuántica de clave, partiendo de datos operativos reales y haciendo uso de herramientas genéricas de análisis de datos y de aprendizaje automático no supervisado. Desde su planteamiento inicial, el proyecto se ha abordado con un enfoque claramente ingenieril, alejándose de demostraciones puramente teóricas y centrado en dar respuesta a problemas prácticos asociados tanto al despliegue como a la operación continuada de enlaces QKD en entornos reales.

A lo largo del desarrollo del trabajo se ha puesto de manifiesto que es posible caracterizar de forma robusta el comportamiento normal de enlaces QKD heterogéneos sin recurrir a reglas *ad hoc*, a umbrales definidos manualmente ni a intervención humana directa durante la fase de monitorización. En particular, los resultados obtenidos muestran que las técnicas de aprendizaje no supervisado, cuando se combinan adecuadamente con conocimiento del dominio y con una fase cuidadosa de ingeniería de características, permiten capturar patrones operativos coherentes incluso en sistemas altamente estables y en escenarios en los que no existe ningún tipo de dato etiquetado que sirva como referencia.

El principal resultado del proyecto es la construcción de un *pipeline* completo, modular y reproducible, que cubre de manera integral todas las etapas necesarias para transformar datos operativos heterogéneos en información útil desde un punto de vista operativo. Este *pipeline* abarca desde la ingestión, limpieza y normalización de conjuntos de datos procedentes de distintos fabricantes, hasta la detección y el análisis de anomalías mediante modelos no supervisados. En su desarrollo se han integrado técnicas estadísticas clásicas, una fase extensa de ingeniería de características y modelos de aprendizaje automático como Isolation Forest y *autoencoders*, manteniendo en todo momento la coherencia física y funcional del sistema QKD analizado.

Este enfoque permite abordar de forma directa uno de los retos más relevantes a

los que se enfrentan las infraestructuras QKD actuales: la dificultad de convertir un volumen creciente de métricas cuánticas y clásicas en señales verdaderamente accionables para la operación diaria del sistema. El trabajo demuestra que la mera disponibilidad de datos no garantiza una monitorización eficaz si no se dispone de herramientas capaces de extraer estructura, identificar incoherencias relevantes y priorizar eventos desde una perspectiva operativa. En este sentido, la propuesta desarrollada constituye un paso sólido hacia sistemas de monitorización más robustos, interpretables y alineados con las necesidades reales de explotación de infraestructuras QKD.

Los objetivos planteados al inicio del proyecto pueden considerarse cumplidos de forma satisfactoria, tanto desde una perspectiva técnica como metodológica:

- Se ha realizado un análisis riguroso del contexto tecnológico y operativo de las infraestructuras QKD actuales, incluyendo métricas cuánticas clave, procesos de estandarización y arquitecturas de control.
- Se han procesado y analizado *datasets* operativos reales procedentes de sistemas QKD comerciales de distintos fabricantes, caracterizando sus diferencias de comportamiento y justificándolas desde un punto de vista físico y de diseño.
- Se ha diseñado un conjunto sólido de características temporales, multivariantes y derivadas, orientadas a capturar tanto el estado instantáneo como la dinámica temporal del sistema.
- Se han implementado y evaluado modelos no supervisados de detección de anomalías, mostrando su carácter complementario y sus distintas sensibilidades ante patrones anómalos de naturaleza diversa.
- Se ha validado el comportamiento del sistema mediante demostradores operativos controlados, reproduciendo escenarios físicamente plausibles y evaluando la coherencia de la respuesta del *pipeline* desde un punto de vista operativo.

Uno de los mensajes clave que se desprenden de este trabajo es que no existe un único detector capaz de cubrir de forma óptima todos los tipos de anomalías que pueden manifestarse en una infraestructura QKD real. En este sentido, Isolation Forest se muestra especialmente eficaz para la identificación de desviaciones estadísticas puntuales y de eventos extremos, mientras que los *autoencoders* aportan una capacidad complementaria para detectar rupturas estructurales y anomalías multivariantes de carácter global. La combinación de ambos enfoques permite construir un sistema de monitorización más robusto y equilibrado, reforzando la idea de que el aprendizaje automático debe entenderse como una herramienta de apoyo al análisis humano y no como un mecanismo de decisión completamente autónomo.

Por último, el trabajo pone de relieve la necesidad de superar enfoques simplistas basados en umbrales fijos o en análisis univariantes. En infraestructuras QKD reales, las relaciones entre métricas son complejas, no lineales y fuertemente dependientes del contexto operativo, lo que hace imprescindible el uso de

modelos capaces de capturar estructura y coherencia más allá del análisis de valores absolutos individuales.

6.2. Líneas de trabajo futuro

El sistema desarrollado constituye una base sólida para la monitorización operativa de infraestructuras QKD, pero no debe considerarse una solución cerrada. A partir de los resultados obtenidos, se identifican varias líneas claras de trabajo futuro que permitirían ampliar el enfoque presentado.

En primer lugar, resulta fundamental validar el sistema mediante experimentación física directa sobre infraestructura QKD real. La introducción controlada de perturbaciones como variaciones térmicas, desalineaciones ópticas, cambios en la atenuación del canal o congestión inducida del plano clásico permitiría contrastar los demostradores artificiales utilizados en este trabajo con escenarios reales y ajustar de forma más precisa la sensibilidad del sistema.

En segundo lugar, la detección de degradaciones progresivas o *drift* lento se ha identificado como un reto específico. Aunque el sistema propuesto evita sobrerreacciones ante variaciones suaves, futuras extensiones podrían incorporar técnicas específicas de detección de deriva a largo plazo, modelos temporales más avanzados o arquitecturas de tipo recurrente, orientadas a tareas de mantenimiento predictivo en enlaces QKD desplegados durante largos periodos.

Otra línea natural de evolución consiste en la integración del *pipeline* con sistemas de control y orquestación de red. En particular, la salida del sistema de monitorización podría alimentar el plano de control de una red cuántica basada en SDN (del inglés *software-defined networking*), permitiendo la reconfiguración dinámica del encaminamiento de claves, el ajuste automático de parámetros operativos o el re-enrutamiento del tráfico ante la detección de degradaciones en un enlace. Esta integración abriría la puerta a redes QKD auto-reparables (*self-healing networks*), capaces de reaccionar de forma autónoma ante fallos o degradaciones.

Finalmente, como continuación directa de este Trabajo Fin de Grado, el tutor del proyecto ha propuesto la ampliación de esta línea de investigación en el marco de un futuro Trabajo Fin de Máster (TFM). Esta continuidad permitiría profundizar en la validación experimental, el procesamiento en tiempo casi real y la integración del sistema en infraestructuras QKD operativas como MadQCI, consolidando el enfoque presentado como una herramienta práctica y aplicable para la gestión y la seguridad de redes cuánticas de próxima generación.

7. Análisis de impacto

7.1. Impacto tecnológico y en la seguridad nacional (EuroQCI)

Las infraestructuras de distribución cuántica de clave se consideran en la actualidad un componente estratégico dentro de las políticas europeas de ciberseguridad y soberanía tecnológica. Esta consideración no responde únicamente

a avances científicos recientes, sino a un contexto geopolítico y tecnológico en el que la protección de las comunicaciones críticas se ha convertido en un elemento central para la autonomía digital de los Estados y de las instituciones supranacionales.

Iniciativas como *EuroQCI* (European Quantum Communication Infrastructure) surgen como respuesta directa a este escenario, con el objetivo de desplegar una red cuántica paneuropea capaz de proteger comunicaciones sensibles frente a amenazas presentes y futuras, incluyendo adversarios con capacidades avanzadas de computación cuántica. A diferencia de proyectos puramente experimentales, EuroQCI se concibe desde su origen como una infraestructura operativa, destinada a integrarse en redes reales utilizadas por administraciones públicas, organismos gubernamentales, operadores de telecomunicaciones e infraestructuras críticas.

En este contexto, resulta especialmente relevante subrayar que la seguridad de una infraestructura QKD no depende únicamente de los fundamentos teóricos de la mecánica cuántica ni de la corrección formal de los protocolos criptográficos. Aunque la seguridad de la QKD se apoya en principios físicos bien establecidos, la seguridad global del sistema depende de forma crítica de su **operación continua, fiable y correctamente monitorizada**. Fallos técnicos, degradaciones progresivas del canal, problemas en el plano clásico o errores de configuración pueden comprometer la disponibilidad de clave segura o inducir estados operativos no deseables, incluso en ausencia de ataques explícitos.

El trabajo desarrollado en este proyecto aborda precisamente esta dimensión operativa de la seguridad QKD. El *pipeline* propuesto permite analizar métricas cuánticas y clásicas reales, identificar patrones anómalos y clasificar eventos en función de su severidad y persistencia, proporcionando una capa adicional de supervisión sobre el estado de salud del sistema. Este enfoque no pretende sustituir los mecanismos de seguridad intrínsecos de la QKD, sino complementarlos mediante herramientas de monitorización avanzadas que faciliten una explotación segura, sostenible y auditable de la infraestructura.

Desde un punto de vista tecnológico, uno de los impactos más relevantes del sistema propuesto es su contribución a la **independencia respecto al fabricante**. En los despliegues actuales, los operadores de redes QKD suelen depender de paneles de monitorización propietarios proporcionados por los fabricantes de los equipos (por ejemplo, Toshiba o QTI). Estos sistemas suelen funcionar como "cajas negras", con lógica interna no documentada y métricas agregadas cuyo significado operativo no siempre es transparente para el operador final.

El enfoque presentado en este trabajo permite al operador de la red (por ejemplo, un organismo público, un operador nacional de telecomunicaciones o una entidad gubernamental) disponer de su **propia métrica de salud del enlace**, construida a partir de datos brutos y procesada mediante herramientas abiertas y reproducibles. Esta capacidad reduce la dependencia de soluciones propietarias, mejora la transparencia del sistema y refuerza la soberanía tecnológica, al permitir que la evaluación del estado de la red no quede supeditada exclusivamente a la interpretación del fabricante.

Este aspecto resulta especialmente crítico en infraestructuras como EuroQCI, caracterizadas por una fuerte heterogeneidad. La coexistencia de múltiples fabricantes, tecnologías QKD distintas, topologías de red variadas y escenarios de despliegue heterogéneos hace inviable una monitorización manual o basada exclusivamente en umbrales estáticos definidos *a priori*. En este sentido, el uso de técnicas de aprendizaje automático no supervisado permite adaptar el sistema de monitorización al comportamiento real de cada enlace, reduciendo la dependencia de configuraciones rígidas y aumentando la resiliencia global de la red.

Desde la perspectiva de la seguridad nacional, disponer de mecanismos automáticos capaces de detectar degradaciones anómalas tanto del canal cuántico como del plano clásico tiene un impacto directo en la protección de comunicaciones gubernamentales, militares y de infraestructuras críticas. La detección temprana de comportamientos anómalos permite actuar antes de que se produzcan fallos graves, minimizando riesgos operativos y contribuyendo a la continuidad del servicio. En este sentido, el enfoque presentado se alinea con los objetivos estratégicos de la Unión Europea en materia de ciberseguridad post-cuántica, autonomía digital y soberanía tecnológica.

7.2. Contribución a los Objetivos de Desarrollo Sostenible (ODS)

Aunque este trabajo se sitúa en un ámbito eminentemente técnico y especializado, su impacto puede relacionarse de forma indirecta, pero significativa, con varios de los Objetivos de Desarrollo Sostenible (ODS) definidos por Naciones Unidas.

En particular, el proyecto contribuye al **ODS 9: Industria, innovación e infraestructura**, al desarrollar y evaluar herramientas orientadas a la monitorización de infraestructuras de comunicación de nueva generación. La mejora de la fiabilidad, la seguridad y la eficiencia operativa de redes QKD favorece el desarrollo de infraestructuras digitales resilientes, capaces de soportar servicios críticos y de evolucionar de forma sostenible a largo plazo.

Existe también una contribución al **ODS 16: Paz, justicia e instituciones sólidas**, en la medida en que la protección de las comunicaciones críticas constituye un elemento esencial para el funcionamiento seguro de las instituciones públicas. Garantizar la confidencialidad y la integridad de las comunicaciones refuerza la confianza en los sistemas de información y contribuye a la estabilidad de las estructuras institucionales.

Finalmente, el uso de técnicas de análisis de datos y aprendizaje automático para optimizar la operación de infraestructuras complejas puede relacionarse con el **ODS 12: Producción y consumo responsables**. La capacidad de detectar degradaciones progresivas y planificar intervenciones de mantenimiento de forma más informada permite reducir actuaciones innecesarias, evitar fallos catastróficos y minimizar el consumo de recursos asociado a paradas no planificadas o a la sustitución prematura de componentes.

En conjunto, aunque el impacto del trabajo es fundamentalmente tecnológico,

sus resultados se alinean con objetivos más amplios relacionados con la innovación responsable, la seguridad de las comunicaciones y el desarrollo de infraestructuras digitales sostenibles, mostrando cómo un trabajo técnico especializado puede contribuir de forma tangible a retos sociales y estratégicos de mayor alcance.

Bibliografía

[1] N. Gisin, G. Ribordy, W. Tittel y H. Zbinden, «Quantum cryptography,» *Rev. Mod. Phys.*, vol. 74, págs. 145-195, 1 mar. de 2002. DOI: 10.1103/RevModPhys.74.145. dirección: https://link.aps.org/doi/10.1103/RevModPhys.74.145.

[2] *The European Quantum Communication Infrastructure (EuroQCI) Initiative*, 2024. dirección: https://digital-strategy.ec.europa.eu/en/policies/european-quantum-communication-infrastructure-euroqci.

[3] V. Martin et al., «MadQCI: a heterogeneous and scalable SDN-QKD network deployed in production facilities,» *npj Quantum Information*, vol. 10, n.º 1, pág. 80, sep. de 2024, ISSN: 2056-6387. DOI: 10.1038/s41534-024-00873-2. dirección: https://doi.org/10.1038/s41534-024-00873-2.

[4] A. Sebastián-Lombraña, H. H. Brunner, R. B. Méndez, C.-H. F. Fung, J. P. Brito y M. Peev, «Towards holistic quantum communications infrastructures,» en *Proceedings of the 2025 International Conference on Quantum Communications, Networking, and Computing (QCNC)*, Nara, Japan: IEEE, 2025, págs. 494-501. dirección: https://qsnp.eu/publications/towards-holistic-quantum-communications-infrastructures/.

[5] ITU-T, *Quantum key distribution networks: Technical framework*, 2020.

[6] ETSI ISG QKD, *Quantum Key Distribution (QKD); General requirements for the security of QKD systems*, 2022.

[7] European Commission, *European Quantum Communication Infrastructure (EuroQCI)*, 2023.

[8] Ministerio de Asuntos Económicos y Transformación Digital, *Plan Complementario de Comunicaciones Cuánticas*, Government report, 2022.

[9] Universidad Politécnica de Madrid – CeDInt, *Infraestructura cuántica y comunicaciones seguras*, Institutional report, 2022.

[10] V. Scarani, H. Bechmann-Pasquinucci, N. Cerf, M. Dusek, N. Lutkenhaus y M. Peev, «The security of practical quantum key distribution,» *Reviews of Modern Physics*, 2009.

[11] C.-H. F. Fung, L. Lydersen y V. Makarov, «Quantum hacking and security of practical quantum key distribution,» *IEEE Journal of Selected Topics in Quantum Electronics*, 2015.

[12] J. F. Dynes, S.-H. Tam, A. Plews, Z. Yuan, A. Sharpe y A. J. Shields, «Real-time diagnostics on a quantum key distribution system using time-series analysis,» *Optics Express*, 2015.

[13] F. T. Liu, K. M. Ting y Z.-H. Zhou, «Isolation Forest,» 2008.

[14] F. T. Liu, K. M. Ting y Z.-H. Zhou, «Isolation-Based Anomaly Detection,» *ACM Transactions on Knowledge Discovery from Data*, 2012.

[15] G. E. Hinton y R. R. Salakhutdinov, «Reducing the Dimensionality of Data with Neural Networks,» *Science*, 2006.

[16] M. Sakurada y T. Yairi, «Anomaly Detection Using Autoencoders with Non-linear Dimensionality Reduction,» 2014.

[17] V. Chandola, A. Banerjee y V. Kumar, «Anomaly Detection: A Survey,» *ACM Computing Surveys*, vol. 41, n.º 3, 2009.

[18] C. C. Aggarwal, *Outlier Analysis*. Springer, 2017.

[19] R. J. Hyndman y G. Athanasopoulos, *Forecasting: Principles and Practice*, 2.ª ed. OTexts, 2018.

[20] T. Smolen., «Comparing Autoencoder and Isolation Forest in Network Anomaly Detection,» *IEEE / MDPI / ResearchGate*, 2020.

Anexos

1. Descripción detallada de los datasets utilizados

Esta primera parte del anexo describe de forma sistemática los conjuntos de datos empleados a lo largo del trabajo. Dado que el objetivo principal del TFG es el análisis y monitorización de datos operativos en infraestructuras de red cuántica, resulta fundamental detallar el origen, la estructura y las limitaciones de los datos utilizados, con el fin de garantizar la reproducibilidad del análisis y facilitar su evaluación técnica.

Los datasets considerados corresponden a dos fuentes diferenciadas, representativas de distintos niveles de granularidad y disponibilidad de métricas en sistemas QKD reales o verosímiles.

1.1. Resumen global de los datasets

La Tabla 2 resume las principales características de los conjuntos de datos empleados.

Dataset	Fabricante / Origen	Frecuencia	Variables principales
QTI	QTI (entorno experimental)	$\sim$10–15 min	QBER, SKR, pérdidas
Toshiba-2024-W25	Toshiba QKD	$\sim$1–2 min	QBER, SKR
Toshiba-2025-W27	Toshiba QKD	$\sim$1 min	QBER, SKR

Cuadro 2: Resumen de los datasets utilizados en el trabajo.

1.2. Dataset QTI

El dataset `QTI` corresponde a un entorno experimental de QKD con un mayor nivel de instrumentación interna, lo que permite disponer de métricas tanto del plano cuántico como de parámetros físicos del canal.

1.3. Variables disponibles

Las principales variables incluidas en este conjunto de datos son:

- `datetime`: instante temporal de la medición.
- `channel_loss`: pérdidas del canal cuántico, expresadas en decibelios (dB).
- `secure_key_length`: longitud total de clave segura generada.
- `secure_key_rate`: tasa de generación de clave segura (bps).
- `qber`: tasa de error cuántico (Quantum Bit Error Rate).
- `phi`: parámetro interno del protocolo relacionado con la estimación de fase.
- `mu_1`, `mu_2`: intensidades medias de los pulsos ópticos.

Características y limitaciones

Este dataset contiene un total de 87 muestras, con una resolución temporal relativamente baja (del orden de varios minutos entre muestras). A cambio, proporciona una visión rica del estado físico del sistema, lo que resulta especialmente útil para el análisis de coherencia entre métricas cuánticas y parámetros del canal.

No obstante, su reducido tamaño limita la aplicación de modelos de aprendizaje automático complejos y lo convierte principalmente en un conjunto de validación cualitativa y exploratoria.

1.4. Dataset Toshiba-2024-W25

El dataset `Toshiba-2024-W25` corresponde a un sistema QKD comercial de Toshiba, con un conjunto de métricas más reducido y orientado a la operación.

Variables disponibles

Las variables incluidas son:

- `Time`: lo que nos permite saber el momento exacto de la medición.
- `QBER`: tasa de error cuántico.
- `SecureKeyRate(bps)`: tasa de generación de clave segura en bits por segundo.

Características y limitaciones

El dataset consta de aproximadamente 90 muestras, con una frecuencia de muestreo cercana al minuto. La ausencia de métricas físicas del canal (como pérdidas ópticas o parámetros internos del protocolo) limita el análisis a relaciones directas entre QBER y SKR.

Este conjunto de datos resulta representativo de escenarios operativos reales, donde únicamente se dispone de métricas de alto nivel expuestas por el sistema de gestión del enlace QKD.

1.5. Dataset Toshiba-2025-W27

El dataset `Toshiba-2025-W27` constituye el conjunto de datos principal del trabajo y sobre el que se realizan la mayoría de los experimentos y demostradores.

1.6. Variables disponibles

Las variables disponibles son idénticas a las del dataset Toshiba-2024-W25:

- `Time`: momento de la medición.
- `QBER`: tasa de error cuántico.

- `SecureKeyRate(bps)`: tasa de generación de clave segura.

Características y limitaciones

Este dataset contiene 10472 muestras, con una frecuencia de muestreo cercana a un minuto, lo que permite capturar tanto la variabilidad normal del sistema como eventos transitorios y degradaciones progresivas.

Su principal limitación es la falta de métricas físicas directas del canal cuántico, como pérdidas ópticas o parámetros internos del protocolo. En consecuencia, el análisis se basa en la coherencia estadística y temporal entre QBER y SKR, en lugar de en un modelo físico completo del enlace.

A pesar de esta limitación, el tamaño y la resolución temporal del dataset lo convierten en un candidato adecuado para la aplicación de técnicas de detección de anomalías y análisis multivariante en un contexto operativo realista.

1.7. Consideraciones finales

La combinación de datasets con distintos niveles de detalle permite evaluar el comportamiento del pipeline propuesto en escenarios heterogéneos, desde entornos altamente instrumentados hasta sistemas operativos reales con métricas limitadas. Esta diversidad refuerza la validez del enfoque adoptado y su aplicabilidad práctica en infraestructuras de red cuántica.

2. Descripción del pipeline de análisis y monitorización

Esta segunda parte describe el flujo completo del pipeline desarrollado en el proyecto, desde la ingesta de datos operativos hasta la generación de resultados y visualizaciones finales. El objetivo de este apartado es proporcionar una visión clara y reproducible del proceso seguido, separando explícitamente el diseño metodológico de los resultados presentados en el cuerpo principal del trabajo.

El pipeline ha sido diseñado con un enfoque modular, reproducible y orientado a la operación, permitiendo analizar conjuntos de datos heterogéneos de infraestructuras QKD de forma consistente. Pero con especial énfasis en los datos generados por máquinas QTI y Toshiba.

2.1. Visión general del flujo de trabajo

El pipeline completo sigue una secuencia bien definida de etapas, que pueden agruparse en los siguientes bloques funcionales:

1. Ingesta y normalización de datos operativos.
2. Preprocesado y unificación de formatos.
3. Ingeniería de características multivariantes y temporales.
4. Detección de anomalías mediante modelos no supervisados.

5. Fusión y análisis comparativo de resultados.
6. Generación de métricas y visualizaciones operativas.

Cada una de estas etapas se implementa como un módulo independiente, ejecutado de forma secuencial por un proceso principal que orquesta todo el flujo de análisis.

2.2. Ingesta y normalización de datos

El pipeline parte de conjuntos de datos operativos almacenados en formato CSV, procedentes de diferentes fabricantes o entornos experimentales. Dado que estos datasets presentan estructuras heterogéneas, la primera etapa consiste en detectar automáticamente el tipo de dataset y normalizar sus campos principales a un esquema común.

En esta fase se garantiza que todas las fuentes queden representadas mediante un conjunto mínimo de variables operativas homogéneas, incluyendo marcas temporales, métricas cuánticas principales (como QBER y SKR) y un identificador de origen (`source`). Esta normalización permite tratar posteriormente todos los datos de forma conjunta sin perder la trazabilidad de su procedencia.

2.3. Preprocesado y alineación temporal

Una vez normalizados, los datasets se combinan en una única estructura temporal ordenada. Durante esta etapa se realizan operaciones estándar de preprocesado orientadas a facilitar el análisis posterior, tales como:

- Conversión y validación de marcas temporales.
- Ordenación cronológica por fuente.
- Alineación temporal mediante remuestreo cuando es necesario.

El resultado de esta fase es un dataset operativo unificado, listo para ser explotado por los módulos de análisis y modelado, manteniendo la coherencia temporal dentro de cada fuente.

2.4. Ingeniería de características

Sobre el dataset preprocesado se construye un conjunto enriquecido de características destinadas a capturar tanto el estado instantáneo del sistema como su evolución temporal. Esta etapa constituye una de las piezas centrales del pipeline.

Las características generadas pueden agruparse en varias categorías:

- **Características temporales**: hora del día, día de la semana, segmentación temporal y variables indicadoras de periodos nocturnos.
- **Características retardadas**: valores pasados de métricas clave (lags) para capturar dependencias temporales.

- **Características de variación**: incrementos y cambios relativos entre muestras consecutivas.
- **Estadísticos locales**: medias, desviaciones típicas y extremos calculados sobre ventanas móviles de distinto tamaño.
- **Interacciones físicas**: combinaciones entre métricas (por ejemplo, relaciones QBER–SKR) coherentes con el comportamiento esperado de sistemas QKD.

Este enfoque multiescala permite representar tanto dinámicas rápidas como tendencias más lentas, proporcionando una base rica para la detección de comportamientos anómalos.

2.5. Detección de anomalías

El pipeline incorpora dos enfoques complementarios de detección de anomalías, ambos basados en aprendizaje no supervisado y entrenados exclusivamente sobre datos operativos:

- **Isolation Forest**, orientado a la detección de observaciones estadísticamente aisladas en el espacio de características.
- **Autoencoder**, orientado a la detección de rupturas estructurales mediante el análisis del error de reconstrucción.

Cada modelo se entrena de forma independiente por fuente, con el objetivo de capturar el comportamiento normal específico de cada enlace o sistema. Esta separación evita introducir sesgos derivados de mezclar dinámicas operativas heterogéneas.

2.6. Fusión y clasificación de eventos

Los resultados generados por ambos modelos se combinan en una etapa de fusión que permite comparar su consistencia y clasificar los eventos detectados según su severidad relativa. En particular, se distinguen:

- Eventos no anómalos.
- Anomalías detectadas por un único modelo.
- Eventos detectados simultáneamente por ambos modelos.

Esta clasificación proporciona una visión más robusta del estado operativo del sistema, reduciendo la dependencia de un único criterio de detección y facilitando la interpretación posterior.

2.7. Análisis final y visualización

Como última etapa, el pipeline genera métricas resumen y visualizaciones orientadas a la interpretación operativa de los resultados. Entre ellas se incluyen:

- Evolución temporal conjunta de métricas cuánticas y etiquetas de anomalía.
- Representaciones en el espacio de fase QBER–SKR.
- Evolución de los scores de anomalía de los modelos.

Estas salidas constituyen la base para los demostradores presentados en el cuerpo principal del trabajo y permiten evaluar de forma cualitativa y cuantitativa el comportamiento del sistema ante distintos escenarios operativos.

2.8. Resumen

Este anexo documenta el camino completo seguido por los datos dentro del pipeline, desde su ingesta hasta la generación de resultados finales. La separación en módulos independientes y la claridad del flujo permiten que el sistema sea reproducible, extensible y aplicable a otros entornos de monitorización de infraestructuras QKD, más allá de los casos analizados en este trabajo.

3. Detalles de implementación y parámetros de los modelos

Este anexo describe las decisiones de implementación adoptadas en los modelos de detección de anomalías utilizados en el proyecto, así como los hiperparámetros seleccionados y sus implicaciones computacionales. El objetivo de este apartado es justificar de forma transparente las elecciones realizadas, situándolas dentro del alcance razonable de un Trabajo Fin de Grado y de un contexto operativo realista.

3.1. Isolation Forest

Isolation Forest se ha utilizado como detector principal de anomalías estadísticas puntuales. Se trata de un algoritmo no supervisado ampliamente utilizado en contextos de monitorización, debido a su buen compromiso entre capacidad de detección, interpretabilidad y coste computacional.

Hiperparámetros seleccionados

Para cada fuente de datos, el modelo se ha entrenado de forma independiente utilizando los siguientes parámetros principales:

- **Número de árboles (`n_estimators = 600`)**: se seleccionó un valor elevado para reducir la varianza del modelo y estabilizar las puntuaciones de anomalía en datasets con alta dimensionalidad tras la ingeniería de características.
- **Contaminación (`contamination = 0.003`)**: se fijó un valor conservador, coherente con un entorno operativo sano, donde se espera que los eventos verdaderamente anómalos sean escasos.

- **Semilla aleatoria fija**: utilizada para garantizar reproducibilidad en los resultados.

La elección de estos valores prioriza la estabilidad del modelo frente a la sensibilidad extrema, evitando una sobre-detección de anomalías ante variaciones normales del sistema.

Justificación metodológica

Isolation Forest está diseñado para aislar observaciones que requieren menos particiones aleatorias para ser separadas del conjunto principal. En este trabajo se utiliza como detector de desviaciones estadísticas instantáneas, especialmente eficaz ante cambios abruptos o eventos extremos en el espacio de características.

El entrenamiento se realiza mediante aprendizaje *one-class*, asumiendo que la mayor parte del dataset representa comportamiento normal, una hipótesis razonable en el contexto de infraestructuras QKD operativas.

3.2. Autoencoder

El Autoencoder se ha empleado como modelo complementario, orientado a detectar rupturas estructurales en la relación multivariante entre métricas, más allá de outliers individuales.

Arquitectura y parámetros

El modelo implementado corresponde a un Autoencoder denso y simétrico, con las siguientes características:

- **Dimensión latente**: 16.
- **Capas**: dos capas densas en el codificador y dos en el decodificador.
- **Función de activación**: ReLU.
- **Función de pérdida**: error cuadrático medio (MSE).
- **Número de épocas**: 25.
- **Tamaño de batch**: 128.
- **Optimizador**: Adam con tasa de aprendizaje 10^{-3}.

Esta arquitectura busca un equilibrio entre capacidad de representación y simplicidad, evitando modelos excesivamente profundos que resultarían difíciles de justificar y entrenar con datasets de tamaño limitado.

Umbral de detección

La detección de anomalías se basa en el error de reconstrucción del Autoencoder. Se ha definido un umbral estático correspondiente al percentil 99.5 del error de reconstrucción observado durante el entrenamiento.

Este criterio permite identificar observaciones que se desvían significativamente de la estructura latente aprendida, sin introducir etiquetas externas ni supuestos adicionales.

Limitaciones del enfoque

El uso de un umbral fijo implica que el modelo no está diseñado para adaptarse automáticamente a cambios muy lentos del comportamiento normal del sistema a lo largo de largos periodos de tiempo. Esta decisión es coherente con el objetivo principal del sistema, centrado en la detección de eventos que requieren atención operativa inmediata, y no en la gestión de deriva a largo plazo, que correspondería a mecanismos de mantenimiento planificado.

3.3. Fusión de modelos

Los resultados de Isolation Forest y Autoencoder se combinan mediante una clasificación lógica que distingue entre detecciones individuales y eventos detectados simultáneamente por ambos modelos. Esta fusión permite aumentar la robustez del sistema, reduciendo la dependencia de un único criterio de anomalía y facilitando la interpretación operativa de los resultados.

3.4. Consideraciones computacionales

Desde el punto de vista computacional, el pipeline completo se ejecuta en modo batch y presenta un tiempo de ejecución total del orden de aproximadamente dos minuto para el dataset más grande analizado.

Este tiempo incluye preprocesado, ingeniería de características, entrenamiento de modelos y generación de visualizaciones. En particular:

- El entrenamiento de Isolation Forest presenta un coste moderado, escalando linealmente con el número de muestras.
- El entrenamiento del Autoencoder constituye la parte más costosa, aunque se mantiene dentro de límites razonables para un entorno de análisis offline.

Estos tiempos son compatibles con un escenario de monitorización diferida, donde el análisis se realiza de forma periódica (por ejemplo, cada pocos minutos), y no requieren capacidades de computación en tiempo real estricto.

3.5. Resumen

Las decisiones de implementación y parametrización adoptadas reflejan un compromiso consciente entre complejidad, interpretabilidad y viabilidad operativa. Los modelos utilizados no pretenden maximizar métricas abstractas de detección, sino proporcionar señales útiles, estables y defendibles para el análisis operativo de infraestructuras QKD reales.

4. Reproducibilidad y estructura del proyecto

Este anexo describe la organización del proyecto, la estructura de carpetas utilizada y el procedimiento necesario para reproducir los experimentos presentados en el trabajo. El objetivo es facilitar la comprensión del flujo de ejecución y permitir la replicación completa del análisis a partir del código y los datos proporcionados.

4.1. Estructura general del proyecto

El proyecto se organiza siguiendo una estructura modular que separa claramente los datos, el código fuente y los resultados generados. La Figura 4.1 muestra la jerarquía principal de directorios utilizada.

- **data/**: contiene todos los datasets y resultados intermedios.
  - `raw/`: datos originales sin procesar.
    - `qkd/`: datasets QKD utilizados en el trabajo.
    - `original/`: copias de respaldo de los datasets originales.
    - `network/`: ficheros auxiliares de simulación que fueron entregados por el laboratorio pero finalmente no fueron utilizados (al no tener métricas cuánticas puras como QBER o SKR).
  - `processed/`: datos ya preprocesados y unificados.
  - `feature_engineered/`: datasets enriquecidos con variables temporales y estadísticas.
  - `models/`: resultados de los modelos (Isolation Forest, Autoencoder y fusiones).
  - `plots/`: visualizaciones finales generadas por el pipeline.
- **src/**: código fuente del proyecto.
  - `constants/`: definición centralizada de rutas y parámetros globales. Esto sigue las buenas prácticas en un entorno de análisis de datos.
  - `python/`: módulos principales del pipeline.
    - `demostradores/`: generación de escenarios sintéticos controlados.
    - `preprocessing.py`: limpieza y unificación de datasets.
    - `feature_engineering.py`: ingeniería avanzada de características.
    - `train_iforest.py`: entrenamiento de Isolation Forest.
    - `train_autoencoder.py`: entrenamiento del Autoencoder.
    - `merging_datasets.py`: fusión de resultados de modelos.
    - `analyze_anomalies.py`: análisis operativo final.

    - `plots.py`: generación de visualizaciones.

- **main.py**: desde este archivo se ejecuta el pipeline completo.

4.2. Ejecución del pipeline

La ejecución completa del sistema se realiza a través del script principal `main.py`. Este fichero actúa como orquestador del pipeline y coordina de forma secuencial todas las etapas del análisis. El propio código pregunta si le quieres injectar una anomalía al dataset, si no estas probando el pipeline y realmente quieres hacer un monitoreo de los datos en producción puedes saltarte ese paso.

Para reproducir el análisis completo, basta con ejecutar: python main.py (siempre que estés dentro de la carpeta src/).

Independientemente del escenario elegido, el pipeline ejecuta de forma ordenada las siguientes fases:

1. Restauración del estado inicial de los datos.
2. Preprocesado y unificación de los datasets.
3. Ingeniería avanzada de características.
4. Entrenamiento de Isolation Forest.
5. Entrenamiento del Autoencoder.
6. Fusión de resultados y análisis de anomalías.
7. Generación de visualizaciones finales.

Si en algún momento la persona que ejecuta quiere saltarse alguna de las fases o probar uno solo de los modelos podría comentar las funciones de run que se encuentran dentro de la funcion main() de main.py.

4.3. Reproducibilidad de los experimentos

El proyecto ha sido diseñado para garantizar la reproducibilidad de los resultados bajo las mismas condiciones de ejecución. Para ello:

- Se utilizan semillas aleatorias fijas en los modelos de aprendizaje automático.
- Los datasets originales se conservan sin modificaciones y se restauran automáticamente antes de cada ejecución.
- Las rutas y parámetros globales se centralizan en un único módulo de constantes.

Estas medidas permiten repetir los experimentos y obtener resultados coherentes, facilitando tanto la validación académica como futuras extensiones del trabajo.

4.4. Dependencias principales

El pipeline se apoya en herramientas y librerías estándar ampliamente utilizadas en análisis de datos y aprendizaje automático. Las principales las tenemos bien descritas en el archivo enviroment.yml:

- `name: qkd-ml dependencies: - python=3.10 - numpy - pandas - scikit-learn - matplotlib - seaborn - tensorflow - pip - pytorch`

Asimismo, se ha incluido un script en R para análisis exploratorio inicial, aunque su uso es opcional y no afecta a la ejecución principal del sistema.

4.5. Resumen

La estructura del proyecto y el diseño del pipeline permiten una ejecución reproducible, modular y controlada del análisis propuesto. Este enfoque refuerza la validez técnica del trabajo y facilita su comprensión, evaluación y posible reutilización en contextos académicos u operativos similares.

5. Limitaciones técnicas y líneas de trabajo futuro

Este anexo recoge de forma explícita las principales limitaciones técnicas del trabajo realizado, así como las líneas de trabajo futuro. El objetivo no es enumerar carencias de forma genérica, sino delimitar con precisión qué aspectos quedan fuera del alcance de este Trabajo Fin de Grado y qué extensiones serían necesarias para avanzar hacia un sistema de monitorización plenamente operativo en infraestructuras QKD reales.

5.1. Limitaciones inherentes a los datos operativos disponibles

Una primera limitación fundamental del trabajo está relacionada con la naturaleza de los datos operativos utilizados. Aunque los datasets analizados proceden de sistemas QKD reales y comerciales, se trata de volcados históricos de métricas ya agregadas, exportadas en formato CSV y sin acceso al flujo completo de telemetría en tiempo real. Así como la limitación que hemos tenido de material (solo dos datasets de Toshiba y uno de QTI).

En particular, no se dispone de:

- métricas de bajo nivel del hardware cuántico (por ejemplo, contadores de detección individuales o información detallada de sincronización).
- variables internas del plano clásico, como latencias de procesado, colas internas del KMS o estados detallados de los módulos de post-procesado.
- etiquetas fiables que asocien intervalos temporales concretos con fallos, ataques reales o eventos operativos documentados.

Estas limitaciones condicionan el tipo de análisis que puede realizarse y refuerzan la necesidad de adoptar un enfoque no supervisado y orientado a patrones, en lugar de modelos predictivos clásicos entrenados con ejemplos etiquetados.

5.2. Limitaciones del enfoque de detección no supervisada

El sistema propuesto se basa en modelos de aprendizaje automático no supervisado (Isolation Forest y Autoencoder), entrenados directamente sobre datos históricos considerados representativos del comportamiento normal del sistema. Este enfoque presenta ventajas claras en ausencia de etiquetas, pero introduce también limitaciones conocidas.

En particular, los modelos:

- no distinguen de forma explícita entre tipos semánticos de anomalías (por ejemplo, fallo físico frente a problema de configuración);
- pueden absorber degradaciones lentas y persistentes como comportamiento normal si estas dominan el conjunto de entrenamiento;
- dependen de forma crítica de la fase de ingeniería de características, que introduce supuestos implícitos sobre qué información temporal y física es relevante.

Por diseño, el sistema prioriza la detección de rupturas estructurales y anomalías abruptas que requieren atención operativa inmediata, y no pretende resolver el problema más amplio de diagnóstico causal ni de clasificación exhaustiva de eventos.

5.3. Limitaciones computacionales y de modo de operación

El pipeline desarrollado sigue un enfoque claramente orientado a procesamiento por lotes (*batch processing*). En el entorno de pruebas utilizado, el tiempo de ejecución completo del pipeline incluyendo carga de datos, preprocesado, ingeniería de características e inferencia de los modelos es del orden de un minuto por ventana de análisis.

Este tiempo de procesamiento descarta su uso en escenarios de tiempo real estricto o de control en lazo cerrado. No obstante, esta limitación es coherente con el contexto operativo actual de las infraestructuras QKD, donde las métricas se evalúan típicamente sobre ventanas temporales agregadas y las decisiones operativas no requieren latencias del orden de milisegundos.

El sistema debe entenderse, por tanto, como una herramienta de monitorización *near real-time* orientada a la supervisión continua, al análisis post-hoc y al soporte a operadores humanos, y no como un mecanismo de reacción inmediata integrado directamente en el control del enlace.

5.4. Limitaciones de validación experimental

La validación del sistema se ha realizado mediante demostradores operativos artificiales, basados en la inyección controlada de perturbaciones físicamente plausibles sobre datos reales. Aunque este enfoque permite evaluar la coherencia del sistema en condiciones realistas, no sustituye a la validación experimental directa sobre infraestructura QKD física.

No fue posible, en el marco temporal de este Trabajo Fin de Grado, introducir perturbaciones controladas directamente sobre los sistemas QKD del laboratorio ni disponer de campañas experimentales diseñadas específicamente para la validación del pipeline. Como consecuencia, los resultados deben interpretarse como una validación funcional y metodológica, pero no como una certificación completa del sistema para su despliegue en producción.

5.5. Líneas de trabajo futuro

A partir de las limitaciones identificadas, se abren varias líneas claras de trabajo futuro. En primer lugar, resulta prioritario validar el sistema mediante experimentación física directa, introduciendo perturbaciones controladas en enlaces QKD reales y comparando las detecciones obtenidas con observaciones operativas documentadas.

En segundo lugar, la incorporación de mecanismos específicos de detección de deriva a largo plazo permitiría complementar la detección de anomalías abruptas con herramientas orientadas al mantenimiento predictivo. Esto podría abordarse mediante análisis de tendencias, detectores de cambio de régimen o modelos temporales explícitos.

Otra extensión natural consiste en la adaptación del pipeline a un modo de operación en streaming, reduciendo la latencia de procesamiento y permitiendo su integración progresiva con sistemas de control y orquestación de red. En este contexto, la salida del sistema podría alimentar decisiones automáticas o semiautomáticas, como la reconfiguración del enlace o la priorización de tareas de mantenimiento.

Finalmente, una línea de trabajo especialmente relevante sería la integración del sistema de monitorización con arquitecturas de red cuántica definidas por software (SDN), permitiendo explorar conceptos como redes QKD auto-reparables y políticas de gestión dinámica de claves basadas en el estado operativo real de los enlaces.

Este documento esta firmado por

	**Firmante**	CN=tfgm.fi.upm.es, OU=CCFI, O=ETS Ingenieros Informaticos - UPM, C=ES
	Fecha/Hora	Thu Jan 15 16:30:53 CET 2026
	Emisor del Certificado	EMAILADDRESS=camanager@etsiinf.upm.es, CN=CA ETS Ingenieros Informaticos, O=ETS Ingenieros Informaticos - UPM, C=ES
	Numero de Serie	561
	Metodo	urn:adobe.com:Adobe.PPKLite:adbe.pkcs7.sha1 (Adobe Signature)