



**UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA CIVIL**

**DISEÑO EXPERIMENTAL PARA LA MODELACIÓN DEL DETERIORO DE PUENTES
CONVENCIONALES TIPO VIGA, USANDO HERRAMIENTAS DE IA**

POR

Cristobal Cisternas Casas

Memoria de Título presentada a la Facultad de Ingeniería de la Universidad de Concepción para
optar al título profesional de Ingeniero Civil

Profesor Guía
Tomás Echaveguren Navarro

Profesores Comisión
Víctor Aguilar Vidal
Juan Carrasco Montagna

28 de Enero del 2026
Concepción (Chile)

© 2026 Cristobal Cisternas Casas

© 2026 Cristobal Cisternas Casas

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento

AGRADECIMIENTOS

Agradezco a mi profesor guía Tomás Echaveguren por su apoyo constante durante este proceso, acompañándome en mi aprendizaje y en mi desarrollo como futuro investigador. También quiero agradecer a los profesores Juan Carrasco y Alejandro Tudela por dedicarle tiempo a cada duda y cuestionamiento que tuve durante mi formación como ingeniero. Agradezco a mi familia por su apoyo constante, en especial a mi tía Lilian que desde pequeño me mostró lo lindo que es aprender. Finalmente, agradezco a mis amigos por su ayuda y compañía, en especial a Felipe que me ha acompañado y apoyado con su conocimiento técnico durante todo este proceso.

RESUMEN

La gestión de la infraestructura vial en Chile ha dependido históricamente de enfoques reactivos basados en inspecciones visuales, careciendo de modelos predictivos de deterioro adaptados a la realidad nacional. El objetivo de este trabajo es desarrollar un diseño experimental para modelar la degradación en puentes convencionales de tipo viga, integrando principios de ingeniería con técnicas de procesamiento y análisis de datos. En la metodología se definió un vector de características compuesto por diez variables explicativas críticas, abarcando factores climáticos, de tráfico, materialidad y diseño. El procedimiento incorpora técnicas de Inteligencia Artificial en un esquema secuencial: el uso de *embeddings* de Modelos de Lenguaje Grandes (LLM) para la vectorización de variables categóricas, algoritmos de *Machine Learning* para la reducción dimensional y el Método de las Ventanas para la generación de series temporales de datos. Finalmente, se formuló un protocolo de calibración y validación comparativa para nueve arquitecturas de regresión, contrastando métodos probabilísticos y deterministas con modelos de aprendizaje profundo (*Deep Learning*). El estudio demuestra que el diseño experimental propuesto permite transformar los registros históricos de inspección visual en herramientas técnicas de predicción para la gestión de redes a gran escala.

ABSTRACT

Road infrastructure management in Chile has historically relied on reactive approaches based on visual inspections, lacking predictive deterioration models adapted to the national context. The objective of this work is to develop an experimental design to model degradation in conventional beam-type bridges, integrating engineering principles with data processing and analysis techniques. Within the methodology, a feature vector composed of ten critical explanatory variables was defined, encompassing climatic factors, traffic, material properties, and design. The procedure incorporates Artificial Intelligence techniques in a sequential framework: the use of Large Language Model (LLM) embeddings for the vectorization of categorical variables, Machine Learning algorithms for dimensional reduction, and the Sliding Window method for the generation of time-series data. Finally, a calibration and comparative validation protocol was formulated for nine regression architectures, contrasting probabilistic and deterministic methods with Deep Learning models. This study demonstrates that the proposed experimental design enables the transformation of historical visual inspection records into technical prediction tools for large-scale network management.

ÍNDICE DE CONTENIDOS

CAPÍTULO 1: INTRODUCCIÓN.....	1
1.1. Motivación	1
1.2. Objetivos	2
1.3. Alcances	2
1.4. Plan de trabajo.....	2
1.5. Principales resultados	3
1.6. Organización de la memoria de título	3
CAPÍTULO 2: MODELOS DE DETERIORO DE PUENTES	4
2.1. Introducción	4
2.2. El deterioro en puentes.....	4
2.3. Modelación del deterioro de puentes	4
2.4. Taxonomía de modelos de deterioro de puentes	6
2.5. Discusión.....	12
2.6. Conclusiones	14
CAPÍTULO 3: FORMULACIÓN DEL MODELO DE DETERIORO	16
3.1. Introducción	16
3.2. Conceptualización del modelo	16
3.3. Variables explicativas	16
3.4. Conclusiones	19
CAPÍTULO 4: CÁLCULO DEL TAMAÑO MUESTRAL.....	20
4.1. Introducción	20
4.2. Diseño factorial	20
4.3. Tamaño de muestra	21
4.4. Conclusiones	22
CAPÍTULO 5: PROCEDIMIENTO DE CALIBRACIÓN	23
5.1. Introducción	23

5.2.	Procedimiento general.....	23
5.3.	Preparación de datos.....	24
5.4.	Herramientas de modelación.....	30
5.5.	Calibración, validación y sensibilización del modelo.....	38
5.6.	Conclusiones	42
CAPÍTULO 6:	CONCLUSIONES.....	43
GLOSARIO	45
REFERENCIAS	46
ANEXOS	51
ANEXO 1.1	Contribución a los Objetivos de Desarrollo Sostenible	51
ANEXO 3.1	Contenido de base de datos de inspección de puentes	52
ANEXO 3.2	Contenido bases de datos de inventario de puentes.....	53
ANEXO 3.3	Clasificación Koppen-Geiger	55
ANEXO 3.4	Contenido base de datos de tráfico.....	57
ANEXO 3.5	Cálculo del Índice de Calificación de la Condición	58
ANEXO 3.6	Tipos de deterioro	64
ANEXO 5.1	Zonificación Climática de Chile	65

ÍNDICE DE TABLAS

Tabla 2.1 Fortalezas y limitaciones de modelos mecanicistas versus empíricos	12
Tabla 2.2 Fortalezas y limitaciones de modelos deterministas versus estocásticos.....	13
Tabla 3.1 Variables Dependientes.....	17
Tabla 3.2 Variables Independientes	18
Tabla 4.1 Niveles de variables para diseño factorial.....	20
Tabla 5.1 Campos de la base de datos depurada	25
Tabla A. 1 Campos base de datos de inspección visual de puentes	52
Tabla A. 2 Campos base de datos de inventario de puentes	53
Tabla A.3 Codificación clasificación climática Koppen-Geiger	55
Tabla A. 4 Campos de base de datos de tráfico	57
Tabla A. 5 Ponderación Gravedad y Extensión en Calificación de deterioro.....	58
Tabla A. 6 Tipos de Causas con ponderación	58
Tabla A. 7 Tabla de pesos de ponderación por unidad	59
Tabla A.8 Tipos de deterioros con ponderación	64

ÍNDICE DE FIGURAS

Figura 2.1 Modelo de deterioro de la condición de un puente	5
Figura 2.2 Clasificador ortogonal de las herramientas utilizadas en modelos de deterioro	6
Figura 5.1 Procedimiento general del modelo	23
Figura 5.2 Ilustración genérica de ventanas ensambladas (de Solminihac, 2018).....	30
Figura 5.3 Ilustración genérica de una red neuronal artificial.....	35
Figura A.1 Zonificación climática del Norte Grande de Chile.....	65
Figura A.2 Zonificación Climática del Norte Chico y Región Metropolitana de Chile	66
Figura A.3 Zonificación Climática Zona Centro de Chile	67
Figura A.4 Zonificación Climática Zona Centro-Sur de Chile	67
Figura A.5 Zonificación Climática Zona Sur y Austral de Chile.....	68
Figura A.6 Zonificación Climática Zona Austral de Chile	69

CAPÍTULO 1: INTRODUCCIÓN

1.1. Motivación

A lo largo de su vida útil, los puentes están sometidos a un proceso de deterioro continuo, resultado de la interacción entre la estructura y las cargas a las que está sometida. La magnitud de esta degradación se cuantifica a través de Índices de Condición, variable que refleja el estado de la estructura al momento de su inspección.

Sin embargo, la degradación de la condición no presenta una tasa constante en el tiempo. La evidencia empírica muestra que el deterioro suele ser lento en los primeros años y se acelera exponencialmente en etapas avanzadas, debido a la fatiga de los materiales y la acumulación de daños. Dado que las inspecciones proporcionan solo una visión estática y puntual del estado del puente, es necesario el desarrollo de modelos predictivos de la pérdida progresiva del Índice de Condición.

La modelación del deterioro se ha abordado desde diversas escalas operativas, distinguiendo entre el seguimiento detallado a nivel de proyecto y la gestión estratégica a nivel de red. La calibración de los modelos de deterioro frecuentemente se ve condicionada por la discontinuidad temporal y la dispersión de las inspecciones visuales, sumado a la alta heterogeneidad de las variables explicativas, factores que limitan la aplicación de algoritmos de regresión que presuponen una evolución continua de la condición del puente.

En consecuencia, resulta necesario el desarrollo de un diseño experimental que permita calibrar un modelo capaz de reconstruir las trayectorias de deterioro y de integrar la fenomenología del deterioro a partir de datos heterogéneos, logrando predecir la degradación de la condición de los puentes con la mayor precisión posible, realizando una comparación de las distintas herramientas de regresión disponibles actualmente.

Esta memoria de título tuvo como objetivo proponer un diseño experimental para la predicción del deterioro en puentes convencionales tipo viga, considerando las distintas variables que influyen directamente en la condición de los puentes, acompañado de un proceso de selección y comparación de la variedad de herramientas de modelación disponibles.

1.2. Objetivos

1.2.1 Objetivo general

Proponer un diseño experimental para la predicción del deterioro en puentes convencionales tipo viga, considerando factores de tráfico, clima, estructuración, geometría y materialidad.

1.2.2 Objetivos específicos

- a) Conceptualizar el fenómeno del deterioro de puentes y sus variables explicativas.
- b) Formular un modelo de deterioro de puentes.
- c) Determinar el tamaño muestral necesario para asegurar la potencia estadística del modelo.
- d) Seleccionar técnicas de calibración del modelo de deterioro de puentes.

1.3. Alcances

El modelo de deterioro por desarrollar es a nivel de red y de primera fase, lo que implica que no se considerará el efecto de las acciones de mantenimiento ni la influencia de eventos naturales o antrópicos, que puedan causar un daño súbito. La calibración considera solo los puentes convencionales tipo viga, excluyendo del análisis los puentes de madera.

1.4. Plan de trabajo

El plan de trabajo se estructuró en cuatro etapas, comenzando con una revisión exhaustiva del estado del arte para establecer una taxonomía de las técnicas de predicción existentes y definir el enfoque de la investigación. A partir de este fundamento teórico, se procedió a la conceptualización del modelo y al diseño experimental, definiendo las variables explicativas, junto al diseño factorial y su tamaño muestral necesario para asegurar la potencia estadística. Finalmente, el trabajo culminó estableciendo las fases necesarias para la calibración, validación y sensibilización del modelo de deterioro, específicas para cada una de las herramientas de regresión seleccionadas.

1.5. Principales resultados

Se desarrolló un diseño experimental para predecir el deterioro de puentes convencionales tipo viga en Chile. Se identificaron preliminarmente diez variables explicativas basadas en la fenomenología del deterioro, las cuales son la edad del puente, el número de cepas, el número de vigas, el espesor de la losa, el tipo de carpeta, la tipología de las vigas, el ancho total, la longitud total, el Tránsito Medio Diario Anual (TMDA) y la clasificación climática. Mediante un análisis de potencia estadística, se determinó que se requiere un tamaño muestral mínimo de 1.400 puentes para asegurar la validez de las predicciones. Finalmente, se seleccionaron técnicas de Inteligencia Artificial en un esquema secuencial, integrando un proceso de vectorización para el tratamiento de datos, algoritmos de selección de variables para la reducción dimensional, junto a un protocolo de calibración, validación y sensibilización para nueve herramientas de regresión.

1.6. Organización de la memoria de título

La memoria se organiza en seis capítulos. En el primero se presenta la motivación, objetivos, el plan de trabajo y los principales resultados. En el capítulo dos se describen los distintos modelos de deterioros presentes en el estado del arte, junto a una propuesta de clasificación. En el tercer capítulo, se presenta la conceptualización del modelo, junto a su fenomenología, variables explicativas y los procesos subyacentes. En el capítulo cuatro se presenta el diseño factorial junto al procedimiento aplicado para determinar el tamaño muestral mínimo. En el capítulo cinco se concentra en las fases del modelo de deterioro de puentes, incluyendo la selección de herramientas para cada una de ellas y su proceso de calibración, junto a su protocolo de validación y sensibilización. Finalmente, en el sexto capítulo se presentan las conclusiones y las líneas futuras de investigación.

CAPÍTULO 2: MODELOS DE DETERIORO DE PUENTES

2.1. Introducción

Este capítulo aborda la fenomenología del deterioro estructural y funcional en puentes, destacando su rol en los sistemas de gestión de puentes (SGP). Asimismo, se presenta una taxonomía que clasifica las investigaciones según sus enfoques mecanicistas, empíricos, deterministas y estocásticos. Finalmente, se presenta una discusión de los hallazgos principales de la revisión del estado del arte.

2.2. El deterioro en puentes

El deterioro de puentes es el resultado de la interacción entre la estructura y las cargas a las que está sometida. Estas cargas pueden ser cíclicas o permanentes. Las cargas climáticas y de tráfico se consideran cíclicas, mientras que el peso propio de la estructura es considerado una carga permanente. (AASHTO, 2020)

Debido a las sollicitaciones recurrentes que está sometido el puente, el deterioro evoluciona a lo largo del tiempo. Esta trayectoria de degradación se postula como un proceso dependiente de factores climáticos, materialidad, estructuración, geometría y tráfico. (Frangopol y Soliman, 2016)

2.3. Modelación del deterioro de puentes

Los modelos de deterioro de puentes son una representación matemática o computacional que estima la velocidad a la que la condición de un puente se degradará con el tiempo (De Dieu et al., 2020). La modelación de este deterioro es un proceso necesario para mantener los puentes operativos y en buen estado, estando presente en la gran mayoría de los sistemas de gestión de puentes (SGP). Un SGP es un procedimiento formal el cual permite analizar los datos de los activos con el propósito de predecir las condiciones futuras del puente, estimando las necesidades de mantenimiento de una red (AASHTO, 1993).

Uno de los pilares de estos sistemas es la recolección de datos, la cual se realiza mediante catastro e inspección de puentes, procesos sistemáticos y detallados que se llevan a cabo para evaluar el estado

y la condición de los elementos de este (de Solminihac, 2018). A partir de los resultados de una inspección, se puede obtener un conjunto de indicadores para definir o calcular índices de condición, que determinan el estado global de la infraestructura. Calificar la condición o funcionalidad de los puentes o de sus componentes individuales es esencial para la modelación del deterioro (Iacovino et al., 2021).

Los modelos de deterioro permiten predecir la condición del puente a lo largo de su ciclo de vida. La eficacia y fiabilidad de gran parte de los procesos de un SGP dependen directamente de la precisión de pronóstico de sus modelos de deterioro. Un modelo impreciso puede llevar a intervenciones prematuras e innecesarias o en acciones tardías que permiten que problemas menores se conviertan en reparaciones mayores y costosas, comprometiendo potencialmente la funcionalidad y seguridad estructural (De Dieu et al., 2020).

La Figura 2.1 ilustra un ejemplo de un tipo modelo de deterioro basado en la degradación de la condición de un puente en función de su edad. La curva exhibe un comportamiento no lineal con tres fases claras en el ciclo de vida. Un periodo inicial de estabilidad donde la condición se mantiene cercana a 100, indicando una degradación lenta. Luego, un punto de inflexión donde la pendiente se vuelve pronunciada, reflejando una aceleración en el deterioro de la estructura y finalmente una estabilización en las calificaciones inferiores, representando el estado de servicio mínimo de la infraestructura.

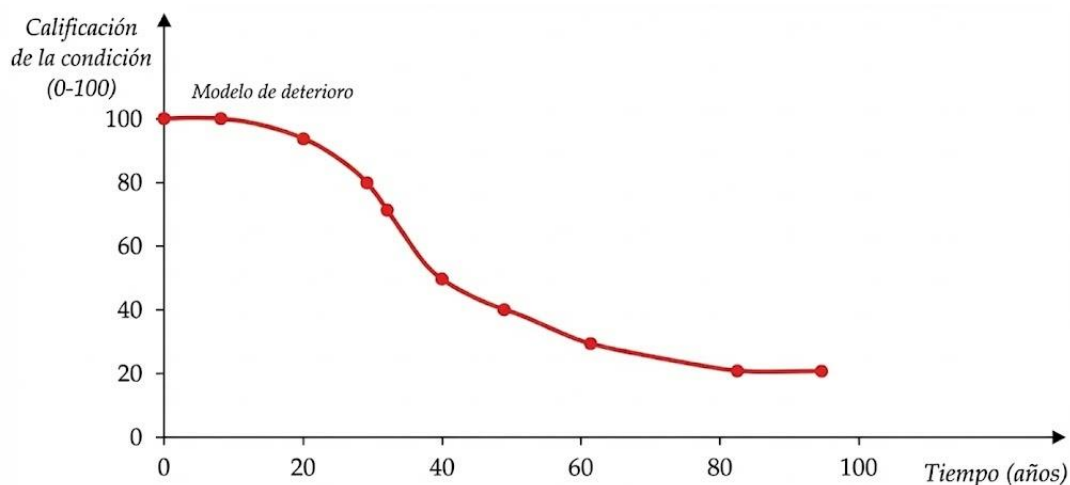


Figura 2.1 Ejemplo de Modelo de deterioro de la condición de un puente

2.4. Taxonomía de modelos de deterioro de puentes

Los modelos de deterioro de puentes se pueden agrupar distinguiendo entre distintos enfoques de modelación. La primera agrupación se realiza en base a las variables explicativas que se utilizan para modelar el deterioro. Los modelos mecanicistas se basan en los procesos físicos de la degradación, *versus* a modelos empíricos, que se basan en patrones estadísticos de datos históricos. La siguiente agrupación considera la naturaleza de la predicción entregada por el modelo. Los modelos deterministas, que predicen un único futuro cierto, frente a modelos estocásticos, que contemplan la incertidumbre en la predicción, entregando una distribución de probabilidad de condiciones futuras (Moghtadernejad et al. 2024).

Estos dos esquemas de clasificación son ortogonales. Un modelo puede pertenecer a cualquiera de los cuatro cuadrantes resultantes, ya que un modelo mecanicista puede ser determinista o estocástico, y lo mismo ocurre con un modelo empírico. La clasificación resultante se muestra en la Figura 2.2, donde se encuentran las distintas investigaciones encontradas en la revisión del estado del arte.

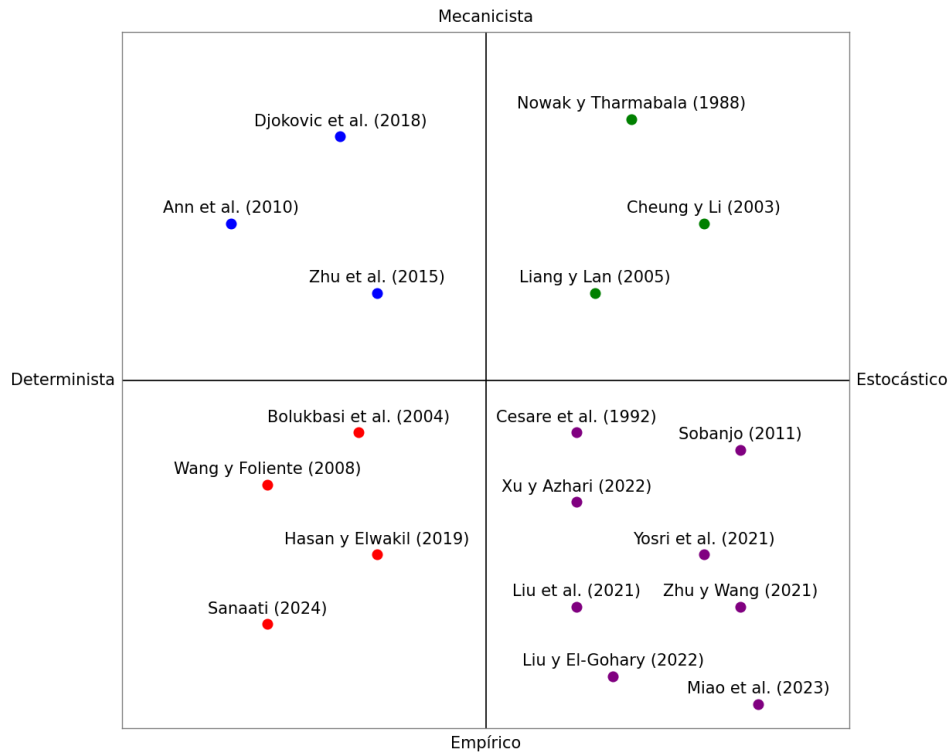


Figura 2.2 Clasificador ortogonal de los modelos de deterioro

2.4.1 Modelos mecanicistas-deterministas

Los modelos de deterioro mecanicistas-deterministas se fundamentan en los procesos físicos que gobiernan la degradación de los materiales. El enfoque mecanicista utiliza leyes fundamentales, como las leyes de difusión de Fick o la ley de Paris, para cuantificar el comportamiento de los componentes de la estructura a lo largo del tiempo. Por otro lado, el componente determinista establece que, para un conjunto definido de parámetros de entrada, el modelo calculará un resultado único y cuantificable.

Ann et al. (2010) evaluaron la vida útil de un puente de hormigón, basándose en el riesgo de carbonatación. Los autores midieron la profundidad de carbonatación (X) y de recubrimiento (R) en 113 puntos de una estructura con 18 años de servicio, clasificando el estado del hormigón en sano, fisurado y juntas. A partir de estos datos determinaron experimentalmente el coeficiente de carbonatación (C) mediante la Ecuación 2.1, la cual es una variante de la ley de difusión de Fick. Posteriormente, definieron la función de estado límite $g(t)$ mediante la Ecuación 2.2, donde $R(t)$ representa la profundidad de recubrimiento y $X(t)$ profundidad de carbonatación. Se estableció que el fin de la vida útil ocurre cuando la carbonatación ha superado el recubrimiento ($g(t) < 0$).

$$X = C\sqrt{t} \quad (2.1)$$

$$g(t) = R(t) - X(t) \quad (2.2)$$

Zhu et al. (2015) proponen un método integral para evaluar la vida útil residual de puentes de hormigón armado, centrándose en el efecto de la corrosión y la fatiga. Utilizaron un modelo de elementos finitos para simular el paso de vehículos pesados sobre el puente, con el fin de obtener el coeficiente de impacto dinámico (μ) y el historial de tensiones ($\Delta\sigma$). Para la medición del efecto de la corrosión (λ), usaron fórmulas empíricas preestablecidas. La predicción de la vida útil residual la obtuvieron a partir de la Ecuación 2.3, donde T es la vida útil residual, $ADTT$ es el tráfico diario promedio de camiones y α es la tasa anual de crecimiento de tráfico.

$$T = \frac{\ln\left(\frac{10^{8.98} * \ln(1 + \alpha)}{365 * ADTT * (\mu * \lambda * \Delta\sigma)^{1.5}} + 1\right)}{\ln(1 + \alpha)} \quad (2.3)$$

Djoković et al. (2018) desarrollan un modelo para estimar la vida útil remanente de puentes de acero, basándose en el efecto de la fatiga. Los autores analizaron la propagación de una grieta desde un tamaño inicial (α_i) hasta una longitud crítica (α_{cr}), estimando la tasa de crecimiento mediante la Ley de Paris. La vida útil remanente (N) la calcularon mediante la Ecuación 2.4, la cual es la integral de la Ley de Paris, donde C y m corresponden a constantes propias del material y ΔK al factor de intensidad de tensiones. Estos datos fueron obtenidos a partir de la información técnica del puente.

$$N = \int_{\alpha_i}^{\alpha_{cr}} \frac{da}{C(\Delta K)^m} \quad (2.4)$$

2.4.2 Modelos mecanicistas-estocásticos

Los modelos mecanicistas-estocásticos reconocen que los fenómenos de deterioro están sujetos a una variabilidad intrínseca. Si bien el componente mecanicista mantiene el uso de leyes físicas fundamentales para modelar la degradación, el enfoque estocástico acepta la incertidumbre inherente del proceso de deterioro, entregando una distribución de probabilidad del estado de la estructura o una curva de probabilidad de fallo a lo largo del tiempo.

Nowak y Tharmabala (1988) es un artículo considerado seminal dentro de los modelos de deterioro de puentes. Los investigadores proponen un marco metodológico, basado en la confiabilidad, para evaluar la seguridad real de un puente existente. Realizaron un análisis de confiabilidad probabilística, calculando un índice de confiabilidad inicial (β) para los componentes de la estructura. Posteriormente, fue calibrado por los datos de capacidad de carga obtenidos de la realización un ensayo con vehículos pesados en el puente real. En la Ecuación 2.5 se encuentra el índice de confiabilidad utilizado por los autores, donde m_R y S_R son la media y varianza de la resistencia de los componentes, y m_Q y S_Q son la media y varianza del efecto de carga que existe sobre el puente.

$$\beta = \frac{m_R - m_Q}{\sqrt{S_R^2 + S_Q^2}} \quad (2.3)$$

Cheung y Li (2003) desarrollan un análisis probabilístico de fatiga y fractura en puentes de acero. Proponen el método de las fajas finitas para la estimación de la falla por fatiga, simulando el paso de vehículos pesados. La falla por fractura ocurre cuando la grieta llega hasta la longitud crítica, calculado mediante la ley de Paris. Junto a estos resultados, realizan un análisis de confiabilidad, donde mediante una distribución gaussiana integran la probabilidad de falla de los componentes del puente.

Liang y Lan (2005) evalúan el deterioro de pilotes de hormigón armado considerando la combinación de la corrosión por carbonatación, el ingreso de iones de cloruro y el ataque por sulfato. Realizaron la extracción de núcleos de hormigón cada 5 años de un puente real para medir la profundidad de carbonatación, ingreso de cloruros y expansión por sulfatos. Aplican la teoría de probabilidad de límites estrechos, que integra la probabilidad de fallo de cada mecanismo de deterioro con la capacidad de carga del pilote. El resultado es un cálculo de la probabilidad de falla total de los pilotes a lo largo de su vida útil.

2.4.3 Modelos empíricos-deterministas

Los modelos empírico-deterministas buscan predecir la condición futura de los componentes a partir de datos de inspecciones pasadas. Mientras que el componente empírico implica que las ecuaciones de deterioro se calibran mediante regresiones ajustadas a grandes volúmenes de datos de inspección, el componente determinista establece una relación funcional fija entre la edad, las variables explicativas y su nivel de degradación.

Bolukbasi et al. (2004) utilizaron las bases de datos del *National Bridge Inventory* (NBI), usándolo de base empírica para desarrollar un modelo predictivo de la condición futura de los componentes del puente, estimado a través de una regresión mediante un polinomio de tercer grado. Las variables explicativas propuestas por los investigadores son la edad del puente, materialidad, tipo de vía y volumen de tráfico. Wang y Foliente (2008) también utilizan las bases de datos del NBI. Proponen un método que identifica los factores más influyentes mediante un análisis de varianza (ANOVA). Posteriormente, aplican regresión lineal a una serie de indicadores estadísticos anuales para cuantificar la tasa de deterioro y analizar cómo cambia la condición de un puente a lo largo del tiempo.

Hasan y Elwakil (2019), desarrollaron modelos de deterioro para superestructuras de puentes de hormigón pretensado en California, utilizando análisis de regresión múltiple sobre datos del NBI para identificar las variables más significativas que afectan la condición de la superestructura. Se generaron cuatro ecuaciones de regresión lineal múltiple específicas para cada tipo de superestructura (Losa, Vigas Múltiples, Vigas T, Vigas Cajón). Las variables explicativas utilizadas fueron de tráfico, geometría, edad y porcentaje de vehículos pesados.

Sanaati (2024) utilizó un modelo de Máquinas de soporte vectorial para predecir la profundidad de la socavación en pilares de puentes. El modelo entrega la profundidad de socavación predicha en milímetros, a partir del diámetro del pilar, velocidad de flujo, profundidad de flujo y tamaño medio del sedimento.

2.4.4 Modelos empíricos-estocásticos

Los modelos empíricos-estocásticos combinan la naturaleza de los datos de entrada, basados en registros históricos de inspección visual y calificaciones de condición, con un tratamiento matemático estocástico que reconoce la incertidumbre inherente al proceso de deterioro. Este tipo de modelos estiman una distribución probabilística de posibles condiciones futuras.

Cesare et al. (1992) establecieron un método basado en cadenas de Markov a partir de bases de datos de inspección e inventario del Departamento de Transporte del Estado de Nueva York. En su enfoque, cada calificación de condición corresponde a un estado en el proceso markoviano, prediciendo el cambio de un estado a otro de los componentes del puente. La materialidad, estructuración y geometría fueron las variables explicativas utilizadas en este artículo.

Sobanjo (2011) también desarrolla un modelo basado en cadenas de Markov, utilizando las bases de datos del NBI. Sin embargo, demostró empíricamente que la probabilidad de que un puente se deteriore a un estado peor no es constante, sino que aumenta con el tiempo que el elemento ha permanecido en su estado actual. Propone el uso de la distribución de Weibull para describir este fenómeno, proponiendo en su lugar el uso de un modelo semi-markoviano como una representación más realista del proceso de deterioro por envejecimiento.

Xu y Azhari (2022) utilizan un modelo markoviano bidimensional, en la cual cada estado combina simultáneamente la edad de la estructura y su condición, utilizando bases de datos del NBI. Complementariamente, el modelo también incorpora una función de riesgo proporcional basada en la distribución de Weibull, lo que introduce una dependencia explícita del deterioro respecto del tiempo. Este estudio solo contempla puentes de acero pertenecientes al estado de New York.

Yosri et al. (2021) utilizan un modelo de deterioro Markoviano dentro de un marco de optimización mediante algoritmos genéticos (GA), logrando replicar las trayectorias de deterioro observadas en puentes de concreto y acero, tanto rehabilitados como no rehabilitados. Analizaron solo la materialidad del puente y la edad como variables explicativas del deterioro. Los datos fueron extraídos del Ministerio de Transporte de Ontario, Canadá.

Liu et al. (2021) propone un modelo híbrido basado en *deep learning*, donde una red neuronal convolucional profunda (CNN) es entrenada para estimar los parámetros de una cadena de Markov. Este método aprovecha la capacidad de la CNN para aprender automáticamente las relaciones entre múltiples variables y el proceso de deterioro. Incluyen 24 factores de entrada, entre ellos se encuentra la edad, variables de tráfico, diseño, material, protección del tablero, clima y ubicación. Usaron los datos disponibles del NBI.

Zhu y Wang (2021) también implementan un modelo híbrido basado en CNN y unidades recurrentes (RNN), con el objetivo de predecir las calificaciones futuras de condición de elementos de puentes en Texas. Además, incluye un algoritmo mejorado de ReliefF para seleccionar de manera objetiva las variables más influyentes en la calificación de la condición. Inicialmente se consideraron 25 características, de las que solo 13 fueron seleccionadas, incluyendo condición previa, edad, tipo de diseño, material, tráfico y ancho de calzada.

Liu y El-Gohary (2022) también utilizan las bases de datos del NBI junto a una RNN para modelar el deterioro y predecir la condición futura. Sin embargo, extienden el enfoque al incorporar datos heterogéneos, como informes técnicos de inspección y variables climáticas, manejándolos mediante una reducción de dimensionalidad de los datos y un aprendizaje de sensibilidad al costo, mejorando la precisión de la predicción en un 15%.

Miao et al. (2023) proponen un marco de análisis que procesa y fusiona datos de inspección heterogéneos, incluyendo tanto series temporales como imágenes de defectos superficiales. Para ello, emplean un conjunto de herramientas de aprendizaje profundo: una RNN para analizar los datos de series temporales, y modelos de CNN para identificar y cuantificar defectos en las imágenes. Posteriormente, los resultados de cada análisis son convertidos en grados de deterioro probabilísticos mediante un Sistema de Inferencia Difusa Refinado (RFIS). El estudio utiliza datos de inspección de puentes en Japón, donde incluyen variables como edad, tipo de tablero, nivel de iones cloruro, nevadas, temperatura mínima y volumen de tráfico.

2.5. Discusión

A partir de la clasificación taxonómica y el análisis de las herramientas presentes en el estado del arte, se determina que no existe un modelo universalmente superior, sino que la elección depende de la disponibilidad de datos y los objetivos de gestión. La Tabla 2.1 y 2.2 sintetizan las disyuntivas fundamentales identificadas en la literatura.

Tabla 2.1 Fortalezas y limitaciones de modelos mecanicistas versus empíricos

Clase de Modelo	Principio Central	Fortalezas	Limitaciones
Mecanicista	Modela los procesos físicos y químicos subyacentes del deterioro utilizando ecuaciones fundamentales	Proporciona una explicación causal; alta precisión si está bien calibrado; bueno para evaluar intervenciones específicas	Computacionalmente intensivo; requiere datos que a menudo no están disponibles para redes enteras
Empírico	Identifica patrones y correlaciones estadísticas en datos históricos de inspección para predecir condiciones futuras	Práctico para grandes redes; puede usar datos fácilmente disponibles; puede actualizarse a medida que se recopilan nuevos datos	Naturaleza de "caja negra"; limitado por la calidad y disponibilidad de los datos.

Tabla 2.2 Fortalezas y limitaciones de modelos deterministas versus estocásticos

Clase de Modelo	Principio Central	Fortalezas	Limitaciones
Determinista	Asume que el deterioro ocurre a una tasa conocida y fija, prediciendo una única condición futura cierta	Simple de desarrollar y entender; computacionalmente trivial	No tiene en cuenta la aleatoriedad e incertidumbre inherentes en el proceso de deterioro
Estocástico	Trata el deterioro como un proceso aleatorio, prediciendo una distribución de probabilidad de posibles condiciones futuras	Cuantifica la incertidumbre; representación más realista del proceso	Puede ser más complejo matemáticamente; los resultados dependen de la distribución de probabilidad elegida

Las investigaciones con un enfoque mecanicista se centraron en el deterioro químico del hormigón y del acero, utilizando leyes de difusión para modelar la degradación de los puentes. Sin embargo, se evidenció que la aplicación determinista pura de estas leyes es insuficiente; la precisión predictiva mejora en un 140% cuando se integran factores de inspección, como la presencia de grietas que aceleran la difusión en el modelo de Hájková et al. (2018). Por otro lado, Borah et al. (2020) demostró que la incorporación de variables probabilísticas para manejar la incertidumbre de los materiales y el ambiente pueden reducir la vida útil proyectada en un 30%.

Los modelos empíricos analizados se centran en capturar la manifestación macroscópica del deterioro, utilizando bases de datos de inspecciones para evaluar la degradación de la condición de los puentes a lo largo de su ciclo de vida. Los enfoques deterministas como el de Bolukbasi et al. (2004) demostraron que es posible generar curvas de deterioro polinómicas, utilizando la edad y su calificación de condición. Sin embargo, la mayoría de estos modelos asumen una tasa de deterioro constante a lo largo del tiempo, lo cual no refleja la realidad. Factores como la edad, el tráfico y el clima alteran estas tasas, por lo que es necesario un enfoque que contemple estas variables (Moghtadernejad et al., 2024).

La frontera actual de los modelos de deterioro reside en la capacidad para capturar la incertidumbre propia del fenómeno de deterioro y la validación con datos reales. Aunque las variables utilizadas en los modelos estocásticos y deterministas comparten una base común (dentro de los modelos empíricos) su definición y medición varían según la fuente de datos y la metodología específica. La variable principal son los resultados de las inspecciones visuales o BCI obtenidos mediante evaluaciones rutinarias (Moomen et al., 2016).

La edad del puente es otro factor determinante, medida generalmente desde el año de construcción o ajustada a la fecha de la última inspección. Respecto a las variables estructurales y de operación, se destacan el material y el diseño geométrico del puente, junto a las variables de tráfico, medidas cuantitativamente a través del Tráfico Medio Diario Anual (TMDA) y el porcentaje de vehículos pesados. Toda esta información se extrae de grandes bases de datos de gestión, como los archivos del Departamento de Transporte del Estado de Nueva York y el NBI de EE. UU., lo que asegura que los modelos se basen en un comportamiento histórico real.

Los modelos de deterioro estructural de puentes basados en inteligencia artificial son una evolución metodológica frente a los enfoques estadísticos tradicionales, permitiendo una representación más precisa, adaptable y escalable de los procesos de degradación. Liu et al. (2021), Zhu y Wang (2021), Liu y El-Gohary (2022), Choi y Kong (2023), evidencia una clara transición hacia el uso de redes neuronales, que modelan directamente la evolución del deterioro estructural a partir de datos históricos y variables explicativas, prescindiendo de supuestos explícitos sobre funciones de transición o formulaciones paramétricas rígidas, logrando precisiones superiores al 90%.

2.6. Conclusiones

El deterioro de puentes se define como la interacción entre la estructura y las cargas cíclicas y permanentes, constituyendo un proceso dependiente de factores climáticos, materialidad, estructuración, geometría y tráfico. La modelación de este fenómeno permite estimar la degradación del índice de condición en el tiempo, lo cual es un componente de los sistemas de gestión de puentes para predecir necesidades de mantenimiento.

Los modelos de deterioro de puentes se clasifican en modelos mecanicistas, basados en leyes físicas fundamentales, y modelos empíricos, basados en patrones estadísticos de datos históricos, los cuales operan bajo esquemas deterministas o estocásticos.

Se determinó que la modelación del deterioro depende de un conjunto específico de variables explicativas, identificándose como factores principales los resultados de las inspecciones visuales o Índice de Condición del Puente (BCI), la edad de la estructura, la materialidad, el diseño geométrico, el Tráfico Medio Diario Anual (TMDA) y el porcentaje de vehículos pesados. Esto define al enfoque empírico como el más apto para calibrar modelos de deterioro a nivel de red.

El estado del arte muestra que la selección de un modelo de deterioro está sujeto a la disponibilidad de datos y los objetivos de gestión. La integración de datos de inspección y la cuantificación de la incertidumbre resultan necesarios para incrementar la precisión predictiva de los modelos de deterioro a nivel de red. Bajo este escenario, surge la oportunidad de implementar herramientas de Aprendizaje Automático (ML) para abordar la complejidad multivariante de los datos de inspección. A diferencia de las formulaciones deterministas o estocásticas convencionales, que a menudo se ven limitadas por la alta dimensionalidad, el ML permite integrar un vector de características sin los sesgos de una función de transición predefinida. Esta capacidad de procesamiento masivo, potenciada por la vectorización avanzada de variables categóricas y técnicas de reducción dimensional, posibilita la captura de patrones no lineales en el deterioro

CAPÍTULO 3: FORMULACIÓN DEL MODELO DE DETERIORO

3.1. Introducción

Este capítulo establece la formulación del modelo de deterioro bajo un enfoque empírico, delimitando el alcance del estudio a puentes convencionales tipo viga y excluyendo intervenciones de mantenimiento mayor. Se formalizan las variables dependientes e independientes que componen el vector de características para predecir la calificación de condición.

3.2. Conceptualización del modelo

Para abordar el fenómeno del deterioro, se ha configurado un modelo preliminar, el cual se fundamenta en la identificación de patrones estadísticos en los datos históricos de inspección.

El modelo se centra exclusivamente en puentes convencionales de tipo viga, que constituyen la tipología predominante en el inventario de puentes. Se asume que la degradación es un proceso multifactorial y progresivo, dependiente de características intrínsecas y agentes externos. Se excluye explícitamente el efecto restaurador de las acciones de mantenimiento mayor o rehabilitación, así como los daños súbitos provocados por eventos naturales o antrópicos. El propósito final del modelo es predecir la Calificación de Condición en función de la edad del puente y su vector de características.

Este diseño conceptual permite transformar los datos tabulares disponibles en curvas de comportamiento temporal, utilizando para ello las variables explicativas que se detallan a continuación.

3.3. Variables explicativas

Las variables explicativas del deterioro se extrajeron de las bases de datos de inspección visual de puentes y la base de datos de tráfico de la red vial, ambas pertenecientes al Ministerio de Obras Públicas (MOP). Solo la variable climática se obtuvo del artículo Sarricolea et al. (2017), el cual realizó una regionalización climática del territorio chileno. En el Anexo 3.1 se encuentra el contenido de la base de datos de inspección de puentes. En el Anexo 3.2 se encuentran los campos de la base de

datos de inventario de puentes. En el Anexo 3.3 se encuentra la codificación de la regionalización climática. En el Anexo 3.4 se encuentra el contenido de la base de datos de tráfico.

Se obtuvo una selección de las variables con mayor influencia en la calificación de condición, basándose en el estado del arte y en la hipótesis de que el deterioro de un puente es un fenómeno influenciado y dependiente de factores climáticos, de materialidad, estructuración, inspección, inventario, geometría y tráfico.

3.3.1 Variables dependientes

Las variables dependientes de este modelo representan la cuantificación del desempeño de la estructura. Estas constituyen las variables objetivo que el modelo buscará predecir en función del vector de características del puente. El Número total de deterioros y la Calificación final del puente son variables pertenecientes a este grupo.

La Calificación final del puente corresponde a un Índice de Calificación de Condición calculado automáticamente a partir de los algoritmos del MOP, basados en los datos de inspección visual. Su rango de valores de la Calificación oscila entre 0 y 100, ya que posee una escala porcentual. En el Anexo 3.5 se encuentra el procedimiento de cálculo de esta calificación.

El número total de deterioros es una variable de naturaleza numérica discreta y representa los deterioros observados en la estructura durante el proceso de inspección. Cada entrada de esta base de datos es un deterioro. Puede tomar cualquier valor natural desde 0 hasta N, siendo N el total de entradas por puente encontradas en la base de datos de inspección de la Dirección de Vialidad.

Tabla 3.1 Variables Dependientes

Variabes	Clase	Fuente	Rango de valores
Número total de deterioros	Numérica	MOP	0 - N
Calificación final del puente	Numérica	MOP	0 - 100

3.3.2 Variables independientes

Las variables independientes constituyen el conjunto de factores de entrada del modelo, conformando el vector de características (X) que describen cada puente. Basándose en la hipótesis de que el deterioro es un fenómeno multifactorial, estas son variables obtenidas de las distintas fuentes de información. En la Tabla 3.2 se puede observar las variables independientes con sus respectivos rangos de valores, obtenidos a partir de las bases de datos de inventario e inspección del Ministerio de Obras Públicas y de la regionalización climática de Chile propuesta por Sarricolea et al. (2017).

Tabla 3.2 Variables Independientes

VARIABLES	Clase	Fuente	Rango de valores preliminar
Edad del puente	Numérica	MOP	0 - 100
Número de Cepas	Numérica	MOP	0 - 50
Número de Vigas	Numérica	MOP	2 - 10
Espesor Losa	Numérica	MOP	0 – 0.70
Carpeta	Categórica	MOP	Hormigón / Asfalto
Vigas	Categórica	MOP	Acero / Hormigón
Ancho total	Numérica	MOP	0 - 50
Longitud Total	Numérica	MOP	0 - 3000
TMDA	Numérica	MOP	0 – 50000
Clasificación Climática	Categórica	Sarricolea et al. (2017)	*Anexo 3.3*
Deterioro	Categórica	MOP	*Anexo 3.2*
Causa del deterioro	Categórica	MOP	*Anexo 3.1*
Gravedad del deterioro	Categórica	MOP	*Anexo 3.1*
Extensión del deterioro	Categórica	MOP	*Anexo 3.1*
Unidad del deterioro	Categórica	MOP	*Anexo 3.1*
Calificación del deterioro (Cd)	Numérica	MOP	0-100

3.4. Conclusiones

Se configuró un modelo de deterioro de carácter empírico restringido a puentes convencionales tipo viga, el cual asume la degradación como un proceso multifactorial y progresivo por envejecimiento, excluyendo explícitamente el efecto recuperador de intervenciones mayores o daños por eventos catastróficos. Esta formulación define como objetivo la estimación de la Calificación de Condición en una escala de 0 a 100, permitiendo transformar datos tabulares en curvas de comportamiento temporal mediante la correlación de variables dependientes de desempeño estructural con un vector de características compuesto por diez variables independientes. La selección de estas variables valida la hipótesis de que el deterioro depende de la interacción conjunta de factores de inventario, materialidad, estructuración, geometría y las cargas cíclicas de tráfico y clima, estableciendo los parámetros de entrada necesarios para las fases posteriores de calibración.

CAPÍTULO 4: CÁLCULO DEL TAMAÑO MUESTRAL

4.1. Introducción

Este capítulo establece el diseño experimental estadístico y calcula el tamaño muestral mínimo necesario para asegurar la validez predictiva y capacidad de generalización del modelo de deterioro.

4.2. Diseño factorial

Se agruparon las variables que son completamente independientes entre sí en columnas, como lo son el clima, la edad y el Tráfico Medio Diario Anual (TMDA) del tramo. Luego las demás variables se agruparon en filas, generando 7776 celdas, las cuales corresponde a un par ordenado de número de deterioros y la Calificación de la condición. En el Anexo digital 4.1 se encuentra la matriz factorial completa utilizada para este análisis. En la Tabla 4.1 se observan las variables seleccionadas.

Para realizar el diseño factorial balanceado se ocupó un set de variables a 10. Además, se dividieron las variables entre dos y tres niveles. La división por niveles se realizó en base a una distribución de las variables, separándolos de tal manera que queden n cantidad de puentes en cada nivel.

Tabla 4.1 Niveles de variables para diseño factorial

ID	Variable	Niveles	Nivel 1	Nivel 2	Nivel 3
1	edad	2	0-30	31 o más	-
2	carpeta	2	Hormigón	Asfalto	-
3	vigas	2	Acero	Hormigón	-
4	numero_vigas	2	2	3 o más	-
5	numero_cepas	2	0	1 o más	-
6	longitud_total	3	0-40	40-100	100 o más
7	ancho total	3	0-7	7-10	10 o más
8	espesor_losa	2	0-0.25	0.25 o más	-
9	Zonificación Climática	3	Árido	Templado	Polar
10	TMDA	2	Bajo	Alto	-

4.3. Tamaño de muestra

Para asegurar la representatividad estadística del estudio y la capacidad del modelo para generalizar los resultados sobre el deterioro de los puentes, se definió el tamaño muestral total (N) considerando la estructura del diseño factorial propuesto.

El diseño muestral se plantea bajo el supuesto de un diseño balanceado, buscando homogeneidad en el valor de n (sujetos por combinación) a través de los distintos grupos m. Esto para garantizar la ortogonalidad de los factores y maximizar la potencia estadística en el análisis de las interacciones entre variables constructivas y climáticas.

4.3.1 Procedimiento de cálculo

La determinación del valor mínimo de N se realizó mediante un análisis de potencia *a priori* utilizando el software estadístico G*Power (versión 3.1.9.7). Dado que el objetivo es modelar el Índice de Condición a partir del set de variables, se seleccionó la familia de pruebas F-tests y el test estadístico de Regresión Lineal Múltiple.

Los parámetros de entrada para el cálculo fueron los siguientes:

- Nivel de Significancia (α) = 0.05. Se asume un riesgo del 5% de cometer un error de Tipo I.
- Potencia Estadística ($1-\beta$) = 0.95. Se establece una alta probabilidad (95%) de detectar efectos significativos si estos existen, reduciendo drásticamente el riesgo de error.
- Número de Predictores = 14. Este valor resulta de la descomposición de las 10 variables independientes originales, considerando los grados de libertad aportados por las variables dicotómicas y politómicas del diseño.
- Efecto tamaño (f^2) = 0.02. Se optó por un tamaño de efecto pequeño (conservador). Esta decisión se justifica debido a la naturaleza compleja y multifactorial de la degradación de infraestructura, donde se espera que cada variable aporte una influencia sutil al modelo general, requiriendo una mayor sensibilidad para su detección.

4.3.2 Resultados

La Figura 4.1 ilustra el análisis de sensibilidad para el modelo de Regresión Lineal Múltiple con 14 predictores. Se observa una familia de curvas que representan la ganancia de Poder Estadístico en el eje y, en función del aumento del Tamaño de la Muestra Total en el eje X, para distintos Tamaños del Efecto f^2 . El resultado de esta proyección indica que el tamaño muestral mínimo necesario es de $N = 1.400$ puentes.

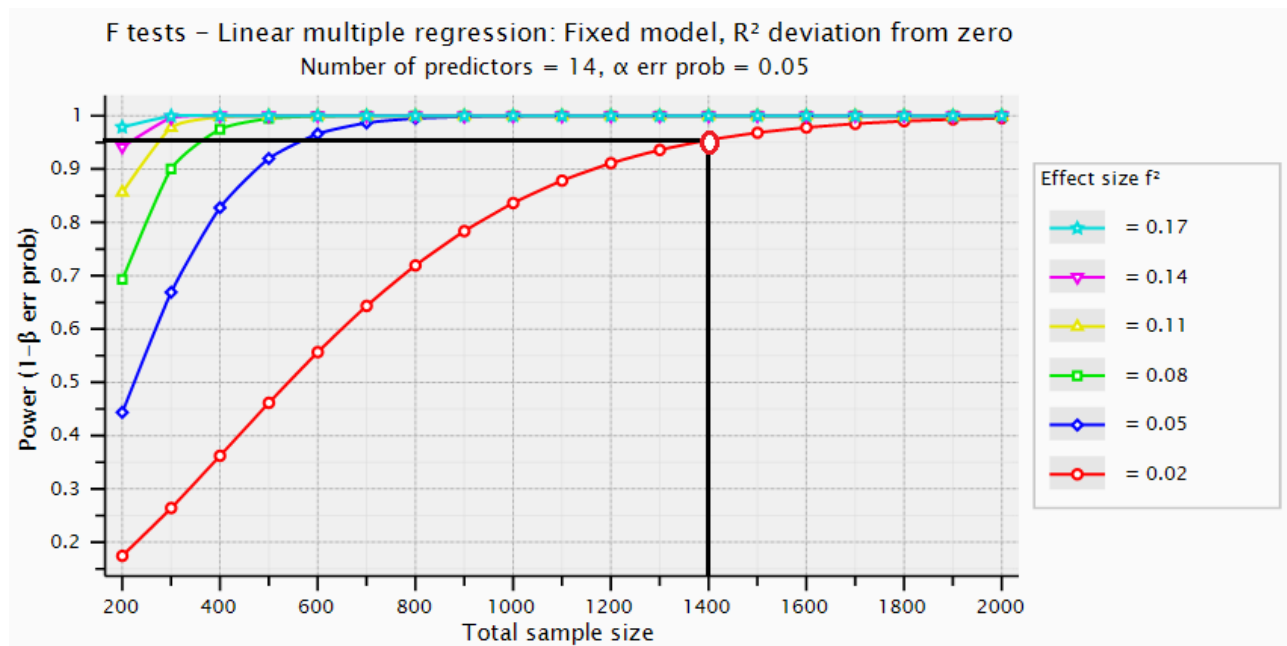


Figura 4.1 Resultados test en software G*Power

4.4. Conclusiones

Se estableció un diseño factorial balanceado seleccionando 10 variables explicativas críticas, las cuales fueron divididas mayoritariamente en dos niveles, con excepción de las variables de longitud, ancho y clima. El procedimiento de cálculo se configuró bajo parámetros estadísticos exigentes para detectar efectos sutiles en un fenómeno complejo.

El análisis de sensibilidad mediante regresión lineal múltiple determinó que, para satisfacer los criterios de potencia estadística con 14 predictores, el tamaño muestral mínimo indispensable para la investigación es de $N = 1.400$ puentes.

CAPÍTULO 5: PROCEDIMIENTO DE CALIBRACIÓN

5.1. Introducción

Este capítulo describe la implementación técnica del diseño experimental, abarcando desde la integración y depuración de bases de datos heterogéneas hasta la formulación matemática de nueve herramientas de regresión, junto a los protocolos de calibración necesarios para asegurar la validez predictiva.

5.2. Procedimiento general

La modelación comienza con la recopilación de datos, obtenidas de diferentes fuentes. Inmediatamente sigue la depuración de datos, proceso que consiste en transformar los datos brutos en un conjunto limpio y estructurado. A partir de esta base depurada, se aplica una reducción dimensional a los datos, realizando un proceso de selección de variables influyentes mediante herramientas de *machine learning*.

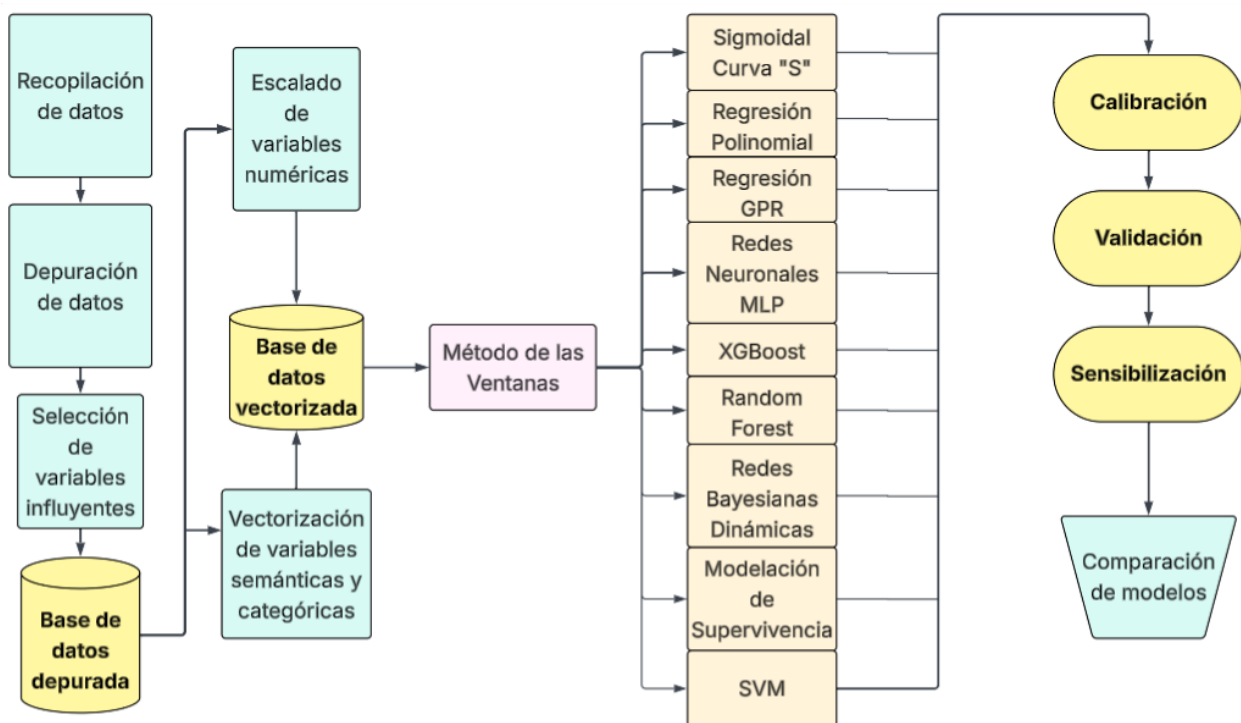


Figura 5.1 Procedimiento general del modelo

Posteriormente, se aplica el escalado de variables numéricas para normalizar sus rangos y la vectorización de variables semánticas y categóricas para convertirlas en formatos numéricos procesables. Estos pasos convergen en una base de datos vectorizada, que sirve como *input* unificado para las ramas de modelación, facilitando la integración de datos heterogéneos en el flujo general del procedimiento.

Luego se realiza un proceso de estratificación de los datos, utilizando herramientas de clustering se crean subgrupos que son internamente homogéneos, a los que llamaremos familias. Dentro de cada familia, es posible cuantificar la contribución real de grupos de variables al deterioro. A continuación de la clusterización, se aplica un proceso de ensamblaje, generando series temporales a partir de una base de datos que solo contiene una inspección por cada puente.

Una vez estructurado los datos y estratificada la red, el siguiente paso es cuantificar la relación entre las características de un puente y su estado de deterioro a lo largo del tiempo, es decir, un proceso de regresión, calibrado mediante herramientas de inteligencia artificial y métodos tradicionales. Este proceso tiene como objetivo desarrollar un modelo predictivo que produce como salida el valor más probable de su calificación de condición (en la escala de 0-100), tomando como entrada un vector de características de un puente.

5.3. Preparación de datos

El objetivo de esta etapa es construir un conjunto de datos unificado e indexado a partir de tres bases de datos diferentes: Una de inventario e inspección de puentes, otra de volúmenes de tráfico y una de zonas climáticas en Chile. El origen de estos datos proviene principalmente del Ministerio de Obras Públicas Chileno (MOP), junto a su Dirección de Vialidad, los cuales tienen 6781 puentes a su cargo. La zonificación climática se obtiene a partir del artículo Sarricolea et al (2017), el cual realiza una regionalización climática de Chile basada en la clasificación Koppen-Geiger. En el Anexo 5.1 se encuentra esta zonificación.

5.3.1 Estructura de base de datos existentes

La base de datos de inventario e inspección obtenida contiene 52 campos, siendo cada entrada un deterioro específico por puente. Se tiene solamente una inspección para cada uno de los puentes del país. Por otro lado, la base de datos de tráfico posee cinco campos, donde cada entrada es un tramo de ruta perteneciente a la red vial chilena. Se seleccionan solo los tramos de ruta donde existan puentes incluidos en el estudio. Finalmente, la base de datos climática posee seis campos, donde cada entrada es un puente, con su respectiva ubicación y zona climática.

5.3.2 Edición de base de datos

Para abordar la depuración de los datos, se implementa un proceso de limpieza en *Python*. Se emplea la librería *Re* para estandarizar formatos de texto mediante expresiones regulares. Posteriormente, se utiliza *FuzzyWuzzy* para consolidar categorías con variaciones tipográficas, como distintos nombres para un mismo tipo de material. Además, se eliminan los campos que no poseen la cantidad de datos suficientes para realizar el análisis.

La integridad estructural del *dataset* final será validada con *Pydantic*, asegurando que cada registro de puente cumpla con un esquema de datos predefinido. Luego de la depuración, las bases de datos son fusionadas, conteniendo datos de materialidad, estructuración, geometría, edad, tráfico y clima, contando con 22 campos, donde cada entrada corresponde a un puente. Los campos se pueden observar en la Tabla 5.1.

Tabla 5.1 Campos de la base de datos depurada

N°	Campos
1	ID
2	Código i3MOP
3	Clase de ruta
4	Edad del puente
5	Latitud
6	Longitud
7	Zona Climática
8	TMDA (veq/h)
9	% Vehículos pesados

10	% Crecimiento vehicular
11	Material Cepas
12	Material Estribos
13	Vigas
14	Losa
15	Carpeta
16	Número de Vigas
17	Número de Cepas
18	Longitud Total
19	Ancho total
20	Espesor Losa
21	Número total de deterioros
22	Calificación automática final

5.3.3 Reducción Dimensional

Con un conjunto de datos limpio y estructurado, la siguiente fase se centra en la reducción dimensional de la base de datos, esto con el fin de eliminar la redundancia y evitar un posible sobreajuste (*overfitting*) durante el entrenamiento y calibración de los modelos. Esta reducción se hará mediante el uso de técnicas de inteligencia artificial adecuadas para medir la influencia relativa de cada variable en la condición de los puentes.

La selección del conjunto de características que se usan para calibrar los modelos de predicción se basa principalmente en la experiencia ingenieril, siendo este criterio subjetivo, llevando a la exclusión de variables que finalmente si aportan directamente al deterioro del puente. El uso de herramientas de IA para la reducción dimensional responde directamente a este problema de la subjetividad, en específico, los algoritmos de *Machine Learning* son herramientas potentes y adecuadas para esta tarea, permitiendo identificar y clasificar de manera objetiva las características más relevantes que influyen en la condición de los puentes. (Zhu y Wang, 2021)

Dentro de las herramientas de *Machine Learning* seleccionadas, se encuentra el algoritmo ReliefF, el cual se basa en la técnica “k-vecinos cercanos”, permitiendo identificar y clasificar de manera objetiva las características más relevantes que influyen en la condición de los puentes. También se seleccionó

el algoritmo *Extreme Gradient Boosting (XGBoost)*, el cual puede utilizarse tanto como técnica de clasificación como también para realizar una regresión supervisada. Este algoritmo se basa en un conjunto de árboles de decisión que se agregan ajustando los residuos o clasificaciones anteriores. Finalmente, también se seleccionó el algoritmo *Random Forest (RF)*. Al igual que *XGBoost*, puede utilizarse tanto para clasificación como para regresión. Esta técnica que utiliza un conjunto de árboles de decisión generados desde distintas muestras del conjunto destinado al entrenamiento, donde se seleccionan los árboles que mayor cantidad de votos reciban entre sí.

Antes de realizar la clasificación, el conjunto de datos se prepara y segmenta para asegurar la validez de los resultados. La base de datos se divide en un conjunto de entrenamiento (80%) y un conjunto de prueba (20%). Para evitar el sobreajuste y garantizar que los modelos generalicen bien a datos no vistos, todo el entrenamiento se realiza utilizando una técnica de validación cruzada de cinco segmentos. Este método divide el conjunto de entrenamiento en cinco partes, entrenando el modelo cinco veces, cada vez utilizando una parte diferente como validación interna y las cuatro restantes para el entrenamiento, garantizando que las herramientas de selección (*XGBoost*, RF y ReliefF) estén correctamente calibradas y que sus rankings de importancia sean estables y fiables.

El proceso se repite 5 veces (una por cada bloque):

- Ronda 1: Se entrena el modelo con los bloques 2, 3, 4 y 5. Se valida con el bloque 1.
- Ronda 2: Se entrena el modelo con los bloques 1, 3, 4 y 5. Se valida con el bloque 2.
- Ronda 3: Se entrena con los bloques 1, 2, 4 y 5. Se valida con el bloque 3.
- Ronda 4: Se entrena con los bloques 1, 2, 3 y 5. Se valida con el bloque 4.
- Ronda 5: Se entrena con los bloques 1, 2, 3 y 4. Se valida con el bloque 5.

El proceso para medir la influencia se realiza en dos etapas para refinar progresivamente los resultados. En la primera etapa, se analizan todas las variables explicativas disponibles para obtener una primera estimación de su importancia.

ReliefF se aplica directamente para calcular un coeficiente de influencia para cada variable. Este coeficiente se basa en la capacidad de la variable para distinguir entre instancias cercanas que

pertenecen a clases diferentes. El valor y rango de este coeficiente dependen del conjunto de datos específico que se está analizando.

XGBoost utiliza métodos de cálculo de importancia para cuantificar la relación entre cada variable y la condición del puente. El algoritmo cuenta el número de veces que una variable se usó para hacer un corte en todos los árboles del modelo y mide la mejora promedio en la precisión que aporta una variable cada vez que se usa para un corte. Durante el entrenamiento, cada algoritmo genera una puntuación para cada característica: *XGBoost* asigna un valor entre 0 y 900, donde se seleccionan las variables con más de 100 puntos.

RF utiliza un método de importancia basada en impurezas, en donde cada vez que una variable se usa para dividir un nodo, se calcula cuánto redujo la impureza. Finalmente, asigna una puntuación de 0 a 0.4, considerándose importantes aquellas con más de 0.2 puntos.

En la segunda etapa, se deben comparar los resultados de los tres modelos, para generar una lista única de variables influyentes. Las puntuaciones generadas por cada modelo son completamente diferentes, por lo que se propone un proceso de normalización, con un rango entre 0 y 100. En el caso de ReliefF, se identifica el valor mínimo y máximo de influencia obtenido, realizando un escalamiento interno.

Este proceso es complementado y validado por el algoritmo Boruta. Su función principal es tomar una decisión binaria: si una variable es estadísticamente relevante (Confirmada) o no (Rechazada). Finalmente, el conjunto de variables influyentes se define como aquellas que son confirmadas por Boruta y que, además, se encuentran entre las más importantes según las puntuaciones promedio de los tres modelos.

5.3.4 Vectorización de variables

El tratamiento de variables semánticas y categóricas es un desafío recurrente en la modelación de datos. Métodos tradicionales como la codificación binaria son incapaces de reflejar la similitud o disimilitud real entre categorías, lo que limita la interpretabilidad del modelo (Koronaki et al., 2024). Para superar esta limitación, se propone un enfoque basado en Modelos de Lenguaje Grandes (LLM) que transforma estas variables en un espacio de características unificado.

Para las variables semánticas y categóricas, se propone una técnica de vectorización avanzada mediante un LLM. A diferencia de métodos tradicionales, esta aproximación permitirá generar *embeddings*, que capturen el significado semántico de las palabras. Las variables numéricas son escaladas para asegurar su correcta interpretación por parte del modelo.

La integración de estos *embeddings* mejora el rendimiento predictivo de los modelos. El resultado final de esta etapa es un vector de atributos único y semánticamente denso para cada uno de los puentes. (Kasneci y Kasneci, 2024)

5.3.5 Organización de base de datos usando método de ensamblaje

El Método de las Ventanas es una técnica de muestreo utilizada para generar series temporales a partir de una base de datos limitada, como en este caso, ya que solo se tiene una inspección por cada puente. Este método se fundamenta en la hipótesis de que los atributos de la red de puentes muestreada pertenecen a la misma población.

El procedimiento consiste en definir una serie de puentes, $B(\theta, t)$, dentro de la red vial, donde cada puente posee un vector de atributos y una edad específica. Dado que los puentes se agrupan previamente en familias con características uniformes y constantes, es posible considerar a cada grupo de puentes, $B_i(\theta_i, t_i)$, como una ventana de un proceso de deterioro estocástico global para esa familia. Posteriormente, estas ventanas se ensamblan para construir un proceso de deterioro único para cada familia, lo que resulta en la creación de varias series temporales agrupadas, como se puede observar en la Figura 5.3. Esto genera una base de datos adecuada y lista para la calibración de los modelos de regresión. (Videla et al, 1996)

Para la generación de series temporales se seleccionó el algoritmo de *clustering* jerárquico basado en densidad *Hierarchical Density-Based Spatial Clustering of Applications with Noise* (HDBSCAN), ya que posee la habilidad intrínseca para clasificar puentes atípicos como ruido, en lugar de forzarlos a una familia, garantizando una segmentación más robusta y realista.

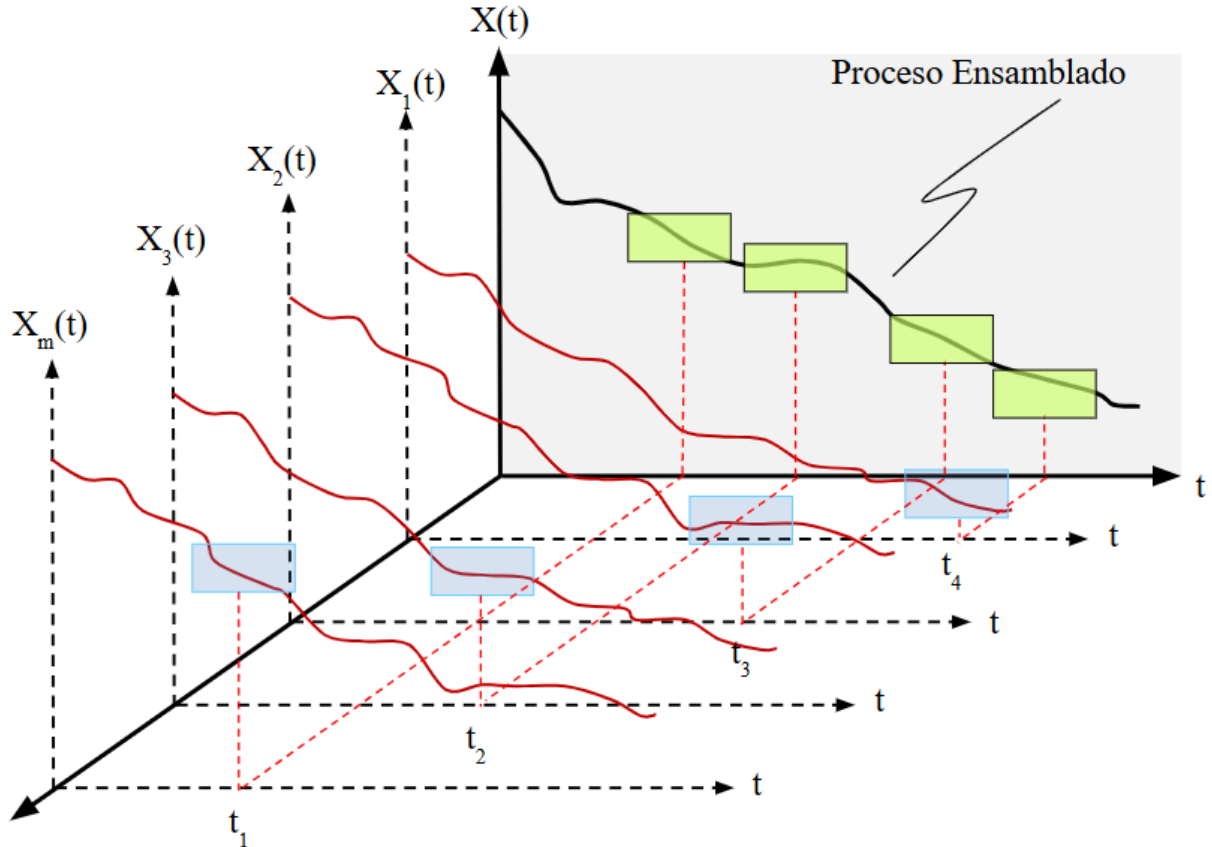


Figura 5.2 Ilustración genérica de ventanas ensambladas (de Solminihac, 2018)

5.4. Herramientas de modelación

La selección de las técnicas para modelar la dependencia funcional del deterioro integra enfoques estadísticos clásicos, probabilísticos y de inteligencia artificial. Dado que el fenómeno de la degradación en puentes exhibe comportamientos tanto deterministas como estocásticos, y considerando la heterogeneidad de la base de datos de la red vial chilena, se ha optado por no restringir el análisis a un único algoritmo.

El propósito es someter estos paradigmas a una evaluación empírica comparativa, buscando identificar la arquitectura que ofrezca el mejor equilibrio entre precisión predictiva y coherencia fenomenológica. En la Tabla 5.2 se resumen los modelos de regresión que se utilizarán para la comparación.

Tabla 5.2 Resumen de los modelos de regresión

ID	Herramienta de modelación	Ecuación
1	Regresión Polinómica	$C(x) = \beta_0 + \beta_1 * (X_1) + \dots + \beta_n * (X_2^m)$
2	Regresión Sigmoidal	$C(t) = \frac{L}{1 + e^{-k(t-t_0)}} + C$
3	Regresión de Procesos Gaussianos (GPR)	<i>Condición</i> $(t) \sim GP(m(t), k(t, t'))$
4	Redes Bayesianas Dinámicas (DBN)	$P(C_{1:T}, X_{1:T}) = P(C_1) \prod_{t=2}^T P(C_t C_{t-1}, X_t)$
5	Modelación de Supervivencia	$\lambda(t x) = \lambda_0(t) * \exp(\beta_1 x_1 + \dots + \beta_p x_p)$
6	Redes Neuronales MLP (FFNN)	$C(t) = W_0 f(W_h x + b_h) + b_0$
7	Extreme Gradient Boosting (XGBoost)	$FO(\theta) = \sum_{i=1}^n (C_i - C_{pr})^2 + \sum_{k=1}^K \Omega(f_k)$
8	Random Forest (RF)	$C(x) = \frac{1}{K} \sum_{k=1}^K f_k(x)$
9	Máquinas de Soporte Vectorial (SVM)	$f(x) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) K(x_i, x) + b$

5.4.1 Regresión polinómica

La regresión polinomial es una extensión de la regresión lineal que permite modelar relaciones no lineales entre variables. En lugar de ajustar una línea recta, se ajusta una ecuación polinómica de un grado n a los datos, donde la condición es la variable dependiente y el vector de características del puente las variables independientes.

La flexibilidad de este modelo reside en el grado del polinomio, ya que un grado mayor permite capturar curvaturas y fluctuaciones más complejas en los datos de deterioro. Sin embargo, se debe

tener cuidado al seleccionar el grado para evitar el sobreajuste, donde el modelo se ajusta demasiado al ruido de los datos de entrenamiento y pierde su capacidad de generalización. La elección del grado óptimo se realiza típicamente mediante técnicas de validación cruzada descritas anteriormente.

$$\text{Condición} = \beta_0 + \beta_1 * (X_1) + \dots + \beta_n * (X_2^m) \quad (4.1)$$

En la Ecuación 4.1 se observa el vector de coeficientes β , el cual capta las relaciones entre variables y el nivel de influencia que tiene sobre la condición del puente. Este método es útil por su capacidad de aproximar una amplia variedad de formas de curvas de deterioro sin asumir un patrón de crecimiento específico de antemano.

5.4.2 Regresión sigmoideal (Curva S)

Este método se basa en el ajuste de los datos a una función sigmoideal, comúnmente representada por la función logística con una curva en forma de S. En el contexto del deterioro de puentes, la curva sigmoideal puede representar el ciclo de vida de una estructura, como se observó en la Figura 2.1. La expresión matemática que modela este comportamiento se encuentra en la Ecuación 4.2, donde L representa la pérdida total de condición, k define la tasa de deterioro y t_0 es el punto de máxima degradación. Dentro del parámetro k, se encuentran las relaciones entre variables y el nivel de influencia que tiene sobre la condición del puente.

$$\text{Condición}(t) = \frac{L}{1 + e^{-k(t-t_0)}} + C \quad (4.2)$$

El ajuste del modelo consiste en encontrar los parámetros de esta función, como el punto de máxima pendiente (t_0), y la asíntota inferior, (C) que minimicen el error con respecto a los datos observados para cada familia de puentes. Su principal ventaja es la alta interpretabilidad de sus parámetros, que se correlacionan directamente con las etapas del proceso de deterioro.

5.4.3 Regresión de Procesos Gaussianos (GPR)

La Regresión de Procesos Gaussianos (GPR) es un enfoque no paramétrico y probabilístico para la regresión. A diferencia de los otros modelos que buscan encontrar los parámetros de una función específica, la GPR define una distribución de probabilidad sobre un conjunto de posibles funciones que podrían ajustarse a los datos. Formalmente, se postula que la función de deterioro sigue un Proceso Gaussiano descrito en la Ecuación 4.3, donde $m(t)$ es la función de media y $k(t, t')$ es un kernel o función de covarianza, el cual define la similitud o cercanía entre dos puntos. El modelo asume que cualquier conjunto de puntos de datos puede ser descrito por una distribución gaussiana (normal) multivariada. Durante el entrenamiento, la GPR utiliza los datos observados para actualizar esta distribución, resultando en una predicción que no es un único valor, sino una media (la curva más probable) y una varianza (un intervalo de confianza).

$$\text{Condición } (t) \sim GP(m(t), k(t, t')) \quad (4.3)$$

La principal fortaleza de la GPR, en este contexto, es su capacidad para cuantificar la incertidumbre. La curva de deterioro generada no solo muestra la condición esperada a una edad determinada, sino que también proporciona una medida de qué tan seguro está el modelo de esa predicción.

5.4.4 Regresión con Redes bayesianas dinámicas

Las Redes Bayesianas Dinámicas (DBN) son modelos gráficos probabilísticos que extienden las redes bayesianas tradicionales para modelar procesos estocásticos que evolucionan con el tiempo. A diferencia de una regresión estática, la DBN actualiza la distribución de probabilidad del estado del puente a medida que avanza el tiempo t , permitiendo inferir la degradación incluso con datos faltantes.

Una DBN se define como la distribución de probabilidad conjunta sobre una secuencia de estados de condición C a lo largo de T pasos de tiempo. Suponiendo que el proceso de deterioro sigue la propiedad de Markov de primer orden (donde el estado futuro depende exclusivamente del estado presente y las variables actuales). La evolución de la condición se modela mediante la Ecuación 4.4.

$$P(C_{1:T}, X_{1:T}) = P(C_1) \prod_{t=2}^T P(C_t | C_{t-1}, X_t) \quad (4.4)$$

Donde:

- $P(C_{1:T}, X_{1:T})$ representa la probabilidad conjunta de toda la trayectoria de deterioro observada y las variables explicativas asociadas.
- $P(C_1)$ es la distribución de probabilidad del estado inicial del puente.
- $P(C_t | C_{t-1}, X_t)$ corresponde al modelo de transición. Este término cuantifica la probabilidad de que el puente evolucione a un estado C dado su estado inmediatamente anterior C_{t-1} modulado por la influencia del vector de variables explicativas X_t en ese intervalo de tiempo.

La estructura del modelo se define mediante un Grafo Acíclico Dirigido (DAG). Esta configuración gráfica establece una jerarquía causal y temporal donde los nodos representan las variables del sistema y las aristas dirigidas indican las dependencias condicionales, permitiendo factorizar la complejidad del problema de deterioro en componentes locales manejables.

5.4.5 Modelación de supervivencia

La modelación de supervivencia se enfoca en predecir el tiempo hasta que suceda un evento, que en este estudio se define como el tiempo que tarda un puente en descender de un nivel de calificación de condición a otro inferior. Este enfoque utiliza la función de riesgo $\lambda(t)$, que describe la tasa instantánea de fallo o deterioro en un tiempo t, dado que el puente ha sobrevivido hasta ese momento. Para modelar la influencia de las variables explicativas sobre esta duración, se emplea comúnmente el Modelo de Riesgos Proporcionales de Cox, descrito en la Ecuación 4.5:

$$\lambda(t|x) = \lambda_0(t) * \exp(\beta_1 x_1 + \dots + \beta_p x_p) \quad (4.5)$$

Donde $\lambda(t|x)$ es el riesgo de deterioro dado el vector de características del puente X, $\lambda_0(t)$ es el riesgo base y β son los coeficientes que indican cómo cada característica del puente acelera o retarda el proceso de deterioro.

5.4.6 Regresión con redes neuronales (MLP)

Las Redes Neuronales de Perceptrón Multicapa (MLP) son modelos de computación inspirados en la estructura del cerebro humano, diseñados para reconocer patrones complejos. Para este problema, se implementará una red neuronal de tipo feedforward, donde el vector de características de cada puente alimenta una capa de entrada. Esta información viaja a través de una o más capas ocultas, donde una serie de transformaciones no lineales, gobernadas por la función de activación (en este caso) *Rectified Linear Unit (ReLU)*, permiten al modelo aprender relaciones jerárquicas y abstractas en los datos. Finalmente, la capa de salida, compuesta por una única neurona con activación lineal, produce la predicción de la calificación de condición. En la Figura 4.3 se muestra la representación gráfica de una red neuronal con 3 capas ocultas.

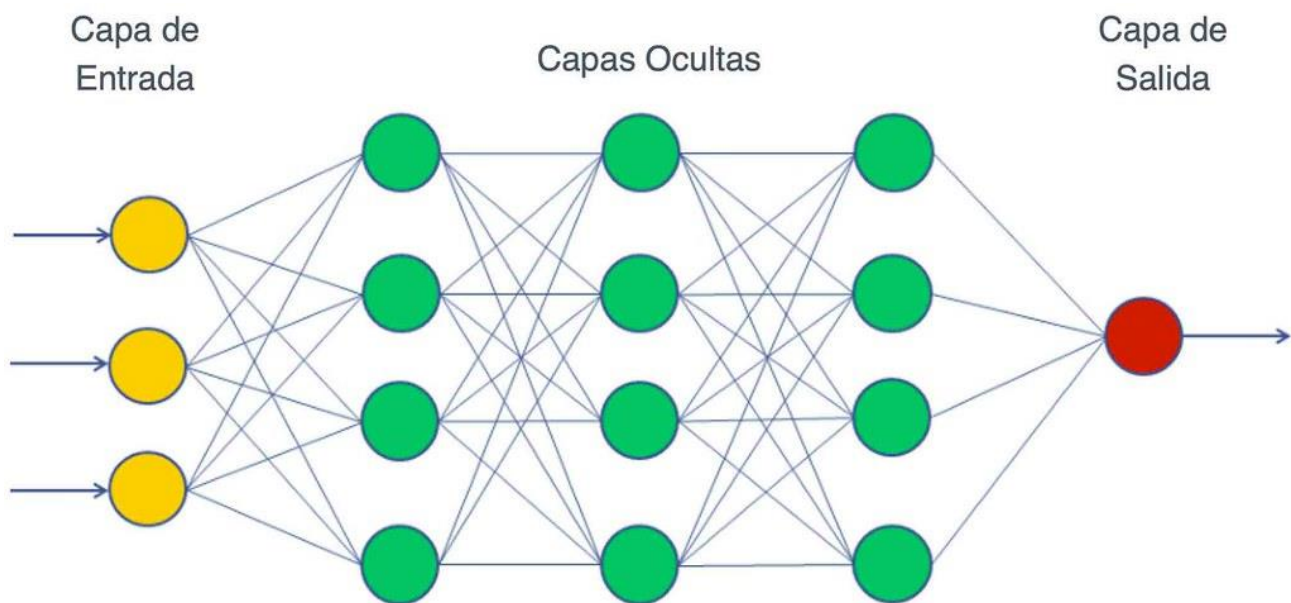


Figura 5.3 Ilustración genérica de una red neuronal artificial

Ahora, considerando una red con una capa, la fórmula que relaciona el vector de características de entrada x con la predicción final del Índice de Condición se encuentra en la Ecuación 4.4, donde W y b son la matriz de pesos y el vector de sesgo, de la capa oculta (h) y la capa de salida (0). Finalmente f , es la función de activación ReLU de la capa oculta.

$$\text{Condición}(t) = W_0 f(W_h x + b_h) + b_0 \quad (4.6)$$

El entrenamiento se realiza mediante el algoritmo de retropropagación, ajustando los pesos de las conexiones para minimizar el Error Cuadrático Medio (MSE) entre las predicciones y los valores reales. La principal fortaleza de este enfoque es su capacidad para modelar relaciones de alta complejidad y no linealidad, aprovechando al máximo la riqueza semántica de los vectores de características generados por los LLM.

5.4.7 Regresión con *XGBoost*

Extreme Gradient Boosting (XGBoost) es un algoritmo de ensemble basado en la técnica de *boosting*, reconocido por su alta eficiencia y precisión predictiva. El modelo funciona de manera secuencial: construye una serie de árboles de decisión donde cada nuevo árbol se entrena para corregir los errores residuales cometidos por el conjunto de árboles anteriores. A diferencia de otros métodos de *boosting*, este modelo integra una regularización en su función objetivo, lo que ayuda a controlar la complejidad de los árboles y a prevenir eficazmente el sobreajuste. El entrenamiento se guía por la optimización de una función objetivo que incluye un término de regularización para evitar el sobreajuste.

$$FO(\theta) = \sum_{i=1}^n (\text{Condición}_i - \text{Condición}_{\text{predi}})^2 + \sum_{k=1}^K \Omega(f_k) \quad (4.7)$$

En la Ecuación 4.5 se observa la función objetivo de *XGBoost*, donde existe un error cuadrático entre la predicción y la condición real. En lugar de enfocarse únicamente en minimizar el error de predicción, *XGBoost* también se asegura de que los árboles de decisión que construye no sean excesivamente complejos, mediante el término de regularización $\sum_{k=1}^K \Omega(f_k)$. Un modelo demasiado complejo puede memorizar los datos de entrenamiento, pero fallará al predecir datos nuevos. Además, este modelo está optimizado para el rendimiento computacional mediante técnicas como el procesamiento en paralelo y el manejo eficiente de valores nulos.

5.4.8 Regresión con *Random Forest*

El modelo Random Forest (Bosque Aleatorio) es otro potente método de *ensemble*, pero a diferencia de XGBoost, se basa en la técnica de *bagging*. Este algoritmo construye una gran cantidad de árboles de decisión independientes durante la fase de entrenamiento. Cada árbol se entrena sobre una submuestra aleatoria de los datos (muestreo con reemplazo) y, en cada nodo del árbol, solo se considera una submuestra aleatoria de las características para decidir la división. Para una tarea de regresión, la predicción final del *Random Forest* no es más que el promedio de las predicciones de todos los árboles individuales. La Ecuación 4.6 refleja este proceso, donde K es el total de árboles del bosque y f_k es la predicción de cada árbol individual.

$$\text{Condición} = \frac{1}{K} \sum_{k=1}^K f_k(x) \quad (4.8)$$

Este proceso de promediado reduce drásticamente la varianza y hace que el modelo sea resistente al sobreajuste. Su simplicidad conceptual, combinada con su robustez y buen rendimiento sin una afinación exhaustiva de hiperparámetros, lo convierte en una excelente opción para establecer una línea base de alto rendimiento en la predicción del deterioro.

5.4.9 Regresión con Máquinas de Vectores de Soporte (SVR)

Las Máquinas de Vectores de Soporte (SVM) para regresión, conocidas como SVR (Support Vector Regression), son algoritmos de aprendizaje supervisado diseñadas para predecir variables continuas. Su objetivo es encontrar una función $f(x)$ que se desvíe a lo más un valor ε de las calificaciones de condición reales y_i . Este enfoque introduce el concepto de la banda de insensibilidad, donde los errores de predicción que caen dentro de este margen de tolerancia ε son ignorados, lo que otorga al modelo una gran robustez frente a la variabilidad inherente en las inspecciones visuales.

Dado que el deterioro de los puentes es un proceso no lineal, la SVR supera las limitaciones de los modelos lineales mediante el uso de funciones núcleo (kernels). Esta técnica permite realizar un mapeo implícito de los vectores de características originales hacia un espacio dimensional superior

(Espacio de Hilbert). En este nuevo espacio transformado, la relación de deterioro, que originalmente era compleja y no lineal, puede ser aproximada eficientemente mediante un hiperplano lineal.

$$f(x) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (4.9)$$

Donde representa $K(x_i, x)$ la función núcleo, donde es seleccionado el Kernel de Base Radial (RBF) por su capacidad de adaptación local, y α_i son los coeficientes de los vectores de soporte que definen la topología de la curva de degradación.

5.5. Calibración, validación y sensibilización del modelo

5.5.1 Re-muestreo de datos para calibración

Para garantizar la capacidad de generalización de los modelos y evitar el *overfitting*, se implementa un re-muestreo de los datos. El conjunto de datos depurado y vectorizado es dividido de forma aleatoria en dos subconjuntos, uno destinado al entrenamiento, el cual tendrá el 80% de los datos. El 20% restante es reservado como conjunto de prueba para la evaluación final. Durante la fase de entrenamiento, se aplica la técnica de validación cruzada de 5 iteraciones. Esto implica que cada modelo es entrenado y evaluado cinco veces en distintas particiones de los datos de entrenamiento, utilizando las métricas de desempeño correspondientes a cada herramienta.

5.5.2 Procedimiento de calibración y validación

Las métricas de desempeño que se utilizarán para validar cada uno de los modelos se presenta a continuación.

- **Regresión Polinomial:** La calibración se centra en determinar el grado óptimo n del polinomio (Ecuación 4.1). Se iteran polinomios de grados 2 a n , seleccionando aquel que minimice el error en los pliegues de validación sin aumentar desproporcionadamente la varianza, penalizando la complejidad innecesaria para mantener la robustez en proyecciones a largo

plazo. En paralelo se utiliza la raíz del error cuadrático medio (RMSE), cuantificando la desviación promedio entre la condición predicha y la real en las mismas unidades de la variable dependiente. También se utiliza el coeficiente de determinación ajustado o R cuadrado. Esta métrica penaliza la inclusión de términos polinómicos adicionales que no aporten una mejora significativa a la explicación de la varianza de los datos, esto para evitar seleccionar polinomios de alto grado que, aunque reduzcan el error de entrenamiento, carecen de capacidad de generalización.

- **Regresión Sigmoidal (Curva S):** Para este modelo no lineal (Ecuación 4.2), la calibración se enfoca en la convergencia de los algoritmos de optimización (como Levenberg-Marquardt) para estimar los parámetros de forma k (tasa de deterioro), t_0 (punto de inflexión) y la asíntota C . Se analizan los intervalos de confianza al 95% de estos parámetros. Si estos intervalos son excesivamente amplios, el modelo se considera inestable fenomenológicamente. También se utiliza el R cuadrado, junto al RMSE.
- **Regresión de Procesos Gaussianos (GPR):** La calibración del GPR se fundamenta en la inferencia bayesiana. El objetivo es optimizar los hiperparámetros del *kernel*, que mejor expliquen la covarianza de los datos observados. Para la validación se utiliza la verosimilitud marginal logarítmica (Log-Marginal Likelihood), optimizando los hiperparámetros y controlando la suavidad de la curva de deterioro (Ecuación 4.3). Además, se utiliza la Probabilidad de Cobertura (Coverage Probability), verificando empíricamente que porcentaje de los datos de prueba reales caen dentro del intervalo de confianza predicho. También se utilizan las métricas estándar RMSE y R cuadrado, evaluadas sobre la curva central predicha, con el fin de comparar su capacidad de pronóstico con los modelos tradicionales.
- **Redes Bayesianas Dinámicas (DBN):** La calibración y validación de las Redes Bayesianas Dinámicas se enfoca en la calidad de la inferencia probabilística y evaluación de la coherencia lógica del grafo. Dado que la estructura del Grafo Acíclico Dirigido (DAG) se define en base a las relaciones causales teóricas, la calibración se centra en el aprendizaje de parámetros bayesiano. El log-verosimilitud se utiliza como la métrica principal de ajuste, midiendo la probabilidad logarítmica que el modelo asigna a los datos de validación. También se utiliza el RMSE y el R cuadrado para establecer una comparación directa con demás modelos.

- **Modelación de supervivencia:** A diferencia de los modelos de regresión que predicen un valor puntual de condición, el modelo de Riesgos Proporcionales de Cox estima la probabilidad de que un puente mantenga su estado a lo largo del tiempo. Por lo tanto, su calibración y validación requieren métricas especializadas. La calibración de los coeficientes β se realiza mediante la maximización de la Verosimilitud Parcial de Cox. Como métricas de desempeño se utiliza el índice de concordancia de Harrell, el cual es la métrica estándar para modelos de supervivencia, evaluando la capacidad del modelo para ordenar correctamente los tiempos de predicción. Además, se utiliza el *Integrated Brier Score* (IBS), donde se mide la distancia cuadrática media entre la probabilidad predicha y el estado real (sobrevivir o fallar), penalizando la incertidumbre.
- **Redes Neuronales (MLP):** La calibración de la Red Neuronal Multicapa (MLP) presenta desafíos únicos debido a su naturaleza de "caja negra" y su alta capacidad para modelar no linealidades, lo que conlleva un riesgo inherente de *overfitting*, dado el tamaño muestral disponible. El proceso de validación ajusta la arquitectura de la red (Ecuación 4.4), variando el número de capas ocultas y la cantidad de neuronas por capa. Adicionalmente, se calibran la tasa de aprendizaje y el número de épocas, utilizando el conjunto de validación para detener el entrenamiento cuando el Error Cuadrático Medio (MSE) deja de disminuir, preveniendo así la memorización del ruido. A diferencia de las regresiones paramétricas donde los coeficientes β son interpretables, la MLP requiere técnicas de Inteligencia Artificial Explicable. Se emplean Gráficos de Dependencia Parcial y valores SHAP para validar que la respuesta de la red ante variaciones en las entradas sea físicamente coherente y no presente oscilaciones producto del sobreajuste. Finalmente, el R cuadrado se utiliza con el fin de comparar la capacidad de predicción con los demás modelos.
- **XGBoost:** La calibración de esta herramienta de *boosting* se realiza optimizando la función objetivo regularizada. Mediante una búsqueda en cuadrícula, se ajustan hiperparámetros críticos como la profundidad máxima de los árboles, la tasa de aprendizaje y, fundamentalmente, los términos de regularización Ω . La validación busca el equilibrio exacto que maximice la precisión predictiva manteniendo la complejidad de la modelo controlada. Se evalúan las métricas estándar de RMSE y R cuadrado tanto en el conjunto de entrenamiento

como en el de prueba. Dado que XGBoost es un modelo de "caja negra", se valida su coherencia fenomenológica mediante SHAP, esto con el fin de verificar la dirección del impacto, validando que valores altos generen consistentemente contribuciones negativas al puntaje de condición predicho.

- **Random Forest:** Para el modelo de ensamble por *bagging*, el objetivo de la calibración no es solo ajustar cada árbol individual, sino garantizar que los árboles sean lo suficientemente distintos entre sí, para que su promedio reduzca efectivamente la varianza. La validación se centra en determinar el número óptimo de árboles K y la cantidad máxima de características a considerar en cada división de nodo. Se evalúa la estabilidad del error "fuera de la bolsa" (*out-of-bag error*) para asegurar que el promedio de las predicciones reduzca efectivamente la varianza sin sesgar el resultado. Se utilizan las métricas estándar de RMSE y R cuadrado tanto en el conjunto de entrenamiento como en el de prueba, con el fin de realizar la comparación con los demás modelos.
- **Máquinas de Vectores de Soporte (SVM):** La validación de la regresión SVR requiere la selección del kernel adecuado (radial, polinómico o lineal) y el ajuste simultáneo de los parámetros de penalización C y el margen de tolerancia ϵ . Se busca la combinación que minimice el error de predicción dentro del margen de tolerancia, asegurando un modelo lo suficientemente generalista para manejar datos atípicos. Se monitorea el porcentaje de datos de entrenamiento que se convierten en Vectores de Soporte. Si la ratio es cercana al 100%, el modelo ha degenerado (está memorizando todo el dataset) y ha perdido su capacidad de generalización. Se utilizan las métricas estándar de RMSE y R cuadrado con el fin de realizar la comparación con los demás modelos.

5.5.3 Procedimiento de Sensibilización

Más allá de las métricas de rendimiento estándar sobre el conjunto de prueba, esta fase se enfoca en validar la robustez y coherencia fenomenológica del modelo con mejor desempeño. Para ello, se realizará un análisis de sensibilidad sistemático empleando un Diseño Factorial Particular. Se seleccionarán las tres a cinco variables de mayor influencia confirmadas en la sección 5.3.3 y se definirán niveles lógicos para cada una, creando una matriz de escenarios virtuales. Estos vectores de

características sintéticos se introducirán en el modelo ya entrenado para generar las predicciones de deterioro correspondientes, con el objetivo de descubrir y cuantificar relaciones complejas que el modelo ha aprendido. Este proceso no solo valida que el comportamiento del modelo es consistente con los principios de la ingeniería, sino que también proporciona un entendimiento más profundo de las dinámicas de deterioro capturadas.

5.6. Conclusiones

En este capítulo se estableció un flujo de trabajo automatizado en Python para la depuración de datos, integrando registros de inventario, tráfico y clima en una base unificada de 22 campos, validada mediante esquemas de integridad estructural. Además, se determinó el uso de *embeddings* generados por Modelos de Lenguaje Grandes (LLM) para la vectorización de variables categóricas y semánticas, permitiendo capturar relaciones contextuales que los métodos de codificación tradicionales no logran representar.

Se implementó un mecanismo de reducción dimensional libre de subjetividad mediante la triangulación de los algoritmos ReliefF, XGBoost y Random Forest, utilizando una validación final con el algoritmo Boruta para confirmar estadísticamente las variables influyentes. Esto junto a una aplicación del Método de las Ventanas, para ensamblar series temporales a partir de datos transversales, superando la limitación de contar con una única inspección por puente.

Se formalizaron nueve modelos de regresión distintos que abarcan enfoques polinómicos, probabilísticos (GPR, DBN), de supervivencia y de inteligencia artificial (MLP, XGBoost, SVR), proporcionando ecuaciones específicas para abordar la no linealidad y la incertidumbre del deterioro. Además, Se definió un esquema de calibración riguroso mediante re-muestreo aleatorio (80% entrenamiento, 20% prueba) y validación cruzada de cinco iteraciones, complementado con un análisis de sensibilidad factorial para verificar la coherencia fenomenológica de los resultados frente a métricas como RMSE y R cuadrado.

CAPÍTULO 6: CONCLUSIONES

El propósito de esta memoria fue desarrollar un diseño experimental para la predicción del deterioro en puentes convencionales tipo viga. A través de una metodología estructurada, se logró integrar la teoría clásica de gestión de puentes con herramientas avanzadas de ciencia de datos, abordando la complejidad multifactorial del fenómeno de deterioro.

El comportamiento de los puentes depende de la interacción conjunta de factores de inventario, materialidad, estructuración, geometría y cargas cíclicas (tráfico y clima), definiendo el vector de características inicial con diez variables independientes específicas. La selección de estas variables permitió estructurar un modelo coherente con los datos disponibles en la Dirección de Vialidad.

Si bien los modelos mecanicistas ofrecen una explicación causal precisa basada en leyes fisicoquímicas (como la Ley de Fick o de Paris), su aplicabilidad en Sistemas de Gestión de Puentes (SGP) a nivel de red es limitada debido a la inviabilidad técnica y económica de obtener parámetros experimentales para cada estructura.

Los modelos empíricos se consolidan como el estándar para la gestión de redes masivas, al aprovechar los datos históricos de inspección visual. Sin embargo, se evidencia una obsolescencia de los enfoques deterministas simples y de las Cadenas de Markov tradicionales, debido a su incapacidad para modelar la dependencia temporal a partir de datos tabulares limitados. La frontera del conocimiento actual se sitúa en los modelos híbridos y de Aprendizaje Profundo (*Deep Learning*), los cuales han demostrado una capacidad superior para procesar datos heterogéneos y capturar patrones latentes complejos que escapan a las regresiones estadísticas convencionales. Esta tendencia valida la elección de las herramientas de IA propuestas en este diseño experimental.

También, se determinó el tamaño muestral necesario para asegurar la validez estadística del diseño experimental. Dado que las 10 variables formuladas generan una matriz factorial compleja, el análisis de potencia realizado indicó que se requiere un mínimo de 1.400 puentes. Esta determinación garantiza que el modelo tendrá una potencia del 95% para detectar efectos pequeños, verificando que la base de datos disponible posea la densidad suficiente para capturar la varianza del deterioro.

La implementación del Método de las Ventanas junto al algoritmo de clustering HDBSCAN demostró ser una solución técnica viable para superar la limitación de contar con una única inspección por puente, permitiendo la generación de series temporales, necesarias para el entrenamiento de modelos de regresión.

Este procedimiento permitió la configuración de un ecosistema de nueve arquitecturas de regresión, donde se contrastan sistemáticamente modelos de aprendizaje profundo con métodos probabilísticos y deterministas. El testeado de este espectro de modelos asegura que la capacidad predictiva del sistema no dependa de un sesgo algorítmico particular, permitiendo identificar la estructura matemática que mejor asimila la fenomenología de la degradación de la condición.

Este trabajo establece una base para construir modelos de puentes convencionales tipo viga. Sin embargo, es fundamental el refinamiento de los datos y las técnicas de modelación.

GLOSARIO

SGP	=	Sistema de gestión de puentes
BCI	=	Bridge Condition Index
NBI	=	National Bridge Inventory
ANOVA	=	Análisis de Varianza
LS-SVR	=	Least Squares Support Vector Regression
GA	=	Genetic Algorithm
CNN	=	Convolutional Neural Network
RNN	=	Recurrent Neural Network
RFIS	=	Recurrent Fuzzy Inference System
TMDA	=	Tráfico Medio Diario Anual
KGC	=	Koppen-Geiger Classification
MOP	=	Ministerio de Obras Públicas Chileno
DBN	=	Dynamic Bayesian Network
RMSE	=	Root Mean Square Error
RF	=	Random Forest
XGBoost	=	Extreme Gradient Boosting
LLM	=	Large Language Model
FFNN	=	Feed Forward Neural Network
SVM	=	Support Vector Machine
HDBSCAN	=	Hierarchical Density-Based Spatial Clustering of Applications with Noise
GPR	=	Gaussian Process Regression
IBS	=	Integrated Brier Score
MLP	=	Multilayer Perceptron
SHAP	=	Shapley Additive Explanations

REFERENCIAS

- AASHTO (1993) **Guidelines for bridge management systems**. American Association of State Highway and Transportation Officials. Washington D.C. Estados Unidos.
- AASHTO (2020). **AASHTO LRFD Bridge Design Specifications (9th Edition)**. Washington, DC.
- Ann, K.Y., Pack, S.-W., Hwang, J.-P., Song, H.-W., & Kim, S.-H. (2010). Service life prediction of a concrete bridge structure subjected to carbonation. **Construction and Building Materials**, **24**, 1494-1501.
- Austroroads (2004) **Austroroads Guidelines for Bridge Management: Structure Information**. Association of Australian and New Zealand road transport and traffic authorities. Sydney. Australia.
- Banco Mundial. (2024). **Transporte: Panorama general**. Sitio web del Banco Mundial.
- Bolukbasi, M., Mohammadi, J., & Arditi, D. (2004). Estimating the Future Condition of Highway Bridge Components Using National Bridge Inventory Data. **Practice Periodical on Structural Design and Construction**, **9(1)**, 16-25.
- Cavalline, T. L., Whelan, M. J., Tempest, B. Q., Goyal, R., & Ramsey, J. D. (2015). **Determination of bridge deterioration models and bridge user costs for the NCDOT bridge management system**. Report No. FHWA/NC/2014-07. North Carolina Department of Transportation. USA.
- Cesare, M. A., Santamarina, C., Turkstra, C., & Vanmarcke, E. H. (1992). Modeling bridge deterioration with Markov chains. **Journal of Transportation Engineering**, **118(6)**, 820-833.
- Cheung, M.S., & Li, W.C. (2003). Probabilistic fatigue and fracture analyses of steel bridges. **Structural Safety**, **23**, 245-26

- De Dieu, G., Chingui, J., & Botir, E. (2020). A Review of Bridge Management System: Deterioration Models. **Asian Journal of Engineering and Technology**, **8(2)**, 38-47.
- de Solminihac, H., Echaveguren, T., Chamorro, A. (2018) **Gestión de Infraestructura Vial**. Tercera edición. Ediciones Universidad Católica de Chile. Santiago.
- Deng, L., & Cai, C. S. (2010). Bridge Scour: Prediction, Modeling, Monitoring, and Countermeasures-Review. **Practice Periodical on Structural Design and Construction**, **15(2)**, 125-134
- Djoković, J. M., Nikolić, R. R., Bujnać, J., & Hadzima, B. (2018). Estimate of the steel bridges fatigue life by application of the fracture mechanics. **IOP Conference Series: Materials Science and Engineering**, **419(1)**, Artículo 012010.
- Frangopol, D. M., & Soliman, M. (2016). *Life-cycle of structural systems: Recent achievements and future directions*. **Structure and Infrastructure Engineering**, **12(1)**, 1-20.
- Ghonima, O., Anderson, J. C., Schumacher, T., & Unnikrishnan, A. (2020). Performance of US Concrete Highway Bridge Decks Characterized by Random Parameters Binary Logistic Regression. **ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering**, **6(1)**.
- Hasan, S., & Elwakil, E. (2019). Stochastic regression deterioration models for superstructure of prestressed concrete bridges in California. **Journal of Structural Integrity and Maintenance**, **4(2)**, 97-108.
- Iacovino, C., Turksezer, Z. I., Giordano, P. F., & Limongelli, M. P. (2022). A survey of bridge condition rating systems. **Proceedings of the 1st Conference of the European Association on Quality Control of Bridges and Structures. EUROSTRUCT 2021** (pp. 13–20).
- Janas, L., Miller, B., & Kaszyński, A. (2018). Computational algorithms supporting the bridge management system. **The Baltic Journal of Road and Bridge Engineering**, **13(4)**, 357-373.

- Jeong, Y., Kim, W., Lee, I., & Lee, J. (2018). Bridge inspection practices and bridge management programs in China, Japan, Korea, and U.S. **Journal of Structural Integrity and Maintenance**, *3*(2), 126-135.
- Kasneci, G., & Kasneci, E. (2024). **Enriching tabular data with contextual LLM embeddings: A comprehensive ablation study for ensemble classifiers** [Preprint]. arXiv.
- Kassem, A., Salaheldin, T. M., Imran, J., & Chaudhry, M. H. (2003). Numerical modeling of scour in cohesive soils around artificial rock island of Cooper River Bridge. **Transportation Research Record: Journal of the Transportation Research Board**, *1851*(1), 45–50.
- Khitab, A., Arshad, M. T., Ali, S. A., Kazmi, S. M. S., & Munir, M. J. (2014). Modeling of chloride ingress in concrete using Fick's laws: Review and historical perspective. **Science International (Lahore)**, *26*(4), 1519-1521.
- Koronaki, E. D., Loachamin Suntaxi, G., Papavasileiou, P., Giovanis, D. G., Kathrein, M., Boudouvis, A. G., & Bordas, S. P. A. (2024). **Implementing LLMs in industrial process modeling: Addressing categorical variables** [Preprint]. arXiv.
- Liang, M.-T., & Lan, J.-J. (2005). Reliability analysis for the existing reinforced concrete pile corrosion of bridge substructure. **Cement and Concrete Research**, *35*(3), 540-550.
- Liang, M.-T., Huang, R., & Fang, S.-A. (2013). Carbonation service life prediction of existing concrete viaduct/bridge using time-dependent reliability analysis. **Journal of Marine Science and Technology**, *21*(1), 94-104.
- Liu, H., Nehme, J., & Lu, P. (2021). An application of convolutional neural network for deterioration modeling of highway bridge components in the United States. **Structure and Infrastructure Engineering**, *19*(6), 731-744.

- Liu, K., & El-Gohary, N. (2022). Deep Learning-Based Analytics of Multisource Heterogeneous Bridge Data for Enhanced Data-Driven Bridge Deterioration Prediction. **Journal of Computing in Civil Engineering**, **36(5)**, 04022023.
- Lu, P., Pei, S., & Tolliver, D. (2016). Regression Model Evaluation for Highway Bridge Component Deterioration Using National Bridge Inventory Data. **Journal of the Transportation Research Forum**, **55(1)**, 5-16.
- Miao, P., Xing, G., Ma, S., & Srimahachota, T. (2023). Deep Learning-Based Inspection Data Mining and Derived Information Fusion for Enhanced Bridge Deterioration Assessment. **Journal of Bridge Engineering**, **28(8)**, 04023048.
- Moghtadernejad, S., Lounis, Z., & Zhang, J. (2024). Bridge deterioration modeling considering the climate change effects. En J. Jensen, D. M. Frangopol, & F. Schmidt (Eds.), **Bridge Maintenance, Safety, Management, Digitalization and Sustainability** (pp. 1487-1495).
- Moomen, M., Qiao, Y., Agbelie, B. R., Labi, S., & Sinha, K. C. (2016). **Bridge deterioration models to support Indiana's bridge management system**. Reporte N.º FHWA/IN/JTRP-2016/03. Purdue University. North Carolina, USA.
- Moomen, M., & Siddiqui, C. (2022). Probabilistic deterioration modeling of bridge component condition with random effects. **Journal of Structural Integrity and Maintenance**, **7(3)**, 151-160.
- Rashidi-Nasab, A., & Elzarka, H. (2023). Optimizing Machine Learning Algorithms for Improving Prediction of Bridge Deck Deterioration: A Case Study of Ohio Bridges. **Buildings**, **13(6)**, 1517.
- Nowak, A. S., & Tharmabala, T. (1988). Bridge reliability evaluation using load tests. **Journal of Structural Engineering**, **114(10)**, 2268-2279.

- Sanaati, B. (2024). Performance Evaluation of LS-SVR Model in Predicting Scour Depth in Bridge Piers. **Environment and Water Engineering**, **10(1)**, 94-108.
- Sarricolea, P., Herrera-Ossandon, M., & Meseguer-Ruiz, O. (2017). Climatic regionalisation of continental Chile. **International Journal of Climatology**, **37**, 4555–4570.
- Sobanjo, J. O. (2011). State transition probabilities in bridge deterioration based on Weibull sojourn times. **Structure and Infrastructure Engineering**, **7(10)**, 747-764.
- Thompson, P.D., E. Small. M. Jhonson y A. Marshall (1998) The PONTIS Bridge Management System. **Structural Engineering International Journal**. **8(4)**. 303-308.
- Videla, C., De Solminihac, H., Gaete, R. y Bustos, M. (1996) **Ajuste de Factores de Calibración para ampliar modelos de deterioro en pavimentos asfálticos**. DGOP-190-08-03-1994. Santiago. Chile.
- Wang, X., & Foliente, G. (2008). Identifying bridge structural condition development trends via categorical inspection condition rating with case studies. **Structures & Infrastructure Engineering**, **4(6)**, 449-466.
- Woodward, R. (2001) **Bridge Management in Europe**. Final report. Inglaterra.
- Xu, G., & Azhari, F. (2022). Data-driven optimization of repair schemes and inspection intervals for highway bridges. **Reliability Engineering and System Safety**, **228**, 108779.
- Yosri, A., Elleathy, Y., Hassini, S., & El-Dakhakhni, W. (2021). Genetic Algorithm-Markovian Model for Predictive Bridge Asset Management. **Journal of Bridge Engineering**, **26(8)**.
- Zhu, J., & Wang, Y. (2021). Feature Selection and Deep Learning for Deterioration Prediction of the Bridges. **Journal of Performance of Constructed Facilities**, **35(6)**, 04021078.

ANEXOS

ANEXO 1.1 Contribución a los Objetivos de Desarrollo Sostenible

Los Objetivos de Desarrollo Sostenible (ODS) constituyen un llamamiento universal a la acción para poner fin a la pobreza, proteger el planeta y mejorar las vidas y las perspectivas de las personas en todo el mundo. Seleccione a cuál o cuáles de los 17 ODS contribuye su trabajo de Memoria de Título:

- ODS-1 : Fin de la pobreza.
- ODS-2 : Hambre cero.
- ODS-3 : Salud y bienestar.
- ODS-4 : Educación de calidad.
- ODS-5 : Igualdad de género.
- ODS-6 : Agua limpia y saneamiento.
- ODS-7 : Energía asequible y no contaminante.
- ODS-8 : Trabajo decente y crecimiento económico.
- ODS-9 : Industria, innovación e infraestructura.
- ODS-10 : Reducción de las desigualdades.
- ODS-11 : Ciudades y comunidades sostenibles.
- ODS-12 : Producción y consumo responsables.
- ODS-13 : Acción por el clima.
- ODS-14 : Vida Submarina.
- ODS-15 : Vida de ecosistemas terrestres.
- ODS-16 : Paz, justicia e instituciones sólidas.
- ODS-17 : Alianzas para lograr los objetivos.

Vinculación
La memoria de título tiene por objetivo desarrollar un diseño experimental para la predicción del deterioro en puentes convencionales tipo viga, integrando la teoría clásica de ingeniería con herramientas avanzadas de ciencia de datos para modelar la complejidad del deterioro estructural. El formular una metodología validada estadísticamente, que asegura la representatividad de los datos y reduce la subjetividad mediante algoritmos de inteligencia artificial, permite modernizar la gestión de activos y proporciona a la Dirección de Vialidad las bases técnicas necesarias para transitar desde acciones reactivas hacia estrategias de mantenimiento predictivo eficientes.

ANEXO 3.1 Contenido de base de datos de inspección de puentes

Tabla A. 1 Campos base de datos de inspección visual de puentes

ID	Contenido	Descripción
1	Fecha inspección principal	Fecha en formato de fecha de Excel.
2	Calificación automática	Numérico. Calificación automática del puente.
3	Calificación final	Numérico. Calificación final del puente (distinta de la automática dado que puede haber intervenciones de los inspectores).
4	Número total de deterioros	Numérico.
5	Índice Estructural	Numérico. Índice interno del sistema I3MOP de cálculo desconocido.
6	Índice Seguridad Vial	Numérico. Índice interno del sistema I3MOP de cálculo desconocido.
7	Índice Contención Sistema de Contención	Numérico. Índice interno del sistema I3MOP de cálculo desconocido.
8	Índice Contención Barandilla	Numérico. Índice interno del sistema I3MOP de cálculo desconocido.
9	Índice Durabilidad	Numérico. Índice interno del sistema I3MOP de cálculo desconocido.
10	Índice Daños Menores	Numérico. Índice interno del sistema I3MOP de cálculo desconocido.
11	Deterioro	*Anexo
12	Causa	*Anexo
13	Gravedad	*Anexo
14	Extensión	*Anexo
15	Medición	Numérico. Medición de deterioro asociado a una unidad de medida.
16	Unidad medición	Unidad de medida de la medición del deterioro: cm, cm2, ml, m2, m3, Ud.
17	Calificación automática del deterioro	Numérico. Calificación específica del deterioro en cuanto a su tipo, causa, extensión, gravedad y unidad donde afecta.
18	Unidad	
19	Código	Alfanumérico. Código asociado a la unidad donde está asociado el deterioro (Ej: E2 1, se refiere a: estribo 2 / primer elemento).

ANEXO 3.2 Contenido bases de datos de inventario de puentes

Tabla A. 2 Campos base de datos de inventario de puentes

ID	Nombre de Variable	Descripción
1	Nombre Puento	Nombre del puente en texto.
2	Código Puento	Código alfa-numérico. Código de la ruta del puente.
3	Código I3MOP	Código alfa-numérico. Código del puente.
4	Región	Texto. Nombre de la Región donde se encuentra el puente.
5	Provincia	Texto. Nombre de la Provincia donde se encuentra el puente.
6	Comuna	Texto. Nombre de la Comuna donde se encuentra el puente.
7	Tuición	Tuición administrativa del puente: Dirección de Vialidad, DDQ-APSA, DDQ Ingenieros Consultores.
8	Código Ruta	Código alfa-numérico. Código abreviado de la ruta del puente.
9	Clase Ruta	Clase de camino: Caminos nacionales, Caminos regionales principales, Caminos regionales comunales, Caminos regionales provinciales, Caminos regionales de acceso.
10	KM	Numérico. Kilómetro de ruta donde se ubica el puente.
11	Tipología General	Tipología del puente: Viaducto, Puente, No observable, Paso Inferior, Paso Superior, Modular.
12	Tipología Estructural	Tipología estructural del puente: Convencional vigas, Convencional cajón, Convencional losa, Convencional celosía, Colgante, Arco, Otros, No observable.
13	Longitud total (m)	Numérico.
14	Ancho total (m)	Numérico.
15	Número de tramos	Numérico.
16	Año de construcción	Numérico. Formato de número, no de fecha.
17	Número de Cepas	Numérico.
18	Ancho de Calzada (m)	Numérico.
19	Ancho de Pasillo (m)	Numérico.
20	Espesor de Losa (m)	Numérico.
21	Número de Vigas	Numérico.
22	Material cepas	Material de cepas: Hormigón, Hormigón armado, Hormigón en masa, Gaviones, Acero, Madera, No existe.
23	Fundación cepas	Tipología de fundación de cepa: Profunda: cajón hincado; Superficial: zapata; Otra tipología; No observable; No existe.
24	Material estribos	Material de estribos: Hormigón, Hormigón armado, Acero, Gaviones, Madera, Tierra reforzada, Otros materiales, No observable, No existe.

25	Fundación estribos	Tipología de fundación de estribos: Superficial: zapata; Superficial: no determinada; Superficial: otra tipología; Profunda: cajón hincado; Profunda: encepado + pilotes; No observable.
26	Material fundación estribos	Material de fundación de estribos: Hormigón armado, No observable.
27	Carpeta	Material de carpeta: Acero, Asfáltico, Hormigón, Hormigón armado, Hormigón en masa, Hormigón no determinado, Madera, Relleno granular, Tierra reforzada, Otros materiales, No aplica, No existe, No observable.
28	Vigas	Material de vigas: Hormigón postensado, Hormigón armado, Acero, Aluminio, Madera.
29	Losa	Material de losa: Hormigón, Hormigón armado, Hormigón no determinado, Acero, Madera, No Existe.

ANEXO 3.3 Clasificación Koppen-Geiger

Tabla A.3 Codificación clasificación climática Koppen-Geiger

Clima	Código	Descripción
Árido	BWh	Hot Desert climate
	BWh (s)	Hot Desert climate with dry summer
	BWk	Cold Desert climate
	BWk(s)	Cold Desert climate with dry summer
	BWk(w)	Cold Desert climate with dry winter
	BSk'(s)	Cold Semi-arid climate with dry summer and oceanic influence
	BSk	Cold Semi-arid climate
	BSk(s)	Cold Semi-arid climate with dry summer
	BSk(w)	Cold Semi-arid climate with dry winter
Templado	Csb'	Mediterranean climate (warm summer) with oceanic influence
	Csb	Mediterranean climate (warm summer)
	Csb(h)	Mediterranean climate (warm summer) with mountain influence
	Csc	Mediterranean climate (mild summer)
	Cfb'(s)	Marine West Coast climate with oceanic influence
	Cfb(s)	Marine West Coast climate (warm and dry summer)
	Cfb'	Marine West Coast climate (warm summer) with oceanic influence
	Cfb	Marine West Coast climate (warm summer)
	Cfc(s)	Marine West Coast climate with mild and dry summer
Cfc	Marine West Coast climate (mild summer)	
Polar	ET	Tundra climate
	ET(s)	Tundra climate with dry summer
	ET(w)	Tundra climate with dry winter
	EF	Ice Cap climate
	EF(s)	Ice Cap climate with dry summer

	EF(w)	Ice Cap climate with dry winter
--	-------	---------------------------------

ANEXO 3.4 Contenido base de datos de tráfico

Tabla A. 4 Campos de base de datos de tráfico

ID	Variable	Descripción
1	CODIGO	Código alfanumérico. Código abreviado de la ruta del puente.
2	KI	Numérico. Kilómetro de inicio del tramo de la ruta.
3	KF	Numérico. Kilómetro de fin del tramo de la ruta.
4	TRAMO	Numérico. Número de tramo de la ruta.
5	TMDA	Numérico. Tráfico Medio Diario Anual del tramo de la ruta.
6	P-A	Numérico. % de TMDA correspondiente a Autos en el tramo de la ruta.
7	P-C	Numérico. % de TMDA correspondiente a Camionetas en el tramo de la ruta.
8	P-C2E	Numérico. % de TMDA correspondiente a Camiones de dos ejes en el tramo de la ruta.
9	P-SR	Numérico. % de TMDA correspondiente a Semiremolques en el tramo de la ruta.
10	P-R	Numérico. % de TMDA correspondiente a Remolques en el tramo de la ruta.
11	P-C+2E	Numérico. % de TMDA correspondiente a Camiones de más de dos ejes en el tramo de la ruta.
12	P-B	Numérico. % de TMDA correspondiente a Buses en el tramo de la ruta.
13	Veloc_tramo	Numérico. Velocidad del tramo de la ruta.

ANEXO 3.5 Cálculo del Índice de Calificación de la Condición

En la ecuación A.3.1 y la ecuación A:3.2 se aprecia el cálculo de la calificación de deterioro, donde se pondera su extensión y su gravedad (Tabla 0.4A2.2.1), su tipo (Tabla 0.5 A2.2.2), su causa (Tabla 0.6 A2.2.3) y la unidad a la que corresponde el deterioro (Tabla 0.7A2.2.4). En las tablas mencionadas se presentan extractos de los ponderadores influyentes en la calificación de condición de deterioro de puentes, aunque los desarrolladores no detallan como se obtuvieron los valores de ponderación.

$$Cd = Co \cdot Pd \cdot Pc \cdot Pu \quad (A.3.1)$$

$$Co = \frac{100 - (V_{Extensión} + V_{Gravedad}) \cdot 4}{100} \quad (A.3.2)$$

Donde:

- Cd se refiere a la calificación de cada deterioro.
- $V_{Gravedad}$ se define para cada deterioro y se le asigna un coeficiente de acuerdo con la Tabla A.1
- $V_{Extensión}$ se define para cada deterioro y se le asigna un coeficiente de acuerdo con la Tabla A.1
- Pd se refiere al coeficiente de cada deterioro de acuerdo con la Tabla A.3
- Pc se refiere al coeficiente de la causa de deterioro de acuerdo con la Tabla A.2
- Pu se refiere al coeficiente de la unidad de un elemento afectado por el deterioro de acuerdo con la Tabla A.3
- Pd se refiere al coeficiente de cada deterioro de acuerdo con la Tabla A.4

Tabla A. 5 Ponderación Gravedad y Extensión en Calificación de deterioro

Gravedad	$V_{Gravedad}$	Extensión	$V_{Extensión}$
Mínima	0.10	Pequeña (<25%)	0.15
Media	3.00	Media (25%<Ext<50%)	0.50
Alta	8.00	Elevada (50%<Ext<75%)	2.00
Muy alta	13.00	Muy Elevada (>75%)	3.50

Tabla A. 6 Tipos de Causas con ponderación

ID	Causa	Ponderación (Pc)
1	Abrasión por tránsito	0,98
2	Acción climática	0,98

3	Acción de la fauna	0,98
4	Asiento diferencial	0,85
5	Ataque químico	0,90
6	Ausencia o deficiente dispositivo de junta	0,97
7	Ausencia o deficiente losa de transición	0,97
8	Cabeceo del muro lateral y existencia de un punto fijo	0,95
9	Calidad deficiente de la pieza/elemento	0,96
10	Cama de nivelación en mal estado o mal ejecutada	0,94
11	Capilaridad	0,99
12	Causa desconocida	0,90
13	Ciclos hielo-deshielo	0,97
14	Corrosión	0,93
15	Corrosión de las armaduras	0,92
16	Deficiente compactación del relleno	0,93
17	Deficiente ejecución	0,97
18	Deformación natural del propio aparato de apoyo	0,99
19	Exceso de compresión	0,95
20	Falta de barbacana	0,97
21	Falta de compresión	0,95
22	Falta de cuña en el neopreno	0,98
23	Falta de limpieza de encofrado	1,00
24	Falta de protección	0,98
25	Falta de protección de anclajes	1,00
26	Falta de protección frente a avenidas	0,91
27	Giro	0,88
28	Giro del estribo	0,85
29	Giro del muro lateral	0,90
30	Golpe o impacto	0,95
31	Impermeabilización defectuosa	0,98
32	Infradimensionamiento de apoyos	0,95
33	Infradimensionamiento del elemento	0,90
34	Mal funcionamiento del sistema de drenaje	0,97

Tabla A. 7 Tabla de pesos de ponderación por unidad

ID	Unidad	Ponderación (Pu)
001	Vigas (Tablero: 02 Convencional (de vigas))	0,91
002	Losa (Tablero: 02 Convencional (de vigas))	0,94
003	Voladizo (Tablero: 02 Convencional (de vigas))	0,94
004	Placas encofrado perdido (Tablero: 02 Convencional (de vigas))	0,98
005	Prelosas colaborantes (Tablero: 02 Convencional (de vigas))	0,96
006	Travesaño (Tablero: 02 Convencional (de vigas))	0,96
007	Imposta (Tablero: 02 Convencional (de vigas))	0,98

008	Cartela (Tablero: 02 Convencional (de vigas))	0,94
009	Costilla (Tablero: 02 Convencional (de vigas))	0,94
010	Revestimiento (Tablero: 02 Convencional (de vigas))	0,98
011	Otros: (Tablero: 02 Convencional (de vigas))	0,96
012	Losa (Tablero: 03 Convencional (losa/cajón))	0,91
013	Losa superior (Tablero: 03 Convencional (losa/cajón))	0,91
014	Losa inferior (Tablero: 03 Convencional (losa/cajón))	0,91
015	Almas (Tablero: 03 Convencional (losa/cajón))	0,91
016	Voladizo (Tablero: 03 Convencional (losa/cajón))	0,94
017	Placas encofrado perdido (Tablero: 03 Convencional (losa/cajón))	0,98
018	Prelosas colaborantes (Tablero: 03 Convencional (losa/cajón))	0,96
019	Travesaño (Tablero: 03 Convencional (losa/cajón))	0,96
020	Imposta (Tablero: 03 Convencional (losa/cajón))	0,98
021	Cartela (Tablero: 03 Convencional (losa/cajón))	0,94
022	Costilla (Tablero: 03 Convencional (losa/cajón))	0,94
023	Revestimiento (Tablero: 03 Convencional (losa/cajón))	0,98
024	Otros: (Tablero: 03 Convencional (losa/cajón))	0,96
025	Cordón superior (Tablero: 04 Convencional (celosía))	0,91
026	Cordón inferior (Tablero: 04 Convencional (celosía))	0,91
027	Diagonal (Tablero: 04 Convencional (celosía))	0,91
028	Montante (Tablero: 04 Convencional (celosía))	0,91
029	Voladizo (Tablero: 04 Convencional (celosía))	0,94
030	Placas encofrado perdido (Tablero: 04 Convencional (celosía))	0,98
031	Losa (Tablero: 04 Convencional (celosía))	0,94
032	Travesaño (Tablero: 04 Convencional (celosía))	0,96
033	Imposta (Tablero: 04 Convencional (celosía))	0,98
034	Cartela (Tablero: 04 Convencional (celosía))	0,94
035	Costilla (Tablero: 04 Convencional (celosía))	0,94
036	Revestimiento (Tablero: 04 Convencional (celosía))	0,98
037	Otros: (Tablero: 04 Convencional (celosía))	0,96
038	Bóveda (Tablero: 05 Bóveda)	0,91
039	Tímpano (Tablero: 05 Bóveda)	0,94
040	Bóveda (tímpano) (Tablero: 05 Bóveda)	0,94
041	Boquilla (tímpano) (Tablero: 05 Bóveda)	0,96
042	Montante (tímpano) (Tablero: 05 Bóveda)	0,96
043	Boquilla (Tablero: 05 Bóveda)	0,96
044	Voladizo (Tablero: 05 Bóveda)	0,96
045	Imposta (Tablero: 05 Bóveda)	0,98
046	Costilla (Tablero: 05 Bóveda)	0,94
047	Revestimiento (Tablero: 05 Bóveda)	0,98
048	Otros: (Tablero: 05 Bóveda)	0,96
049	Dintel (Tablero: 07 Cajón)	0,91
050	Tímpano (Tablero: 07 Cajón)	0,96
051	Voladizo (Tablero: 07 Cajón)	0,96

052	Imposta (Tablero: 07 Cajón)	0,98
053	Revestimiento (Tablero: 07 Cajón)	0,98
054	Otros: (Tablero: 07 Cajón)	0,96
055	Anillo (Tablero: 08 Tubo)	0,91
056	Tímpano (Tablero: 08 Tubo)	0,96
057	Voladizo (Tablero: 08 Tubo)	0,96
058	Imposta (Tablero: 08 Tubo)	0,98
059	Revestimiento (Tablero: 08 Tubo)	0,98
060	Otros: (Tablero: 08 Tubo)	0,96
061	Fustes (Cepa)	0,94
062	Fundación (Cepa)	0,94
063	Travesaño (Cepa)	0,96
064	Cargadero (Cepa)	0,96
065	Orejas (Cepa)	0,96
066	Pedestal (Cepa)	0,96
067	Tajamar (Cepa)	0,96
068	Capitel (Cepa)	0,94
069	Jabalcones (Cepa)	0,94
070	Revestimiento (Cepa)	0,98
071	Otros I: embellecedor perimetral (Cepa)	0,96
072	Otros II: (Cepa)	0,96
073	Fundación (Estribo)	0,94
074	Dintel, cargadero (Estribo)	0,94
075	Muro frontal portante (Estribo)	0,94
076	Muro de contención de tierras (Estribo)	0,96
077	Fustes, machones (Estribo)	0,94
078	Placas de machones (Estribo)	0,96
079	Murete espaldar (Estribo)	0,96
080	Aletas (Estribo)	0,96
081	Orejas (Estribo)	0,98
082	Tajamar (Estribo)	0,96
083	Anclajes (Estribo)	0,96
084	Dispositivo antisismo (Estribo)	0,98
085	Revestimiento (Estribo)	0,98
086	Otros I: (Estribo)	0,96
087	Otros II: (Estribo)	0,96
088	Fundación (Hastial)	0,94
089	Muro frontal (Hastial)	0,94
090	Muro de contención de tierras (Hastial)	0,96
091	Imposta (Hastial)	0,98
092	Otros I: (Hastial)	0,96
093	Otros II: (Hastial)	0,96
094	Calzada Tramo I (Junta)	0,96
095	Calzada Tramo II (Junta)	0,96

096	Pasillo Tramo I (Junta)	0,96
097	Pasillo Tramo II (Junta)	0,96
098	Aparato de apoyo (Apoyo)	0,96
099	Aparato de apoyo (Apoyo; Cuña)	0,96
100	Aparato de apoyo (Apoyo; Cama de nivelación)	0,98
101	Sistema de contención I	0,91
102	Sistema de contención II	0,91
103	Sumideros (ENR)	0,98
104	Barbacanas (ENR)	0,98
105	Tubos de desagüe (ENR)	0,98
106	Vierteaguas (ENR)	0,98
107	Cortagoteras (ENR)	0,98
108	Cunetilla (ENR)	0,98
109	Mechinales (ENR)	0,98
110	Canaleta (ENR)	0,98
111	Cuneta (ENR)	0,98
112	Cámaras (ENR)	0,98
113	Otros I: (ENR)	0,96
114	Otros II: (ENR)	0,96
115	Horizontal (ENR)	0,98
116	Vertical (ENR)	0,98
117	Pórtico no luminoso (ENR)	0,98
118	Pórtico luminoso (ENR)	0,98
119	Manga de viento (ENR)	0,96
120	Otros I: (ENR)	0,96
121	Otros II: (ENR)	0,96
122	Pantalla antirruído (ENR)	0,98
123	Valla antivertido (ENR)	0,98
124	Valla cerramiento (ENR)	0,98
125	Valla antivandálica (ENR)	0,98
126	Protección catenaria (ENR)	0,98
127	Pórtico pregálibo (ENR)	0,98
128	Farolas (ENR)	0,98
129	Luminarias (ENR)	0,98
130	Focos (ENR)	0,98
131	Iluminación ornamental (ENR)	0,98
132	Cámaras de registros (ENR)	0,98
133	Tuberías eléctricas (ENR)	0,98
134	Tuberías telefónicas (ENR)	0,98
135	Tuberías telecomun. (ENR)	0,98
136	Tuberías semaforicas (ENR)	0,98
137	Tuberías gas (ENR)	0,98
138	Tuberías Abastecim. (ENR)	0,98
139	Otras Tuberías (ENR)	0,98

140	Pavimento (ENR)	0,96
141	Bordillo (ENR)	0,98
142	Pasillos (ENR)	0,98
143	Mediana (ENR)	0,98
144	Carril ciclista (ENR)	0,99
145	Bancos (ENR)	0,99
146	Papeleras (ENR)	0,99
147	Estatuas (ENR)	0,99
148	Plantación ornamental (ENR)	0,99
149	Fuentes (ENR)	0,99
150	Otros I: (ENR)	0,98
151	Otros II: (ENR)	0,96
152	Terraplenes (ENR)	0,94
153	Enrocado (ENR)	0,91
160	Tirante (Tablero: 07 Atirantado)	0,89
161	Anclaje inferior (Tablero: 07 Atirantado)	0,89
162	Capot inferior (Tablero: 07 Atirantado)	0,98
163	Macizo anclaje inferior (Tablero: 07 Atirantado)	0,92
164	Tubo guía inferior (Tablero: 07 Atirantado)	0,95
165	Anillo centrador inferior (Tablero: 07 Atirantado)	0,92
166	Tubo antivandálico inferior (Tablero: 07 Atirantado)	0,98
167	Amortiguador inferior (Tablero: 07 Atirantado)	0,95
168	Anclaje superior (Tablero: 07 Atirantado)	0,89
169	Capot superior (Tablero: 07 Atirantado)	0,98
170	Macizo anclaje superior (Tablero: 07 Atirantado)	0,92
171	Tubo guía superior (Tablero: 07 Atirantado)	0,95
172	Anillo centrador superior (Tablero: 07 Atirantado)	0,92
173	Tubo antivandálico superior (Tablero: 07 Atirantado)	0,98
174	Amortiguador superior (Tablero: 07 Atirantado)	0,95
200	Arranques (fundación) del arco (Tablero: 06 Arco)	0,94
201	Anillo (Tablero: 06 Arco)	0,91
202	Riostras entre anillos (Tablero: 06 Arco)	0,94
203	Conexión con tablero (Tablero: 06 Arco)	0,96
204	Péndolas (Tablero: 06 Arco)	0,94
205	Imposta (Tablero: 06 Arco)	0,98
206	Costilla (Tablero: 06 Arco)	0,94
207	Otros: (Tablero: 06 Arco)	0,96

ANEXO 3.6 Tipos de deterioro

Tabla A.8 Tipos de deterioros con ponderación

ID	Deterioro	Ponderación (Pd)
1	Alteración superficial	97
2	Anclajes sueltos	93
3	Armadura vista	96
4	Baches, ahuellamiento	97
5	Cárcavas	96
6	Circulación	0
7	Corrosión	92
8	Deformación	89
9	Depositación	96
10	Despegue	95
11	Desplazamiento	93
12	Desprendimientos de hormigón	92
13	Desprendimientos de hormigón sin armadura vista	99
14	Eflorescencias	98
15	Falta de alineación	99
16	Fisuras	88
17	Fisuras en mapa	93
18	Grafiti	100
19	Grietas (>5 mm)	87
20	Hongos	94
21	Humedades, filtraciones	99
22	Hundimiento	92
23	Iluminación	99
24	Lajación	92
25	Madrigueras	98
26	Nido de piedras	99
27	Obstrucción	94
28	Pátina, mancha de óxido	99
29	Pérdida de finos	93
30	Pérdida de pieza	91
31	Pérdida de tornillos, roblones, anclajes	93
32	Pérdida de tratamiento protector	98
33	Piezas sueltas	92
34	Presencia de agua	99
35	Pudrición / ataque de insectos	93
36	Restos metálicos de ejecución (clavos, barras)	99
37	Rotura	85
38	Socavación local	80
39	Vegetación	99

ANEXO 5.1 Zonificación Climática de Chile

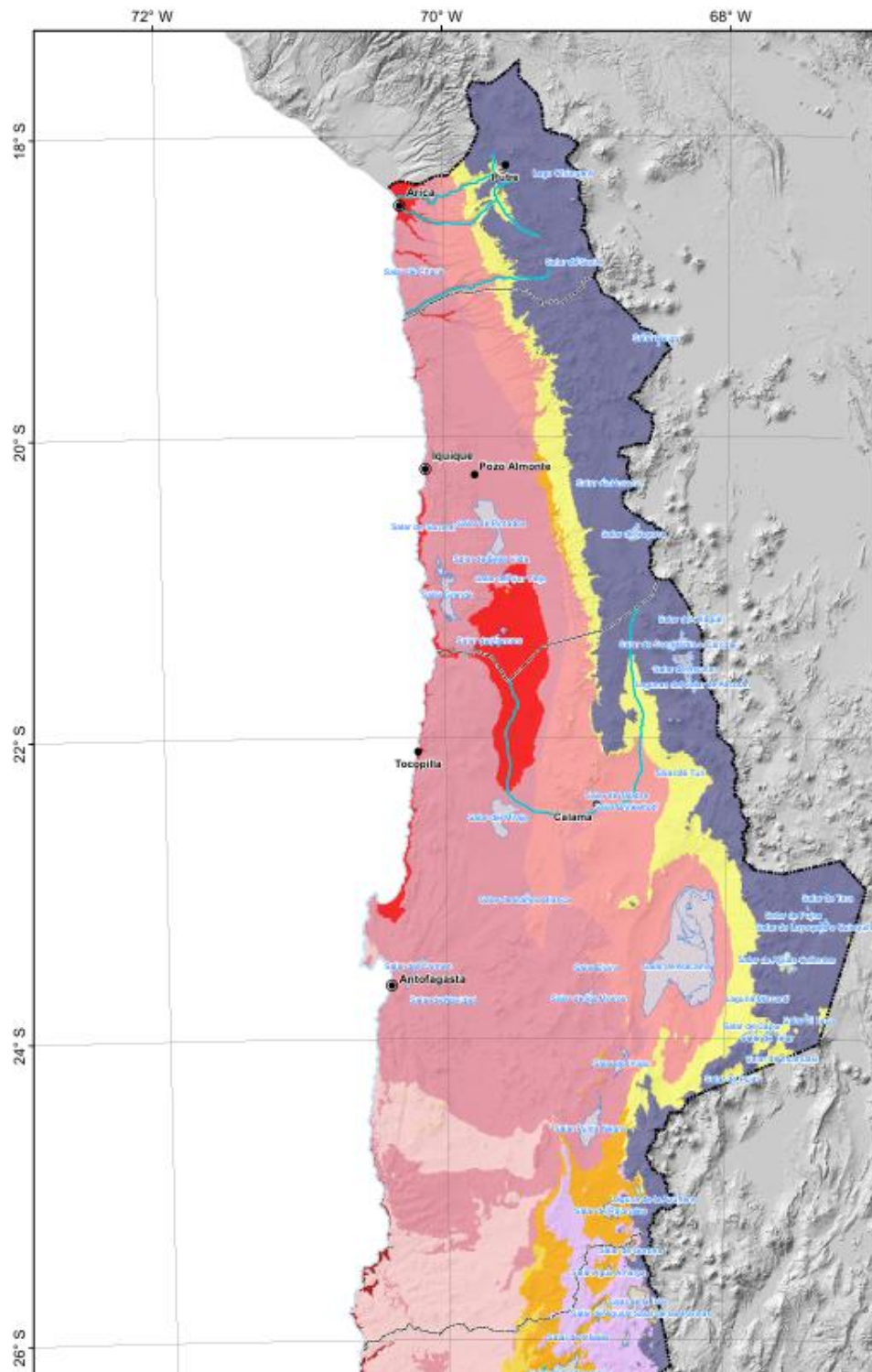


Figura A.1 Zonificación climática del Norte Grande de Chile

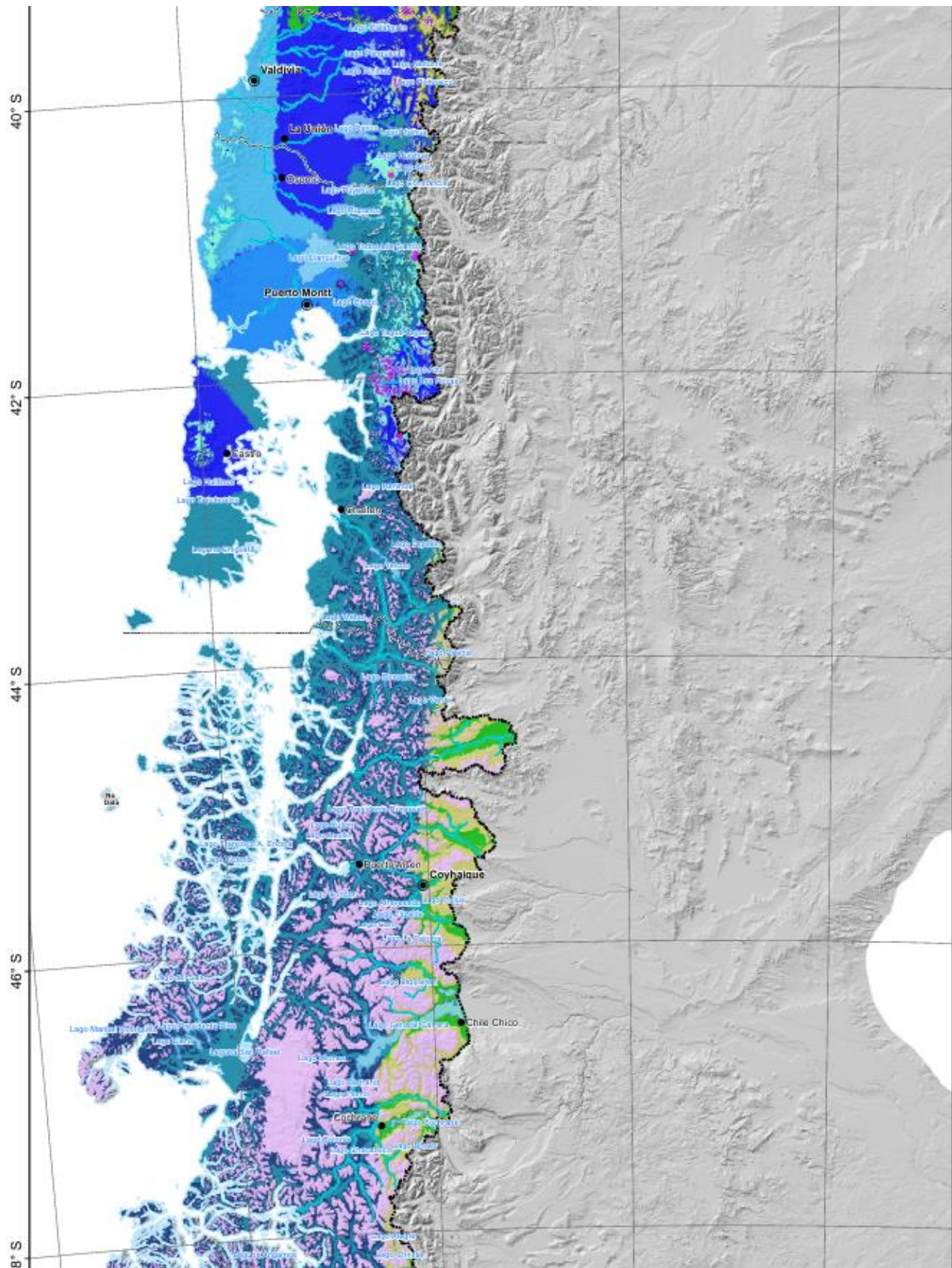


Figura A.5 Zonificación Climática Zona Sur y Austral de Chile

UNIVERSIDAD DE CONCEPCIÓN – FACULTAD DE INGENIERÍA
RESUMEN DE MEMORIA DE TÍTULO

Departamento : Departamento de Ingeniería Civil
Carrera : Ingeniería Civil
Nombre del memorista : Cristobal Cisternas Casas
Título de la memoria : Diseño Experimental para la Modelación del Deterioro de Puentes convencionales tipo viga, usando herramientas de IA
Fecha de la presentación oral : 28 de enero del 2026

Profesor(es) Guía : Tomas Echaveguren Navarro
Profesor(es) Revisor(es) : Juan Carrasco Montagna
Víctor Aguilar Vidal

Concepto :
Calificación :

Resumen

La gestión de la infraestructura vial en Chile ha dependido históricamente de enfoques reactivos basados en inspecciones visuales, careciendo de modelos predictivos de deterioro adaptados a la realidad nacional. El objetivo de este trabajo es desarrollar un diseño experimental para modelar la degradación en puentes convencionales de tipo viga, integrando principios de ingeniería con técnicas de procesamiento y análisis de datos. En la metodología se definió un vector de características compuesto por diez variables explicativas críticas, abarcando factores climáticos, de tráfico, materialidad y diseño. El procedimiento incorpora técnicas de Inteligencia Artificial en un esquema secuencial: el uso de embeddings de Modelos de Lenguaje Grandes (LLM) para la vectorización de variables categóricas, algoritmos de Machine Learning para la reducción dimensional y el Método de las Ventanas para la generación de series temporales de datos. Finalmente, se formuló un protocolo de calibración y validación comparativa para nueve arquitecturas de regresión, contrastando métodos probabilísticos y deterministas con modelos de aprendizaje profundo (Deep Learning). El estudio demuestra que el diseño experimental propuesto permite transformar los registros históricos de inspección visual en herramientas técnicas de predicción para la gestión de redes a gran escala.