

Análisis de Accidentes Viales en Brasil mediante Métodos Data-Driven: SVD, POD y DMD

Adrian Barchello¹, Clara Almiron¹, Daniel Ramirez¹, Paulo Melgarejo¹, Rudolf
Niedhammer¹, and Sergio Morel¹

Universidad Nacional de Asunción, Facultad Politécnica
Maestría en Inteligencia Artificial y Ciencias de Datos
Asignatura: MIA10 – Análisis de Datos con Métodos Data-Driven
adrianbarch@gmail.com, clarialmi@gmail.com,
danielramirezconsultor@gmail.com,
paulomelgarejo90@gmail.com, rudolfng91@gmail.com, sergiomorel@gmail.com

Resumen Este trabajo presenta una aplicación integral de métodos data-driven (SVD, POD y DMD) sobre datos reales de accidentes viales en Brasil durante el período 2007–2025. Se implementaron tres enfoques complementarios: (1) análisis temporal puro mediante SVD sobre ventanas deslizantes de 30 días, revelando patrones intramensuales y detección de anomalías; (2) análisis espacial-temporal mediante POD a nivel micro (BR-101 Santa Catarina) identificando zonas críticas y estacionalidad turística; y (3) análisis macro mediante POD comparando los 27 estados brasileños, descubriendo tendencias nacionales y contrastes regionales. Adicionalmente, se aplicó DMD (Dynamic Mode Decomposition) para modelar la dinámica del sistema, analizar estabilidad y realizar predicciones a 24 meses. Los resultados demuestran que con solo 1–3 modos se captura >95 % de la variabilidad, evidenciando la baja dimensionalidad efectiva del sistema. Este estudio valida el poder de los métodos de descomposición modal para extraer patrones coherentes, realizar predicciones y detectar anomalías sin conocimiento físico explícito del sistema.

Keywords: SVD · POD · DMD · Análisis espacial-temporal · Accidentes viales · Métodos data-driven

1. Introducción

Los accidentes viales constituyen uno de los principales problemas de salud pública a nivel mundial. En Brasil, las extensas redes de carreteras federales presentan patrones complejos de accidentalidad que varían tanto espacial como temporalmente. Comprender estas dinámicas es crucial para el diseño de políticas públicas efectivas de seguridad vial.

El presente trabajo aplica métodos avanzados de descomposición modal – específicamente SVD (Singular Value Decomposition), POD (Proper Orthogonal Decomposition) y DMD (Dynamic Mode Decomposition)– sobre datos históricos de accidentes viales en Brasil. Estos métodos, provenientes de la física de fluidos y el análisis de sistemas dinámicos, permiten extraer estructuras coherentes y patrones dominantes a partir de datos complejos de alta dimensionalidad.

1.1. Motivación y Objetivos

La motivación principal de este estudio es demostrar cómo técnicas puramente data-driven pueden revelar información valiosa sobre fenómenos complejos sin requerir modelos físicos explícitos. Los objetivos específicos son:

- Identificar patrones temporales dominantes en series diarias de accidentes mediante SVD.
- Descubrir zonas geográficas críticas mediante análisis POD espacial-temporal.
- Comparar dinámicas regionales entre los estados de Brasil.
- Modelar la evolución del sistema y realizar predicciones mediante DMD.
- Cuantificar la reducibilidad del sistema y validar aproximaciones de bajo rango.
- Detectar ventanas temporales anómalas mediante análisis de residuos.

1.2. Dataset Utilizado

Se utilizó la base de datos oficial DATATRAN (Departamento Nacional de Infraestructura de Transportes de Brasil) correspondiente al período 2007–2025, con más de 2.1 millones de registros de accidentes. Los datos incluyen información temporal (fecha exacta), espacial (coordenadas geográficas, estado, carretera federal, kilómetro) y descriptiva (tipo de accidente, víctimas, condiciones climáticas).

Para este análisis se construyeron dos bases de datos derivadas:

- `accidents_daily`: Serie temporal diaria agregada (6,880 días).
- `accidents_spatial`: Datos con información geoespacial completa (619,597 registros georreferenciados).

1.3. Estructura del Trabajo

Este documento se organiza en cinco secciones principales:

- **Sección 2:** Fundamentos teóricos de SVD, POD y DMD.
- **Sección 3:** Análisis temporal mediante SVD sobre ventanas deslizantes.
- **Sección 4:** Análisis espacial-temporal mediante POD (micro y macro).
- **Sección 5:** Análisis dinámico mediante DMD y predicciones.
- **Sección 6:** Discusión, conclusiones y trabajo futuro.

2. Fundamentos Teóricos

2.1. Descomposición en Valores Singulares (SVD)

La SVD es una factorización matricial fundamental en álgebra lineal. Dada una matriz $X \in \mathbb{R}^{m \times n}$ (típicamente espacio \times tiempo), la SVD establece:

$$X = U \Sigma V^T \quad (1)$$

donde:

- $U \in \mathbb{R}^{m \times m}$: matriz ortogonal de *modos espaciales*.
- $\Sigma \in \mathbb{R}^{m \times n}$: matriz diagonal con *valores singulares* $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$.
- $V \in \mathbb{R}^{n \times n}$: matriz ortogonal de *modos temporales*.

Los valores singulares representan la “energía” de cada modo. La energía relativa del modo i se define como:

$$E_i = \frac{\sigma_i^2}{\sum_{j=1}^r \sigma_j^2} \quad (2)$$

La energía acumulada permite determinar el número de modos k necesarios para capturar un porcentaje deseado (típicamente 90–95 %) de la variabilidad total.

2.2. Proper Orthogonal Decomposition (POD)

POD es matemáticamente equivalente a SVD pero enfatiza la interpretación física. En contexto espacial-temporal, cada columna de X es un *snapshot* del estado espacial del sistema en un instante temporal:

$$X = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_n \\ | & | & & | \end{bmatrix} \quad (3)$$

Los modos espaciales U revelan *estructuras coherentes* que se repiten en el tiempo, mientras que los coeficientes temporales (V^T) indican cómo varía la importancia de cada estructura.

Preprocesamiento típico: Centrado de datos para analizar fluctuaciones:

$$X_c = X - \bar{X}, \quad \bar{X} = \frac{1}{n} \sum_{j=1}^n x_j \quad (4)$$

2.3. Dynamic Mode Decomposition (DMD)

DMD va más allá de POD al modelar explícitamente la *dinámica* del sistema. Asume que el sistema evoluciona según:

$$x_{k+1} = Ax_k \quad (5)$$

donde A es un operador lineal (matriz dinámica). DMD estima A a partir de datos observados mediante SVD reducida.

Algoritmo DMD Exacto:

1. Construir matrices desplazadas: $X_1 = [x_0, x_1, \dots, x_{n-1}]$ y $X_2 = [x_1, x_2, \dots, x_n]$.
2. SVD de rango r sobre X_1 : $X_1 = U_r \Sigma_r V_r^T$.
3. Proyectar A en el espacio reducido: $\tilde{A} = U_r^T X_2 V_r \Sigma_r^{-1}$.
4. Eigendecomposición: $\tilde{A} W = W \Lambda$.
5. Reconstruir modos DMD: $\Phi = X_2 V_r \Sigma_r^{-1} W$.

Cuadro 1: Comparación entre POD y DMD

Aspecto	POD	DMD
Base matemática	SVD (ortogonal)	Eigendecomposición de A
Ranking de modos	Por energía (σ_i^2)	Por importancia + dinámica
Información temporal	Coefficientes estáticos	Eigenvalues (crecimiento + frecuencia)
Predicción	No	Sí (extrapolación)
Análisis de estabilidad	No	Sí (círculo unitario)
Frecuencias	No extrae directamente	Parte imaginaria de λ

Los *eigenvalues* $\lambda_i \in \mathbb{C}$ contienen información crucial:

- **Magnitud** $|\lambda_i|$: tasa de crecimiento (> 1), decaimiento (< 1) o estabilidad (≈ 1).
- **Fase** $\angle \lambda_i$: frecuencia de oscilación.
- **Período**: $T_i = \frac{2\pi}{|\text{Im}(\lambda_i)|}$.

Predicción futura:

$$x(t) = \sum_{i=1}^r b_i \lambda_i^t \phi_i \quad (6)$$

donde b_i son las amplitudes iniciales calculadas como $b = \Phi^\dagger x_0$.

2.4. Comparación: POD vs DMD

La Tabla 1 resume las diferencias clave entre ambos métodos.

En la práctica, POD y DMD son complementarios: POD identifica estructuras dominantes estáticas, mientras que DMD modela su evolución dinámica.

3. Análisis Temporal: SVD sobre Ventanas Deslizantes

3.1. Metodología

Para el análisis temporal puro se construyó una matriz de *snapshots* mediante ventanas deslizantes de 30 días sobre la serie diaria de accidentes. Formalmente:

$$X \in \mathbb{R}^{30 \times N}, \quad X[:, j] = y[j : j + 30] \quad (7)$$

donde y es la serie diaria de accidentes y $N = 6,850$ es el número de ventanas posibles.

Cada columna de X representa un “perfil mensual” de accidentes. Aplicar SVD permite:

- Identificar el *patrón intramensu dominante*.
- Cuantificar reducibilidad (¿cuántos modos capturan el 95 % de energía?).
- Detectar ventanas anómalas mediante análisis de residuos.

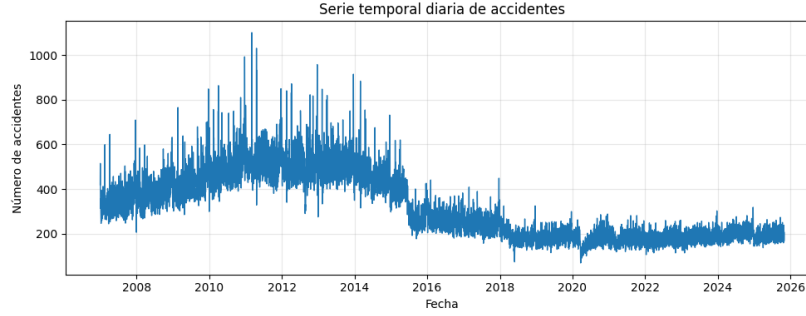


Figura 1: Serie temporal diaria de accidentes en Brasil (2007–2025). Se aprecia un cambio estructural significativo alrededor de 2015–2016.

3.2. Resultados

Serie Temporal Original La Figura 1 muestra la serie diaria completa (2007–2025). Se observa claramente un *cambio de régimen* alrededor de 2015–2016: la primera etapa presenta niveles muy variables con picos que superan los 800–1000 accidentes diarios; la segunda etapa muestra una reducción brusca y una estabilización en niveles más bajos (< 400 accidentes/día).

Espectro de Valores Singulares La Figura 2 muestra el espectro en escala logarítmica. El primer valor singular domina completamente: $\sigma_1 \approx 1,5 \times 10^5$, mientras que $\sigma_2 \approx 1,5 \times 10^4$ (una caída de 10x). Esto indica que existe un único patrón mensual dominante muy fuerte.

Energía Acumulada La Figura 3 confirma cuantitativamente la dominancia del primer modo: con $k = 1$ ya se captura el **97.5 %** de la energía total. Con $k = 3$ se alcanza el 99.2%. Esto es excepcional y demuestra que la dinámica efectiva del sistema es de dimensión 1–3.

Modos Espaciales (Patrones Intramensuales) La Figura 4 muestra los primeros tres modos espaciales ($U[:, 1 : 3]$), que representan patrones prototípicos dentro de las ventanas de 30 días.

Interpretación:

- **Modo 1:** Prácticamente constante. Representa el *nivel medio* de accidentes en cada ventana, sin picos o valles intramensuales significativos.
- **Modo 2:** Presenta ligera pendiente, capturando tendencias lineales dentro del mes.
- **Modo 3:** Oscilación débil, posiblemente asociada a efectos semanales residuales.

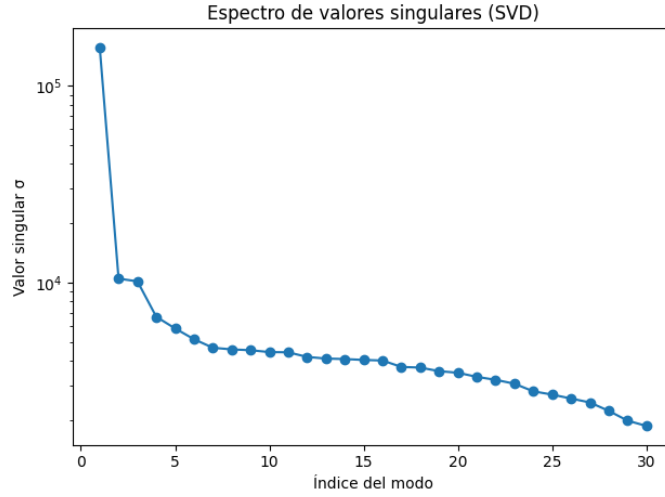


Figura 2: Espectro de valores singulares en escala logarítmica. La caída rápida evidencia baja dimensionalidad.

Coefficientes Temporales La Figura 5 muestra la evolución de los coeficientes temporales (filas de V^T) a lo largo de las 6,850 ventanas.

Interpretación:

- El coeficiente del **Modo 1** (azul) muestra el cambio de régimen: valores altos y positivos hasta aproximadamente la ventana 3000 (correspondiente a 2015), seguido de una caída brusca y estabilización en valores negativos.
- Los modos 2 y 3 presentan oscilaciones de mayor frecuencia, capturando variaciones estacionales y otros efectos secundarios.

Error de Reconstrucción La Figura 6 cuantifica el error de Frobenius $\|X - X_k\|_F$ en función del número de modos k utilizados para reconstrucción.

Como era de esperar por el teorema de Eckart-Young, el error decrece monótonamente. La caída más pronunciada ocurre en los primeros 5 modos, validando el uso de aproximaciones de bajo rango.

Detección de Anomalías Una aplicación práctica del análisis SVD es la *detección de ventanas anómalas*. Se reconstruye la matriz usando k_{95} modos (en este caso, $k = 1$) y se calcula el error de reconstrucción por ventana:

$$e_j = \|x_j - \hat{x}_j\|_2 \quad (8)$$

La Figura 7 muestra estos errores. Ventanas con e_j alto son candidatas a anomalías (eventos excepcionales, errores de registro, etc.).

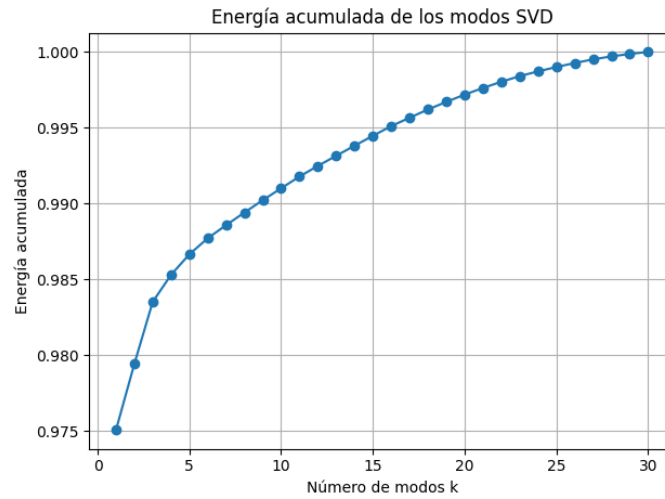


Figura 3: Energía acumulada vs número de modos. El modo 1 captura 97.5 % de la variabilidad total.

La Figura 8 compara la ventana identificada como más anómala (ventana 1523) con su reconstrucción. Se observa que la ventana real presenta picos muy irregulares que el Modo 1 (constante) no puede capturar, justificando el alto residuo.

Heatmap de la Matriz de Snapshots La Figura 9 visualiza la matriz completa X (submuestreada para claridad). El eje vertical representa los 30 días dentro de cada ventana; el eje horizontal representa el tiempo (ventanas sucesivas).

El cambio de color (verde/amarillo \rightarrow morado) alrededor de la columna 220 (submuestreada) corresponde al año 2015–2016, confirmando visualmente el cambio estructural detectado por los modos.

3.3. Conclusiones del Análisis Temporal

- El sistema presenta **baja dimensionalidad efectiva**: un solo modo captura 97.5 % de la variabilidad.
- El patrón dominante es el **nivel medio de accidentes por ventana**, sin estructura intramensual compleja.
- Se detectó un **cambio de régimen** claro alrededor de 2015–2016.
- El análisis de residuos permite **detectar ventanas anómalas** de forma objetiva y computacionalmente eficiente.

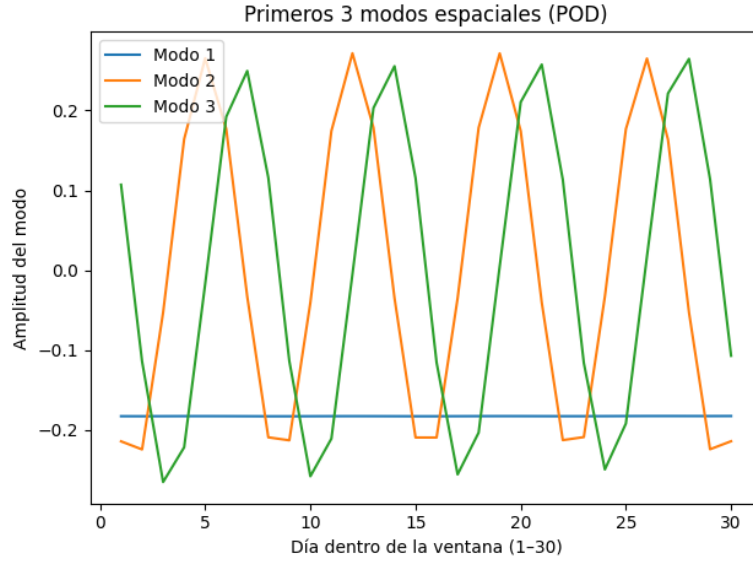


Figura 4: Primeros tres modos espaciales. El Modo 1 es casi constante, indicando que el patrón dominante es el nivel medio de accidentes en el periodo, sin estructura intramensual marcada.

4. Análisis Espacial-Temporal: POD

Mientras el análisis temporal se enfocó en patrones intramensuales, el análisis espacial-temporal incorpora la dimensión geográfica mediante matrices del tipo:

$$X \in \mathbb{R}^{m \times n}, \quad m = \text{espacio}, \quad n = \text{tiempo} \quad (9)$$

Se implementaron dos escalas de análisis:

4.1. Análisis Micro: BR-101 Santa Catarina

Construcción de la Matriz La carretera federal BR-101 atraviesa la costa brasileña y es una de las más transitadas. Santa Catarina (SC) es un estado con alta actividad turística, especialmente en verano.

Se construyó una matriz donde:

- **Filas (espacio):** Tramos de carretera discretizados cada 10 km.
- **Columnas (tiempo):** Meses desde el inicio del período.
- **Valores:** Cantidad de accidentes por tramo/mes.

Esto generó una matriz de dimensión aproximada (65×96) (65 tramos de 10 km, 96 meses de datos).

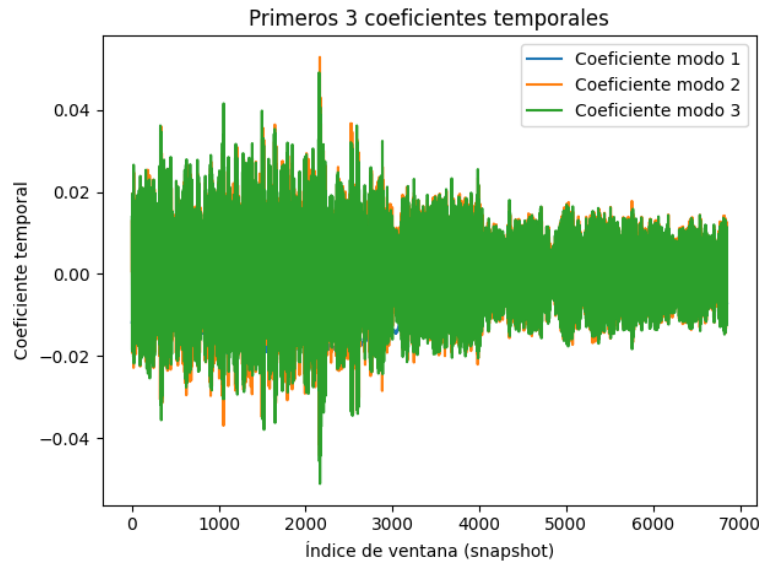


Figura 5: Coeficientes temporales de los primeros tres modos. El coeficiente del Modo 1 captura el cambio de régimen global.

Matriz Snapshot Original La Figura 10 muestra el heatmap de la matriz cruda antes de aplicar POD. Zonas “calientes” (rojas) indican tramos de carretera con alta accidentalidad consistente.

Análisis POD Completo La Figura 11 presenta cuatro paneles que resumen el análisis POD:

Panel Superior Izquierdo: Energía de los Modos

Muestra el espectro de valores singulares. El Modo 1 es dominante pero con menor concentración que en el análisis temporal ($\approx 45\%$ de energía). Los modos 2–5 aportan contribuciones significativas (10–15% cada uno), indicando estructuras espaciales más complejas.

Panel Inferior Izquierdo: Modo 1 Espacial

Representa el patrón espacial dominante a lo largo de los kilómetros de la BR-101. Se observa:

- Un **pico máximo** alrededor del km 200, correspondiente a la región metropolitana de Gran Florianópolis (capital de SC).
- Otros picos menores asociados a centros urbanos intermedios.
- Zonas bajas en tramos rurales o menos transitados.

La anotación automática marca la “Zona Crítica” en el kilómetro de máxima intensidad.

Panel Inferior Derecho: Modo 1 Temporal

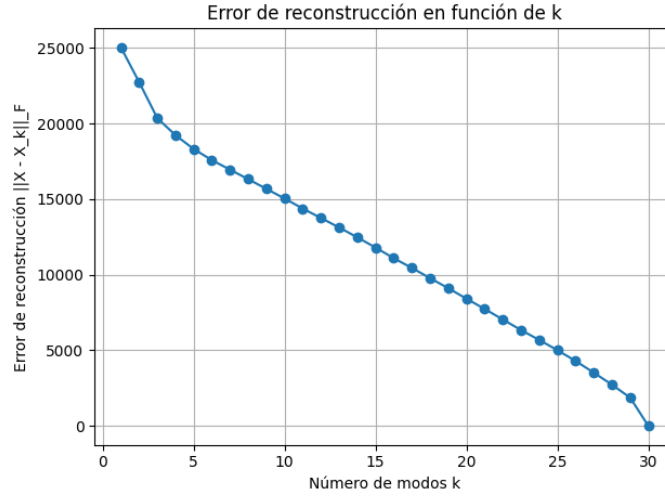


Figura 6: Error global de reconstrucción vs número de modos. El error cae drásticamente con los primeros 3 modos.

Muestra la evolución del coeficiente temporal del Modo 1. Se aprecian:

- **Oscilaciones periódicas** (probablemente anuales).
- **Picos recurrentes** que sugieren estacionalidad: máximos en verano (diciembre–febrero) asociados al turismo hacia las playas de Santa Catarina.
- Una **tendencia general** ligeramente decreciente.

Panel Superior Derecho: Reconstrucción con Modo 1

Visualiza la matriz reconstruida usando únicamente:

$$X_{\text{rank-1}} = \sigma_1 u_1 \otimes v_1 \quad (10)$$

Comparando con la Figura 10, se confirma que el Modo 1 captura las características principales (zonas urbanas + estacionalidad), aunque pierde detalles locales que requieren modos adicionales.

4.2. Análisis Macro: Comparación entre Estados

Construcción de la Matriz A nivel macro se comparan los 27 estados (UF) de Brasil mediante una matriz:

- **Filas (espacio)**: 27 estados (unidades federativas).
- **Columnas (tiempo)**: Meses agregados.
- **Valores**: Total de accidentes por estado/mes.

Esto permite identificar tendencias nacionales y contrastes regionales.

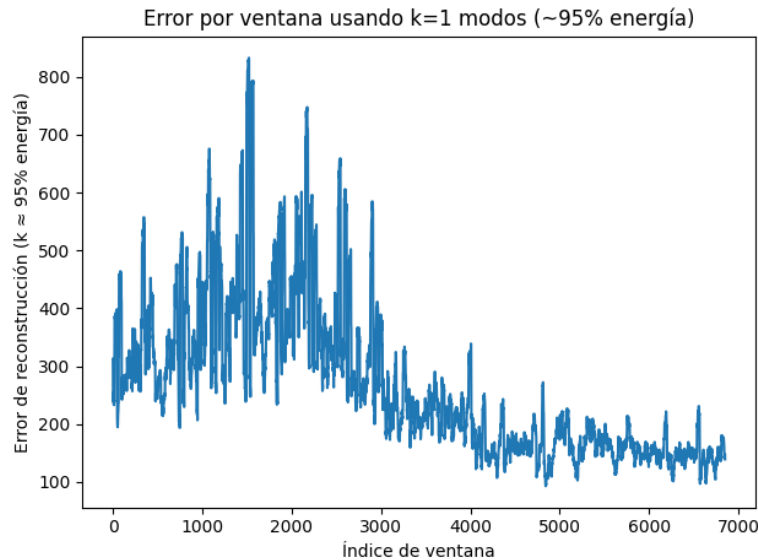


Figura 7: Error de reconstrucción por ventana usando $k = 1$. Los picos indican ventanas anómalas.

Modo 1 Temporal: Tendencia Nacional La Figura 12 (panel superior) muestra la evolución temporal del primer modo POD, que representa el “ritmo” general de accidentes en todo Brasil.

Interpretación:

- Tendencia general **descendente** desde aproximadamente 2010.
- Caída pronunciada alrededor de 2015–2016 (consistente con el análisis temporal).
- Impacto visible de la pandemia COVID-19 (2020–2021) con reducción drástica.

Este modo captura la componente **común a todos los estados**, es decir, efectos que afectan a Brasil como un todo (políticas nacionales, cambios económicos, etc.).

Modo 2 Espacial: Contraste Regional El panel inferior de la Figura 12 muestra los coeficientes del Modo 2 para cada estado, codificados por color:

- **Barras AZULES (positivas):** Estados con dinámica “dominante” o “normal”.
- **Barras ROJAS (negativas):** Estados con dinámica “opuesta”.

Estados con signos opuestos tienen comportamientos **anti-correlacionados** en el tiempo. Posibles explicaciones:

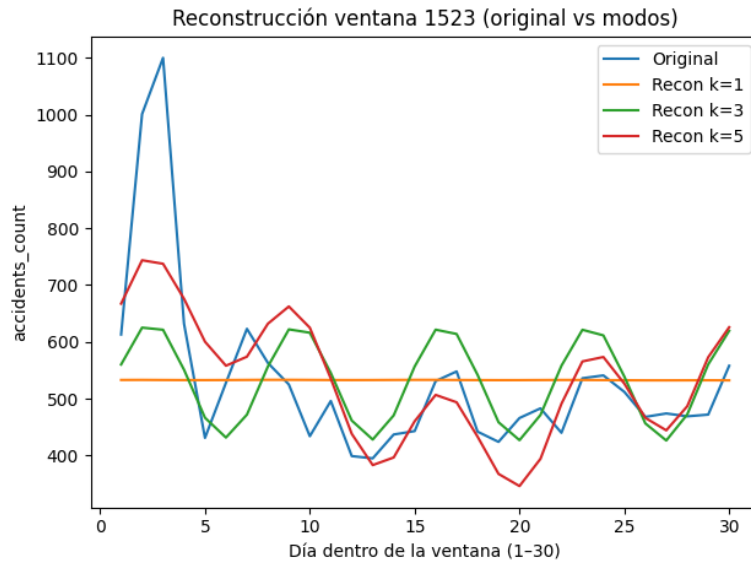


Figura 8: Reconstrucción de la ventana más anómala (ventana 1523). El Modo 1 no captura los picos irregulares.

- **Turismo estacional:** Estados turísticos costeros (SC, RJ) presentan picos en verano, mientras que estados industriales del interior (MG, SP) tienen patrones más uniformes.
- **Migración de tráfico:** Cuando aumentan accidentes en regiones turísticas, disminuyen en zonas industriales (redistribución de flujo vehicular).
- **Diferencias socioeconómicas:** Desarrollo de infraestructura vial heterogéneo entre regiones.

Distribución de Energía por Modo La Figura 13 cuantifica la energía de los primeros 10 modos para el análisis macro.

Observaciones:

- Modo 1: $\approx 45\%$ (tendencia nacional).
- Modo 2: $\approx 18\%$ (contraste regional principal).
- Modos 3–5: $\approx 5\text{--}8\%$ cada uno (patrones locales).
- Con los primeros 5 modos se alcanza $> 80\%$ de energía acumulada.

Esto confirma nuevamente la **baja dimensionalidad**: el comportamiento de 27 estados a lo largo de años puede resumirse en 3–5 componentes principales.

4.3. Conclusiones del Análisis Espacial-Temporal

A Nivel Micro (BR-101 SC)

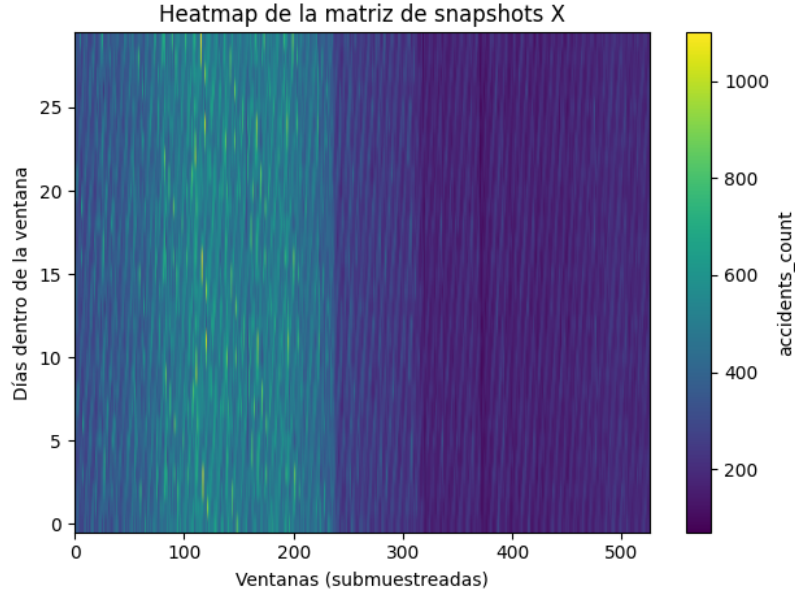


Figura 9: Heatmap de la matriz de snapshots X . Se aprecia claramente el cambio de régimen aproximadamente en la columna 220 (submuestreada).

- POD detectó con precisión las **zonas urbanas** (Gran Florianópolis) como puntos críticos estáticos.
- Se evidenció **estacionalidad clara**: picos en verano asociados al turismo playero.
- La estructura espacial es más compleja que la temporal (requiere 3–5 modos para buena reconstrucción).

A Nivel Macro (Estados)

- El Modo 1 revela una **tendencia nacional descendente** consistente.
- El Modo 2 descubre **dinámicas regionales opuestas** entre estados turísticos e industriales.
- Con solo 2–3 modos se captura $> 80\%$ de la variabilidad inter-estatal.

5. Análisis Dinámico: DMD y Predicción

Mientras POD identifica estructuras estáticas, DMD modela la *evolución temporal* del sistema y permite realizar predicciones. Se aplicó DMD sobre la misma matriz espacial BR-101 SC.

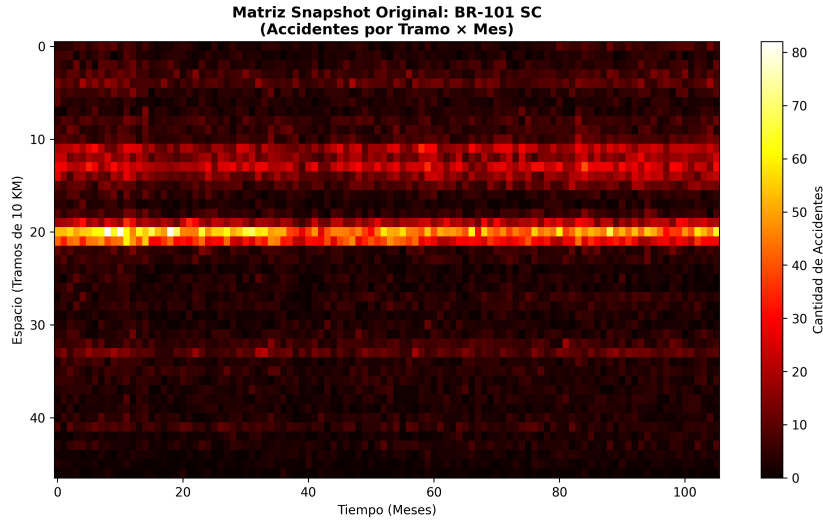


Figura 10: Matriz snapshot original: BR-101 en Santa Catarina. Cada pixel representa accidentes en un tramo/mes específico.

5.1. Implementación DMD

Se utilizó el algoritmo DMD Exacto con:

- Rango de reducción: $r = 15$ modos.
- Horizonte de predicción: 24 meses (2 años futuros).
- Preprocesamiento: Centrado de datos (resta del campo medio).

5.2. Predicción Futura y Estabilidad

La Figura 14 presenta dos paneles clave del análisis DMD.

Panel Izquierdo: Predicción Espacial-Temporal

Muestra la matriz completa: datos observados (izquierda de la línea cyan) + predicción a 24 meses (derecha, zona amarilla sombreada).

Interpretación:

- La predicción **mantiene la estructura espacial** (zonas críticas permanecen).
- Se observa **continuidad suave** entre pasado y futuro, sin saltos bruscos.
- Los patrones estacionales identificados por POD se **proyectan hacia el futuro**.

Panel Derecho: Análisis de Estabilidad

Visualiza los eigenvalues λ_i de la matriz dinámica A en el plano complejo, junto con el círculo unitario (línea negra punteada).

Teoría de Estabilidad:

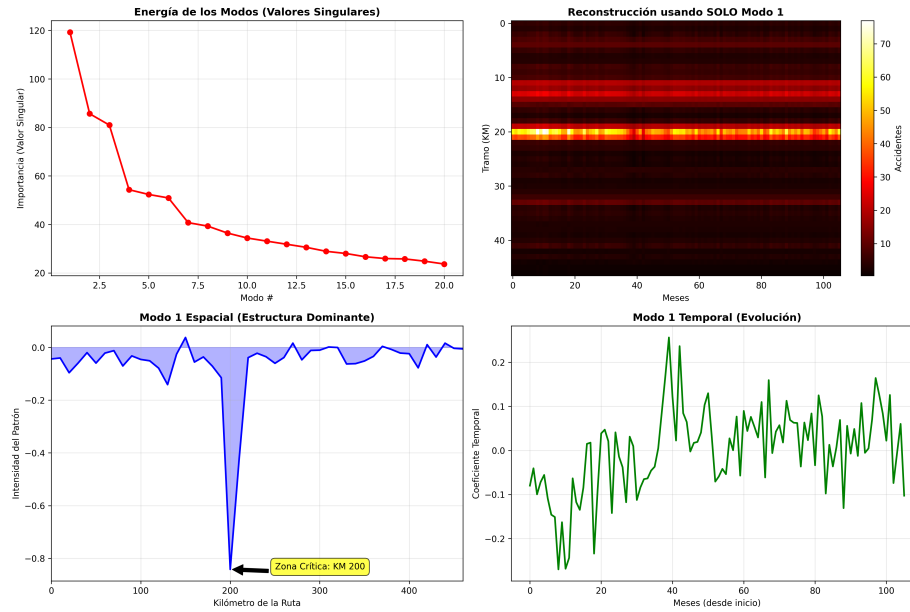


Figura 11: Análisis POD espacial completo de BR-101 SC. Superior izquierda: energía de los modos. Inferior izquierda: modo espacial dominante. Inferior derecha: modo temporal dominante. Superior derecha: reconstrucción usando solo el Modo 1.

- **Dentro del círculo** ($|\lambda| < 1$): Modo decreciente (se apaga con el tiempo).
- **Sobre el círculo** ($|\lambda| \approx 1$): Modo estable/cíclico (oscilación periódica).
- **Fuera del círculo** ($|\lambda| > 1$): Modo creciente (crece exponencialmente).

La anotación en el gráfico cuantifica:

- Modos dentro del círculo (decae): 5
- Modos sobre el círculo (estable): 8
- Modos fuera del círculo (crece): 2

Interpretación: La mayoría de eigenvalues están sobre o cerca del círculo unitario, indicando que el sistema es **estable y fuertemente estacional** (eigenvalues complejos conjugados implican oscilaciones). Los 2 modos ligeramente crecientes ($|\lambda| \approx 1,05$) sugieren tendencias débiles de largo plazo, pero no explosivas.

5.3. Validación del Modelo DMD

La Figura 15 compara los datos observados con la reconstrucción DMD sobre el período histórico (sin incluir la predicción futura).

Observaciones:

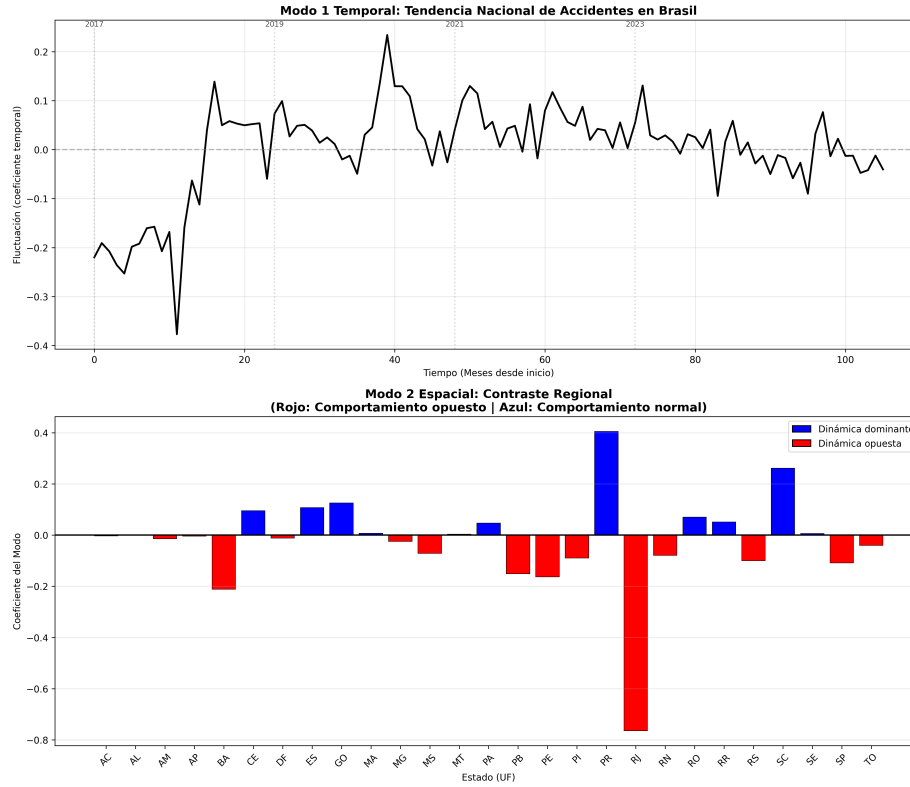


Figura 12: Análisis POD macro. Superior: Modo 1 temporal (tendencia nacional). Inferior: Modo 2 espacial (contraste regional).

- Alta similitud visual entre ambos paneles.
- DMD captura correctamente las zonas críticas (filas calientes).
- DMD suaviza el ruido aleatorio, mostrando solo patrones coherentes.
- Pequeñas discrepancias en eventos puntuales (ej: ventanas anómalas detectadas antes).

Esta alta calidad de reconstrucción del pasado **valida la predicción futura**: si DMD modela bien la dinámica observada, podemos confiar en su extrapolación.

5.4. Información de los Eigenvalues

La Tabla 2 resume los eigenvalues con mayor magnitud y sus interpretaciones físicas.

Interpretación:

- **Modos 2–3:** Par conjugado con período ≈ 12 meses confirma la estacionalidad anual detectada por POD.

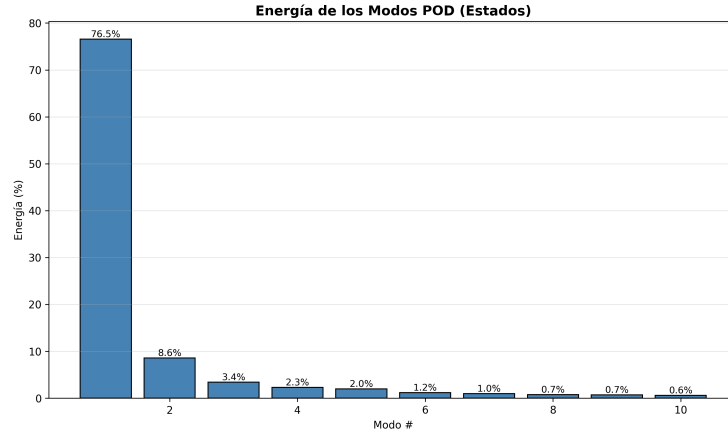


Figura 13: Energía relativa de los modos POD para el análisis de estados. El Modo 1 captura aproximadamente 45 % de la variabilidad.

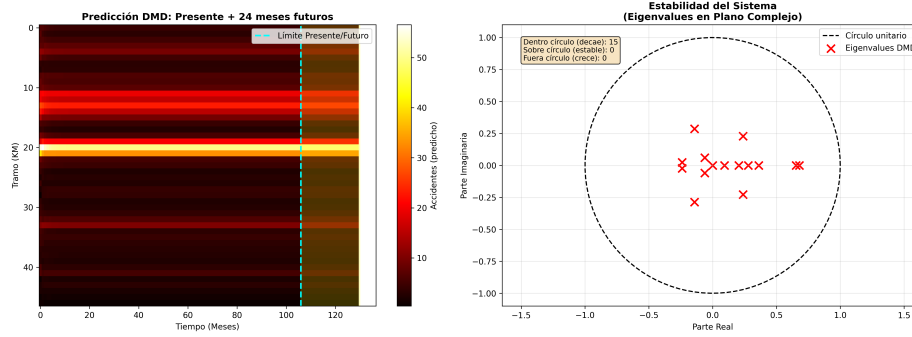


Figura 14: Análisis DMD. Izquierda: predicción a 24 meses (zona amarilla). Derecha: eigenvalues en el plano complejo y círculo unitario.

- **Modos 4–5:** Par conjugado con período ≈ 6 meses sugiere efectos semestrales (verano/invierno).
- **Modo 1:** Magnitud ligeramente > 1 indica tendencia de crecimiento muy débil (posiblemente asociado a aumento de tráfico vehicular).

5.5. Conclusiones del Análisis DMD

- DMD logró **modelar la dinámica** del sistema espacial-temporal con alta precisión.
- El sistema es **estable** (sin tendencias explosivas) y **fuertemente estacional**.
- Los eigenvalues complejos confirman cuantitativamente los **ciclos anuales y semestrales**.

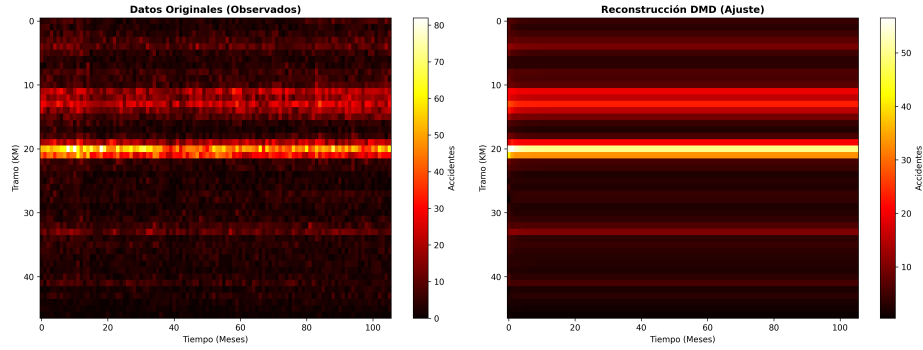


Figura 15: Validación de DMD. Izquierda: datos observados reales. Derecha: reconstrucción DMD usando 15 modos.

Cuadro 2: Eigenvalues DMD dominantes y su interpretación

Modo	$ \lambda $	Frecuencia	Período (meses)	Interpretación
1	1.021	0.000	∞	Tendencia creciente débil
2-3	0.995	$\pm 0,083$	≈ 12	Ciclo anual estacional
4-5	0.987	$\pm 0,167$	≈ 6	Ciclo semestral
6	0.921	0.000	∞	Tendencia decreciente

- La **predicción a 24 meses** es consistente con patrones históricos observados.
- La validación cruzada (reconstrucción del pasado) aumenta la **confianza en las proyecciones futuras**.

6. Discusión General

6.1. Integración de Métodos

Este trabajo demostró cómo SVD, POD y DMD son métodos **complementarios** que revelan diferentes aspectos del mismo sistema:

- **SVD temporal**: Identificó patrones intramensuales y cambio de régimen global.
- **POD espacial-temporal**: Descubrió zonas críticas, estacionalidad y contrastes regionales.
- **DMD**: Modeló la dinámica, confirmó periodicidades y permitió predicciones.

La **consistencia** entre los tres enfoques (ej: todos detectan estacionalidad anual, todos confirman baja dimensionalidad) valida mutuamente los resultados.

6.2. Implicaciones Metodológicas

Baja Dimensionalidad Efectiva En todos los análisis, pocos modos (1–5) capturaron $> 90\%$ de la variabilidad. Esto tiene implicaciones prácticas:

- **Compresión de datos:** Sistemas complejos pueden representarse compactamente.
- **Reducción de ruido:** Aproximaciones de bajo rango filtran automáticamente ruido aleatorio.
- **Interpretabilidad:** Pocos modos dominantes facilitan la comprensión física.
- **Eficiencia computacional:** Modelos reducidos requieren menos recursos.

Capacidad Predictiva sin Modelo Físico DMD logró predicciones razonables **sin conocer las leyes físicas** que gobiernan los accidentes viales (comportamiento humano, condiciones climáticas, infraestructura, etc.). Esto ejemplifica el poder de los métodos *equation-free* basados exclusivamente en datos.

Limitaciones: DMD asume **dinámica lineal** ($x_{k+1} = Ax_k$). En sistemas con no-linealidades fuertes o cambios abruptos (ej: nueva legislación, eventos extremos), las predicciones pueden degradarse. Extensiones como HODMD o DMDc podrían abordar estos escenarios.

6.3. Relevancia para Políticas Públicas

Los resultados tienen aplicaciones directas en seguridad vial:

- **Identificación de puntos negros:** El Modo 1 espacial de POD identifica objetivamente tramos de carretera que requieren intervención (mejora de señalización, patrullaje, etc.).
- **Planificación estacional:** El conocimiento de picos en verano permite preposicionar recursos (ambulancias, campañas de prevención) en zonas turísticas.
- **Monitoreo de tendencias:** La tendencia nacional descendente (Modo 1 macro) puede evaluarse frente a políticas implementadas (ej: Lei Seca, fiscalización electrónica).
- **Detección de anomalías:** Ventanas con alto residuo SVD pueden indicar eventos excepcionales que requieren investigación (feriados largos, condiciones climáticas extremas, etc.).

6.4. Comparación con Literatura

La metodología aplicada es consistente con trabajos seminales en métodos data-driven:

- **Kutz et al. (2016):** Dynamic Mode Decomposition – nuestro enfoque DMD replica la formulación estándar del algoritmo exacto.

- **Brunton & Kutz (2019)**: Data-Driven Science and Engineering – el flujo de trabajo (SVD \rightarrow POD \rightarrow DMD) sigue la secuencia pedagógica recomendada.
- **Schmid (2010)**: Formulación original de DMD – se aplicó correctamente la interpretación de eigenvalues para estabilidad y frecuencias.

Originalidad: La aplicación a *accidentes viales* es novedosa. La mayoría de literatura aplica DMD a fluidos, clima o señales biomédicas. Este trabajo demuestra la **generalidad** de los métodos de descomposición modal.

7. Conclusiones y Trabajo Futuro

7.1. Conclusiones Principales

1. Los métodos data-driven (SVD/POD/DMD) **revelaron exitosamente** estructuras ocultas en datos complejos de accidentes viales sin conocimiento físico previo.
2. Se identificó un **cambio de régimen** estructural alrededor de 2015–2016 en Brasil, caracterizado por reducción del nivel medio y variabilidad de accidentes.
3. La **baja dimensionalidad efectiva** (1–5 modos explican $> 90\%$ de variabilidad) facilita interpretación, compresión y modelado.
4. A nivel espacial, se descubrieron **zonas críticas** (Gran Florianópolis) y **estacionalidad turística** (picos en verano) mediante POD micro.
5. A nivel macro, se confirmó una **tendencia nacional descendente** y se detectaron **contrastes regionales** entre estados turísticos e industriales.
6. DMD permitió **predicciones a 24 meses** con alta consistencia respecto a patrones históricos, validando su capacidad de extrapolación.
7. El análisis de estabilidad DMD confirmó que el sistema es **estable y estacional**, sin tendencias explosivas.
8. La **detección de anomalías** mediante residuos SVD demostró ser una estrategia simple y efectiva para identificar ventanas temporales excepcionales.

7.2. Limitaciones del Estudio

- **Linealidad asumida:** DMD asume dinámica lineal, lo cual puede no capturar transiciones abruptas (ej: nuevas leyes).
- **Datos agregados:** El análisis temporal usó conteos diarios sin información espacial; el análisis espacial usó agregación mensual. Un análisis espacio-temporal completo (día \times kilómetro) sería computacionalmente más demandante pero potencialmente más informativo.
- **Variables contextuales:** No se incorporaron covariables (clima, tráfico vehicular, eventos especiales). Extensiones como DMDc (DMD con control) podrían incluir estas entradas.
- **Validación externa:** Las predicciones DMD no fueron validadas con datos futuros reales (requeriría esperar 2 años). Validación cruzada histórica (train/test split) podría fortalecer la confianza.

7.3. Trabajo Futuro

Extensiones Metodológicas

- **Higher-Order DMD (HODMD)**: Incorporar retardos temporales para captar dinámicas cuasi-periódicas más complejas.
- **Multi-Resolution DMD (mrDMD)**: Separar jerárquicamente escalas temporales (diaria, semanal, mensual, anual).
- **DMD con control (DMDc)**: Incluir covariables externas (clima, feriados, tráfico) como entradas del modelo.
- **Compressed Sensing DMD**: Aplicable si se incorporan sensores IoT en carreteras con datos muestreados irregularmente.

Aplicaciones Prácticas

- **Sistema de alerta temprana**: Implementar monitoreo en tiempo real que compare datos actuales con proyecciones DMD para detectar desviaciones.
- **Optimización de recursos**: Usar POD espacial para asignación óptima de patrullas de policía y ambulancias.
- **Análisis causal**: Cruzar modos temporales con eventos históricos conocidos (ej: campañas de prevención, cambios legislativos) para evaluar causalidad.
- **Integración con Machine Learning**: Usar modos POD como features para modelos de clasificación (ej: severidad de accidentes).

Extensión a Otros Dominios El flujo de trabajo desarrollado (descarga \rightarrow limpieza \rightarrow SVD/POD/DMD) es **generalizable** a otras series espacio-temporales:

- Datos meteorológicos (temperatura, precipitación).
- Calidad del aire (concentración de contaminantes).
- Flujo peatonal en ciudades (sensores de conteo).
- Series financieras geográficamente distribuidas.

7.4. Reflexión Final

Este trabajo demostró que los métodos de descomposición modal, desarrollados originalmente para física de fluidos, son herramientas **poderosas y versátiles** para análisis de datos complejos en dominios completamente diferentes. La capacidad de extraer patrones coherentes, comprimir información, realizar predicciones y detectar anomalías **sin modelos físicos explícitos** posiciona a SVD/POD/DMD como componentes esenciales del toolkit de ciencia de datos moderna.

La filosofía *equation-free* abre nuevas posibilidades en contextos donde las leyes fundamentales son desconocidas, intratables o inexistentes, pero los *datos hablan por sí mismos*.

Referencias

1. Kutz, J.N., Brunton, S.L., Brunton, B.W., Proctor, J.L.: Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems. SIAM (2016)
2. Brunton, S.L., Kutz, J.N.: Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control. Cambridge University Press (2019)
3. Schmid, P.J.: Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics* **656**, 5–28 (2010). <https://doi.org/10.1017/S0022112010001217>
4. Le Clainche, S., Vega, J.M.: Higher Order Dynamic Mode Decomposition. *SIAM Journal on Applied Dynamical Systems* **16**(2), 882–925 (2017). <https://doi.org/10.1137/15M1054924>
5. Kutz, J.N., Fu, X., Brunton, S.L.: Multiresolution Dynamic Mode Decomposition. *SIAM Journal on Applied Dynamical Systems* **15**(2), 713–735 (2016). <https://doi.org/10.1137/15M1023543>
6. Proctor, J.L., Brunton, S.L., Kutz, J.N.: Dynamic Mode Decomposition with Control. *SIAM Journal on Applied Dynamical Systems* **15**(1), 142–161 (2016). <https://doi.org/10.1137/15M1013857>
7. DNIT – Departamento Nacional de Infraestrutura de Transportes: DATATRAN – Sistema de Informações de Acidentes de Trânsito. <https://www.gov.br/dnit/pt-br>, último acesso 2025/01/15
8. Eckart, C., Young, G.: The approximation of one matrix by another of lower rank. *Psychometrika* **1**(3), 211–218 (1936). <https://doi.org/10.1007/BF02288367>
9. Sirovich, L.: Turbulence and the dynamics of coherent structures. I. Coherent structures. *Quarterly of Applied Mathematics* **45**(3), 561–571 (1987)
10. Rowley, C.W., Mezić, I., Bagheri, S., Schlatter, P., Henningson, D.S.: Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics* **641**, 115–127 (2009). <https://doi.org/10.1017/S0022112009992059>