VOLUMEN 48
XXXI Verano De la Ciencia
ISSN 2395-9797
www.jovenesenlaciencia.ugto.mx

Detección de desvíos de material fisible mediante monitoreo de antineutrinos de reactor: física, diseño de detector y calibración de un sistema CUSUM

Detection of fissile material diversion through reactor antineutrino monitoring: physics, detector design, and calibration of a CUSUM system

David Yves Ghislain Delepine[1], Orlando Ricardo Torres Ibarra[2] Victor Hugo Zamora Padilla[3]
[1,2,3]Universidad de Guanajuato delepine@ugto.mx, or.torresibarra@ugto.mx[1], vh.zamorapadilla@ugto.mx

Resumen

El monitoreo de antineutrinos de reactor es una técnica de salvaguardias nucleares que verifica la composición del combustible sin acceso físico al núcleo, aprovechando que los antineutrinos escapan de cualquier blindaje. Este trabajo desarrolla la física de producción y detección, simula un detector de referencia (50 m, ~1 tonelada) y calibra un sistema de detección de desvíos basado en CUSUM simple y espectral. Se diseñó un escenario adversarial de peor caso, una reducción del 15% en la fracción de combustible de Pu-239 combinada con un ajuste de potencia (1.8%) que anula la firma en la tasa total detectada. Una prueba de Neyman-Pearson, es decir una prueba optimizando la detección exacta del escenario previsto mostró que ese escenario deja apenas un Signal-to-Noise Ratio (SNR) de 1.3σ acumuladas en 240 días, lo que lo hace indiferenciable del ruido. Para poder separar la señal característica del escenario, se necesita un mínimo de SNR entre 4-5 σ. Frente a desviaciones sin esa compensación cruzada, el sistema detecta razonablemente bien. Acercar y agrandar el detector, de 50 a 20 metros y de 1 a 2.08 toneladas, eleva el techo teórico de detección al 90%, y una prueba de razón de verosimilitud generalizada (GLRT) sobre un banco de 36 ataques candidatos alcanza 85.5% sin necesidad de conocer el ataque exacto de antemano.

Palabras clave: antineutrinos de reactor; salvaguardias nucleares; no proliferación; CUSUM; espectro Huber-Mueller; GLRT.

1. Introducción: justificación y motivación

Verificar que un reactor nuclear no está desviando material fisible hacia un programa de armas depende hoy, sobre todo, de la contabilidad de material que declara el propio operador: balances de masa, cámaras, sellos y visitas periódicas del Organismo Internacional de Energía Atómica (OIEA). Son mecanismos que funcionan, pero que dependen en último término de accesos físicos programados y de la cooperación del Estado inspeccionado. El monitoreo de antineutrinos de reactor ofrece una vía complementaria: los antineutrinos que produce un reactor son un subproducto de la fisión, no hay blindaje conocido capaz de detenerlos, y su flujo total, junto con la forma de su espectro de energía, lleva codificada información sobre qué isótopo se está fisionando en cada momento (Bernstein et al., 2020). Un detector colocado a distancia razonable del núcleo, sin necesidad de entrar a la planta, puede en principio verificar de forma pasiva y continua que la composición del combustible declarado coincide con la real.

La magnitud que se busca detectar tiene una referencia concreta: el OIEA define la Cantidad Significativa (CS) de plutonio en aproximadamente 8 kg, la cantidad que se estima necesaria para no poder descartar la fabricación de un primer dispositivo nuclear explosivo (OIEA, 2002). Esa cifra fija la escala de sensibilidad y de tiempo que un sistema de monitoreo de antineutrinos necesita alcanzar para ser útil como herramienta de salvaguardias: no basta con detectar cualquier desvío, hay que detectarlo con margen suficiente antes de que se acumule 1 CS.

El monitoreo de antineutrinos tiene ventajas claras frente a los métodos tradicionales: no requiere acceso físico continuo, mide desde fuera del perímetro de la planta; los antineutrinos no pueden bloquearse ni falsificarse con blindaje, así que cualquier cambio real de composición deja una huella física; y es un monitoreo pasivo y continuo, en principio casi en tiempo real, que complementa a los inventarios periódicos en vez de sustituirlos. También tiene desventajas importantes. El flujo cae con el cuadrado de la distancia, así que el detector debe estar razonablemente cerca del reactor, decenas de metros, para acumular suficiente estadística. Los detectores con sensibilidad útil, del orden de toneladas, siguen siendo instalaciones no triviales de construir y operar. La señal es intrínsecamente estadística: un desvío pequeño solo se distingue del ruido acumulando muchos días de datos, así que el sistema es lento comparado con cámaras o sellos. Y, como muestra este trabajo, un operador que conoce el funcionamiento del sistema puede en principio camuflar un desvío de combustible ajustando otro parámetro, la potencia declarada, con precisión suficiente para cancelar la firma en el canal de detección más sensible.

2. El concepto físico

Mecanismo de producción

Cada fisión de U-235 o Pu-239 produce dos fragmentos ricos en neutrones que decaen beta sucesivamente hacia la estabilidad, y cada decaimiento beta libera un antineutrino electrónico. En promedio, cada fisión termina emitiendo entre 6 y 7 antineutrinos, aunque solo una fracción cae en la ventana de energía detectable (2 a 10 MeV en este trabajo).

La potencia térmica declarada P_{th} fija la tasa total de fisiones por segundo, F_{tot}. Si cada fisión de un isótopo i libera, en promedio, una energía E_i (202.5 MeV para U-235, 211.5 MeV para Pu-239), y $f_i(t)$ es la fracción de fisiones debida a ese isótopo en el día t, la energía media liberada por fisión es $\langle Ef\rangle(t)$, y la tasa total de fisiones queda determinada por el balance de energía:

$$\langle Ef\rangle(t) \;=\; f_{235}(t)\,E_{235} \;+\; f_{239}(t)\,E_{239}$$

$$F_{\mathrm{tot}}(t) \;=\; \frac{P_{\mathrm{th}}}{\langle Ef\rangle(t)}$$

El flujo de antineutrinos que llega al detector, a distancia d del núcleo, decae con el inverso del cuadrado de la distancia (emisión isótropa) y conserva la forma espectral por fisión $S_i(E)$ de cada isótopo:

$$\varphi(E) \;=\; \frac{F_{\mathrm{tot}}(t)\,\left(f_{235}(t)\,S_{235}(E) \;+\; f_{239}(t)\,S_{239}(E)\right)}{4\pi d^2}$$

Métodos de detección

El canal estándar de detección es el decaimiento beta inverso (IBD), sobre los protones libres de un blanco rico en hidrógeno:

$$\bar{\nu}_{\mathrm{e}} \;+\; p \;\rightarrow\; e^{+} \;+\; n$$

con umbral de 1.8 MeV y una firma en coincidencia retardada: primero la aniquilación del positrón (señal "prompt"), y microsegundos después la captura del neutrón térmico en un núcleo dopante como gadolinio o litio. La sección eficaz de la reacción, cerca del umbral, se aproxima por:

$$\sigma(E_\nu) \;\approx\; \kappa\,E_{\mathrm{e}}\,p_{\mathrm{e}},\;\; E_{\mathrm{e}} \;=\; E_\nu \;-\; \Delta$$

donde Δ ≈ 1.293 MeV es la diferencia de masa neutrón-protón más la masa del positrón, pe es el momento del positrón y κ es una constante que fija las unidades (9.52×10^{-44} cm²/MeV²). El número de eventos esperado por día es proporcional al número de protones libres del blanco n_p, al flujo y a la sección eficaz, integrados sobre energía:

$$N(t) \;=\; n_{\mathrm{p}}\,\Delta t \int_{E_{\mathrm{min}}}^{E_{\mathrm{max}}} \varphi(E,t)\,\sigma(E)\,dE$$

VOLUMEN 48
XXXI Verano De la Ciencia
ISSN 2395-9797
www.jovenesenlaciencia.ugto.mx

El espectro de Huber-Mueller

Es la parametrización estándar del espectro de antineutrinos por fisión de cada isótopo combustible, obtenida al convertir mediciones acumuladas de espectros beta de productos de fisión a espectros de antineutrinos (Mueller et al., 2011; Huber, 2011). Se usa la forma exponencial de polinomio habitual:

$$S_{\text{iso}}(E) \;=\; \exp\left(\sum_{\text{i=0}}^{5} a_i E^i\right)$$

con coeficientes a_i ajustados de forma independiente para U-235 y Pu-239. El espectro de Pu-239 es sistemáticamente más "duro" que el de U-235, es decir, favorece energías más altas (figura 1); esa diferencia de forma es la base tanto del CUSUM espectral como de la prueba de Neyman-Pearson descritos en la sección 5.

El modelo de Huber-Mueller no es perfecto: experimentos de alta estadística (Daya Bay, RENO, Double Chooz) observaron un exceso de antineutrinos entre 4 y 6 MeV respecto a la predicción, un rasgo conocido como el "bulto de 5 MeV" (An et al., 2016), con significancia local de hasta 4.4σ en Daya Bay. Como la amplitud del bulto parece variar con el quemado del combustible, se especuló que podría servir como una firma adicional para distinguir U-235 de Pu-239. Análisis posteriores de mayor estadística (comparando NEOS y Daya Bay) desfavorecen a Pu-239 y Pu-241 como fuente única del bulto y apuntan más bien a U-235 o a incertidumbres en las bases de datos nucleares (Huber, 2017); su origen sigue sin resolverse, y por eso este trabajo no lo usa como canal de discriminación adicional, aunque sí ilustra que el espectro medido contiene más estructura de la que captura el modelo de Huber-Mueller por sí solo.

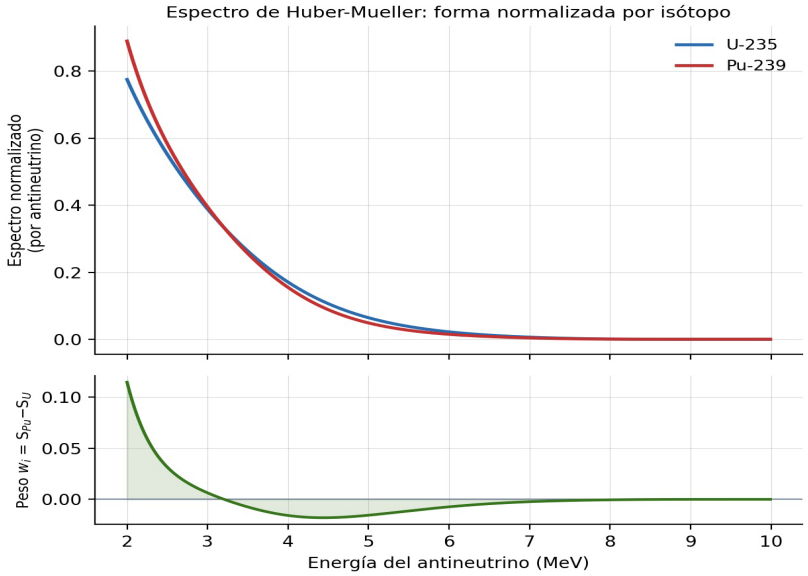


Figura 1. Arriba: forma normalizada del espectro de Huber-Mueller para U-235 y Pu-239 (calculada con los coeficientes usados en la simulación). Abajo: el peso espectral $w_i = S_{^{239}Pu}(E_i) - S_{^{235}U}(E_i)$ que usa el CUSUM espectral, la diferencia de forma entre ambos isótopos en función de la energía.

Evolución temporal en un ciclo de operación

A lo largo de un ciclo de operación, 540 días en este trabajo, típico de un reactor de agua a presión (PWR), la fracción de fisiones debida a cada isótopo cambia por el quemado del combustible, aproximada por una relajación exponencial hacia el equilibrio:

$$f_{235}(t) \;=\; f_{235,\text{BOC}}\, e^{-\lambda_5 t}$$

con $f_{235,BOC} \approx 0.65$ la fracción de U-235 al inicio del ciclo (BOC), relajándose hasta $f_{235,EOC} \approx 0.40$ al final (EOC); el Pu-239, generado por captura neutrónica en U-238, crece de forma complementaria hasta una

fracción de equilibrio cercana a 0.60, valores consistentes con estudios de isotopía de combustible gastado y con mediciones directas del flujo de antineutrinos a lo largo de varios ciclos de operación de reactores PWR comerciales (Bowden et al., 2009). Esa evolución lenta y suave es exactamente el tipo de cambio "declarado" que el sistema de monitoreo debe distinguir de un cambio "no declarado": ambos modifican el espectro observado con el tiempo, pero solo uno es consistente con la trayectoria de quemado normal reportada por el operador.

3. Diseño del detector

El detector simulado sigue una arquitectura de capas concéntricas simplificada, común en detectores de antineutrinos de campo cercano: un núcleo central de centelleador líquido orgánico, rico en hidrógeno y de densidad aproximada 900 kg/m³, con un lado de referencia de 1 m (aproximadamente 0.9 toneladas); una capa de blindaje pasivo de plomo de 5 cm para atenuar la radiación gamma ambiental; y una capa exterior de polietileno de 15 cm para moderar y absorber neutrones rápidos de fondo. El número de protones libres disponibles como blanco de la reacción IBD se estima a partir de la masa del núcleo $M_{\text{núcleo}}$asumiendo una composición tipo CH2 (dos protones libres por cada 14 g de material):

$$n_{\text{p}} \;=\; M_{\text{núcleo}} \left(\frac{N_{\text{A}}}{14}\right) \times 2$$

con N_{A} el número de Avogadro. El fondo instrumental se modela como la suma de varias contribuciones típicas de un sitio de campo cercano, rayos cósmicos, radiactividad ambiental y fondo correlacionado con el reactor entre otras, que suman alrededor de 190 cuentas por día sin atenuar, y se atenúan exponencialmente al atravesar las capas de plomo y polietileno según su coeficiente de atenuación característico μ:

$$B \;=\; B_0 \, e^{-\mu_{\text{Pb}} x_{\text{Pb}}} \, e^{-\mu_{\text{PE}} x_{\text{PE}}}$$

La configuración de referencia usada en la mayor parte de este trabajo es 50 m de distancia al núcleo del reactor y aproximadamente 1 tonelada de masa de detector, una escala deliberadamente conservadora frente a detectores de campo cercano ya construidos, como PROSPECT (aproximadamente 4 t a 7-9 m del núcleo de un reactor de investigación; Ashenfelter et al., 2019) o Nucifer (aproximadamente 0.8 t a 7 m; Boireau et al., 2016).

El número de antineutrinos detectados en un día dado es el resultado de contar eventos raros e independientes entre sí: de los aproximadamente 10^{19}-10^{20} antineutrinos que el núcleo emite por segundo, solo una fracción minúscula e independiente entre partículas interactúa por IBD en el volumen del detector, y lo mismo ocurre con cada cuenta de fondo (un decaimiento cósmico o ambiental aislado). Formalmente, esto es exactamente la condición del teorema del límite de Poisson: un número muy grande de ensayos independientes (N fisiones, o N fotones de fondo posibles), cada uno con una probabilidad de éxito muy pequeña (p, la probabilidad de que ESA fisión en particular produzca un evento detectado), con el producto N_p convergiendo a un valor finito λ, converge a una distribución de Poisson de parámetro λ, sin importar la forma detallada de p. Es la misma justificación estadística detrás de todo conteo de decaimientos radiactivos o de partículas. En este trabajo, λ se fija, para cada día t y cada bin de energía, igual al número de eventos esperado $N_{esp}(t)$ que entrega el modelo físico de la sección 2 (la ecuación de eventos esperados, integrando flujo, sección eficaz y protones libres); ese es el valor que la simulación usa como media de la distribución de Poisson de la que se extrae cada observación sintética, y es también la referencia $N_{esp}(t)$ contra la que el CUSUM compara cada observación real en la ecuación siguiente:

$$P(n;\,\lambda) \;=\; \frac{\lambda^n \, e^{-\lambda}}{n!}, \;\; \text{Var}(n) \;=\; \text{E}(n) \;=\; \lambda$$

Esa igualdad entre varianza y media es la que justifica, en cada canal del CUSUM, normalizar el residuo diario dividiendo entre la raíz cuadrada del número esperado en vez de ajustar una desviación estándar por separado: en un proceso de Poisson, esa raíz es, por construcción, la desviación estándar teórica. La figura 2 muestra una simulación de la tasa total diaria en un ciclo limpio, sin ningún desvío: la curva esperada declina de forma suave por el quemado natural del combustible, mientras que la serie simulada con ruido de Poisson, una sola realización Monte Carlo, fluctúa alrededor de ella.

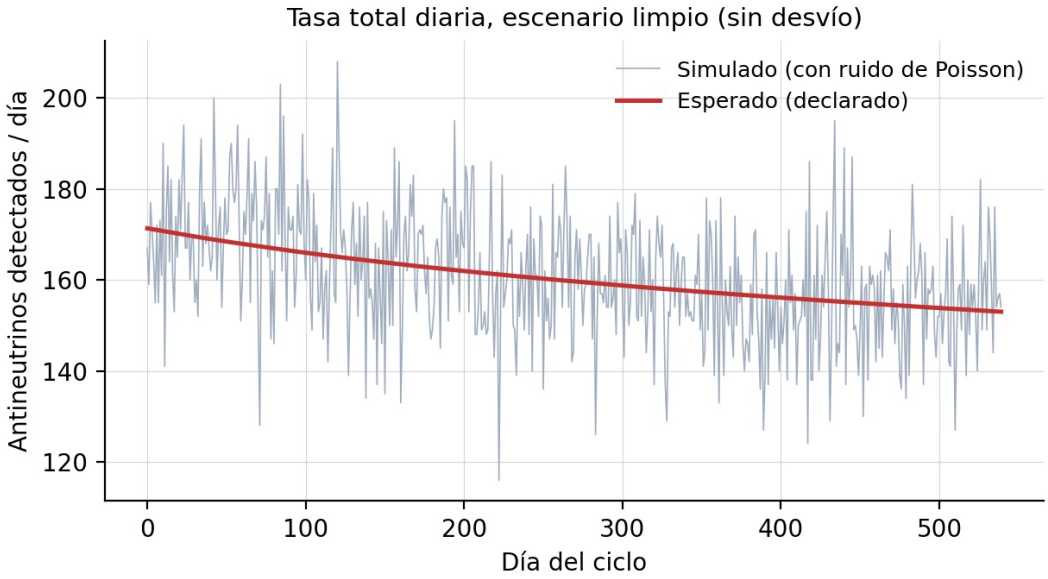


Figura 2. Tasa total diaria de antineutrinos detectados en un ciclo limpio, sin ningún desvío, configuración de referencia (50 m, 1 tonelada). La curva esperada declina de forma suave por el quemado natural del combustible; la serie simulada con ruido de Poisson (una sola realización Monte Carlo) fluctúa alrededor de ella.

4. Escenarios de robo posibles

El escenario de peor caso

El escenario más exigente que se probó combina dos ajustes simultáneos: una reducción del 15% en la fracción de Pu-239 ($x = 0.15$) a partir del día 300 de un ciclo de 540, y una reducción δ de la potencia real respecto a la declarada, calibrada con precisión para dejar la tasa TOTAL de antineutrinos detectados sin cambios. Para entender por qué existe ese δ y de dónde sale su valor (1.8%), conviene escribir la tasa total detectada explícitamente. Integrando la ecuación de eventos esperados de la sección 2 sobre energía, y usando $F_{tot} = P_{th}/\langle Ef \rangle$, la tasa total detectada un día dado puede escribirse como el producto de la potencia declarada por un factor que depende solo de la mezcla de combustible:

donde C agrupa las constantes de geometría y detector ($n_p, \Delta t, 4\pi d^2$), y $I_{\sigma,iso} = \int s_{iso}(E)\sigma(E)dE$ es el número de eventos detectables por fisión de cada isótopo, integrando el espectro crudo (no normalizado) contra la sección eficaz de IBD. El desvío calibrado impone que la tasa real, con $f_{239} \to f_{239}(1-x)$ y $P_{th} \to P_{th}(1-\delta)$, sea idéntica a la declarada:

$$1 - \delta = \frac{f_{235}E_{235} + f_{239}(1-x)E_{239}}{f_{235}E_{235} + f_{239}E_{239}} \times \frac{f_{235}I_{\sigma,235} + f_{239}I_{\sigma,239}}{f_{235}I_{\sigma,235} + f_{239}(1-x)I_{\sigma,239}}$$

Esta es exactamente la ecuación que resuelve numéricamente el código ($I_{\sigma,235}$ e $I_{\sigma,239}$ salen de integrar el espectro de Huber-Mueller contra la sección eficaz de IBD). Evaluándola en el día 300, con $f_{235}(300) \approx 0.50$, $f_{239}(300) \approx 0.54$, $E_{235} = 202.5\ MeV$, $E_{239} = 211.5$ MeV, e integrando numéricamente $I_{\sigma,235}$ e $I_{\sigma,239}$ sobre 2-10 MeV, obtenemos $\delta \approx 1.8\%$.

Los dos factores tienen una lectura física directa. El primero es menor que 1 porque el Pu-239 libera más energía por fisión que el U-235 ($E_{239} > E_{235}$): al retirar Pu-239 sin tocar la potencia, se necesitarían más fisiones por segundo para sostener la misma potencia declarada, un efecto que por sí solo subiría la tasa

detectada en torno a un 8.7% (el inverso de 0.920). El segundo factor es mayor que 1 porque, en este modelo, el Pu-239 produce menos eventos detectables por fisión que el U-235 ($I_{\sigma,239} < I_{\sigma,235}$): perder Pu-239 reduce el rendimiento promedio de eventos por fisión del combustible restante, un efecto que por sí solo bajaría la tasa detectada en torno a un 6.3%. Los dos efectos casi se cancelan, dejando un residuo neto de apenas +1.8-1.9% en la tasa total si no se toca la potencia; ese residuo es precisamente lo que la reducción de potencia $\delta \approx 1.8\%$ cancela. La figura 3 muestra la tasa total diaria en este escenario: pese al desvío que comienza en el día 300, no hay ningún salto visible en la curva simulada.

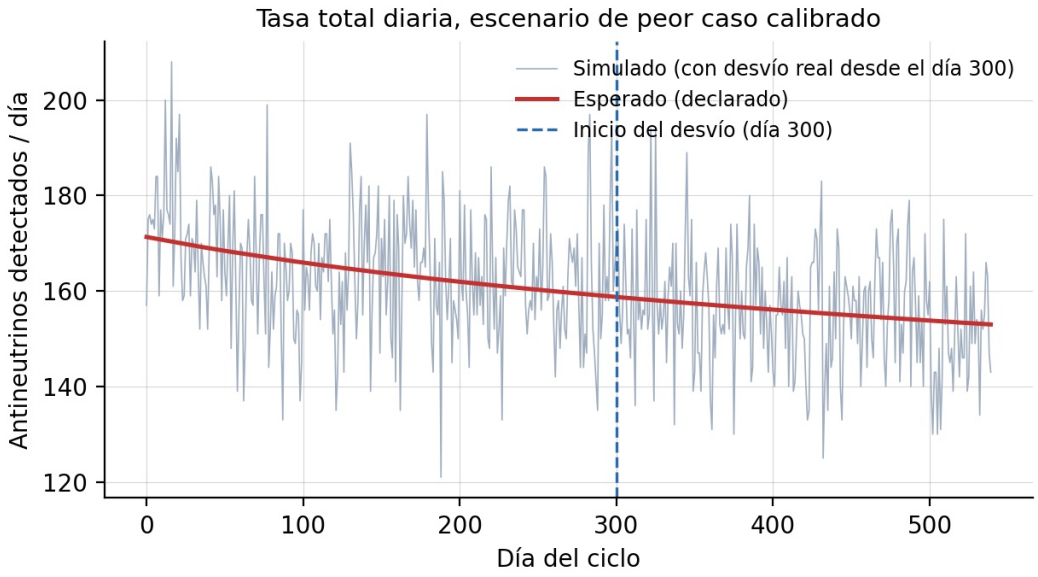


Figura 3. Tasa total diaria de antineutrinos en el escenario de peor caso calibrado. El desvío de combustible comienza en el día 300 (línea vertical), pero el ajuste de potencia que lo acompaña deja la curva esperada declarada prácticamente indistinguible de la real.

Dos escenarios de contraste, sin compensación cruzada

Para comparar, se probaron dos escenarios más simples que varían un solo parámetro sin ningún ajuste compensatorio: una discrepancia de potencia aislada ($\delta \neq 0$, $x = 0$), que se refleja íntegramente en la tasa total; y un desvío de Pu-239 aislado, sin compensar la potencia ($x \neq 0$, $\delta = 0$), que según la ecuación anterior también deja una firma visible en el canal simple, ya que sin el ajuste de potencia los dos efectos (energía y rendimiento) no se cancelan. Ambos escenarios representan a un adversario menos sofisticado, o simplemente un error operativo o una mala calibración de instrumentos, y sus resultados de detección se presentan junto con los del peor caso en la sección 6.

5. Herramientas estadísticas

CUSUM simple

El CUSUM (suma acumulada) formaliza la idea de que un solo día con una diferencia entre lo esperado y lo observado no dice mucho, porque el conteo de eventos fluctúa de forma natural según Poisson, pero si esa diferencia se acumula día tras día en la misma dirección, algo sistemático está ocurriendo (Page, 1954). Cada día se calcula un residuo estandarizado:

$$z_t = \frac{N_{\text{obs}}(t) - N_{\text{esp}}(t)}{\sqrt{N_{\text{esp}}(t)}}$$

y ese residuo se acumula en dos estadísticos, uno para detectar un aumento sostenido y otro para una disminución:

ISSN 2395-9797
www.jovenesenlaciencia.ugto.mx

$$S_t^+ = \max(0, S_{t-1}^+ + z_t - k)$$
$$S_t^- = \min(0, S_{t-1}^- + z_t + k)$$

El parámetro k es una tolerancia que evita que el ruido normal dispare la alarma, típicamente medio sigma; cuando S^+ supera un umbral h o S^- cae por debajo de $-h$, se declara una desviación real. Ni k ni h se fijan a mano: ambos se calibran, junto con los días de persistencia, mediante la búsqueda en rejilla de la sección 6 (reactor_cusum.py, sección 8), buscando en cada caso el par (k, h) que minimiza la tasa de falsas alarmas en corridas limpias sujeto a mantener el mayor poder de detección posible frente al desvío de prueba; un h demasiado bajo dispara con cualquier fluctuación de Poisson, y uno demasiado alto no detecta nada, así que el criterio no es minimizar h ni las falsas alarmas por separado, sino el compromiso entre ambos.

CUSUM espectral

Un CUSUM que solo mira la tasa total tiene un punto ciego, precisamente el que explota el escenario de peor caso de la sección 4: si el reactor quema más Pu-239 del declarado, algunos intervalos de energía reciben más eventos de los esperados y otros reciben menos, porque el espectro de Pu-239 no tiene la misma forma que el de U-235 (figura 1). El CUSUM espectral pondera cada intervalo de energía por un peso:

$$w_i = S_{\mathrm{Pu239}}(E_i) - S_{\mathrm{U235}}(E_i)$$

la diferencia entre los espectros normalizados de cada isótopo, es decir, ya divididos entre su propia integral para que la comparación capture solo la forma y no la amplitud total. El residuo espectral estandarizado en cada intervalo I (grupo de bins fusionados, sección 3) es:

$$Z_I(t) = \frac{\sum_{i \in I} w_i (x_i(t) - N_i(t))}{\sqrt{\sum_{i \in I} w_i^2 \, N_i(t)}}$$

Ponderar así equivale a un filtro acoplado: más peso a los intervalos de energía donde la composición del combustible deja una huella más clara, y menos peso a lo que solo refleja un error de normalización general, como una potencia mal declarada.

La prueba de Neyman-Pearson

Cuando se conoce con exactitud tanto la hipótesis nula H0 (nada raro pasa) como la alternativa H1 (un ataque específico y completamente conocido), el lema de Neyman-Pearson garantiza que la prueba más poderosa posible, el que maximiza la probabilidad de detección para una tasa de falsa alarma fija, es la razón de verosimilitud entre ambas hipótesis (Neyman & Pearson, 1933):

$$\Lambda(x) = \frac{P(x \mid H_1)}{P(x \mid H_0)}$$

Para conteos de Poisson independientes por bin de energía, con media N_i bajo H_0 y $N_i + \mu_i$ bajo H_1, el logaritmo de la razón de verosimilitud tiene forma cerrada:

$$ln\,(\Lambda) = \sum_i x_i \ln\left(\frac{N_i + \mu_i}{N_i}\right) - \mu_i$$

El lema garantiza rechazar H_0 cuando $\ln(\Lambda)$ supera un umbral, elegido para fijar la tasa de falsa alarma deseada α. En este trabajo se usó como cota superior teórica: se calculó $\mu_i(t)$ exacto para el ataque calibrado de la sección 4 en cada bin de energía fino, ponderado por su propia varianza de Poisson (versión gaussiana de la ecuación anterior). A diferencia del CUSUM espectral, que usa un peso genérico (la diferencia de forma U-235/Pu-239), la prueba de Neyman-Pearson usa el peso específico de ESTE ataque, así que ningún otro método basado solo en los datos de antineutrinos puede superarlo. Sirve como referencia teórica: si ni siquiera la prueba óptima detecta el peor caso con margen razonable, ninguna mejora puramente estadística lo hará tampoco.

6. Resultados de la simulación Monte Carlo

Todos los resultados que siguen usan 200 semillas de Monte Carlo por configuración, tanto para el escenario limpio (calibración de falsas alarmas) como para el escenario con desvío (poder de detección), suficientes para que la incertidumbre estadística de cada estimación quede acotada a unos pocos puntos porcentuales.

Contra el escenario de peor caso, la prueba de Neyman-Pearson acumula apenas 1.3σ de señal en los 240 días restantes del ciclo, muy por debajo de las 3 a 5σ que suele exigirse para una detección confiable. Al correr el motor de CUSUM completo, la detección neta, la fracción de corridas con desvío que se detectan, descontando lo que aparecería por puro azar en corridas limpias en la misma ventana, fue de aproximadamente 0% en todas las combinaciones de parámetros probadas (tabla 1), incluido la prueba teóricamente óptima.

***Tabla 1**. Detección contra el escenario de peor caso, 200 semillas de Monte Carlo por configuración*

Estrategia	Falsas/ciclo	Detección neta	Retardo
Solo CUSUM simple	—	0% (en el ruido)	—
Solo CUSUM espectral (por intervalo)	0.88	2.5% ± 3.4 pp	102 días
Solo filtro acoplado continuo	0.16	0% ± 1.9 pp	—
OR (mejor con sus propios parámetros)	0.97	0% ± 3.4 pp	—
AND simple + 1 espectral	0.88	2.5% ± 3.4 pp	102 días
Test óptimo de Neyman-Pearson (límite teórico)	—	~0%	—

Ese resultado no significa que el sistema sea inútil: los dos escenarios de contraste de la sección 4, sin la compensación cruzada que define al peor caso, sí se detectan razonablemente bien y de forma creciente con la magnitud del desvío (tabla 2 y figura 4). El punto ciego del sistema no es general, es una combinación muy específica y difícil de lograr para un adversario: un ajuste de potencia calibrado con precisión de décimas de punto porcentual, coordinado exactamente con la magnitud del desvío de combustible.

***Tabla 2**. Detección neta contra discrepancias sin compensar, por magnitud (200 semillas por magnitud)*

Escenario	Magnitud	Detección neta	Retardo
Discrepancia de potencia sola	1.0%	2.5%	118 días
Discrepancia de potencia sola	2.0%	14.5%	120 días
Discrepancia de potencia sola	3.0%	44.0%	108 días
Discrepancia de potencia sola	5.0%	93.0%	60 días
Desvío de Pu-239 sin compensar	10%	10.5%	111 días
Desvío de Pu-239 sin compensar	15%	22.5%	106 días

Desvío de Pu-239 sin compensar	25%	62.0%	102 días
Desvío de Pu-239 sin compensar	35%	94.0%	61 días

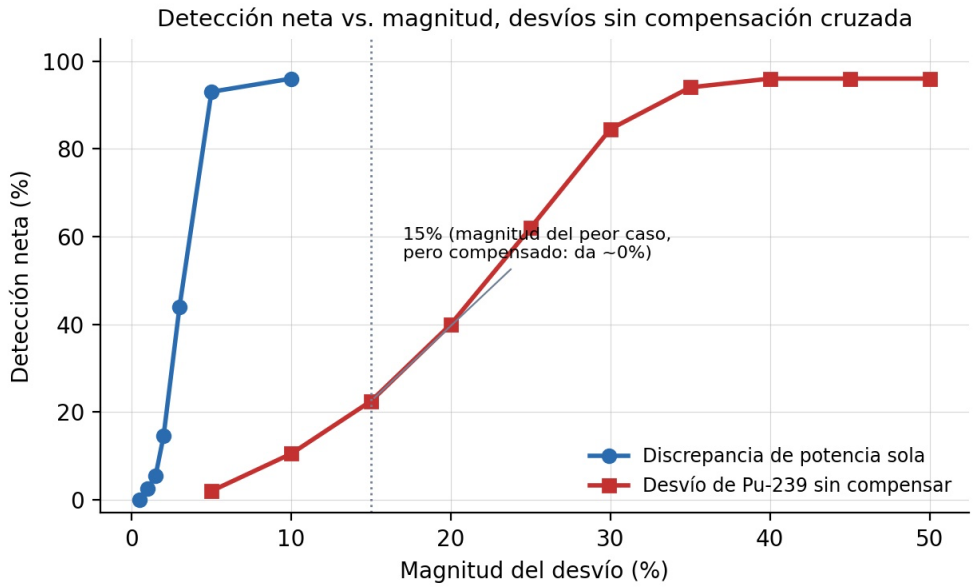


Figura 4. Detección neta en función de la magnitud del desvío, para cada ingrediente del peor caso probado por separado. El mismo 15% de desvío de Pu-239 que resulta casi indetectable cuando se compensa con precisión la potencia (figura 3) da 22.5% de detección neta cuando no se compensa nada.

7. Cómo mejorar la sensibilidad del detector

Mejoras de hardware: distancia y masa

Si el límite frente al peor caso es de estadística y no de método, cabe preguntarse cuánta mejora de hardware se necesita para superarlo. El número de protones libres del blanco escala linealmente con la masa del detector M (ecuación de la sección 3), y el flujo escala como el inverso del cuadrado de la distancia d (ecuación de la sección 2), así que el número de eventos esperado por bin, N_i, escala como:

$$N_i \propto \frac{M}{d^2}$$

La firma de un ataque específico es, en cada bin, una fracción fija δ_i del conteo esperado ($\mu_i = \delta_i N_i$), una propiedad de la física del combustible que no depende del detector. Como la varianza de Poisson es igual a la media ($var = N_i$), al construir la razón señal-ruido acumulada, $SNR^2 = \sum_i \mu_i^2/N_i = \sum_i {\delta_i}^2 N_i$, sobrevive una sola potencia de N_i, ni cero ni dos, así que:

$$SNR^2 = C\,\frac{M}{d^2}$$

con C una constante que depende solo de la física del reactor y del ataque, no del detector. Esta relación implica que acercar el detector rinde más por unidad de esfuerzo que agrandarlo: doblar la masa solo dobla el SNR^2, pero acercarse a la mitad de la distancia lo multiplica por 4. Una configuración de 20 m y 2.08 toneladas, comparable a detectores de campo cercano ya construidos como PROSPECT, alcanza 5.02σ de señal acumulada y eleva el techo teórico de detección del peor caso de prácticamente 0% a 90% (figura 5).

ISSN 2395-9797
www.jovenesenlaciencia.ugto.mx

El algoritmo de fusión de canales usado en la sección 6, sin embargo, solo captura 11.5% de esa mejora si no se actualiza junto con el hardware, lo que motiva las mejoras de software que siguen.

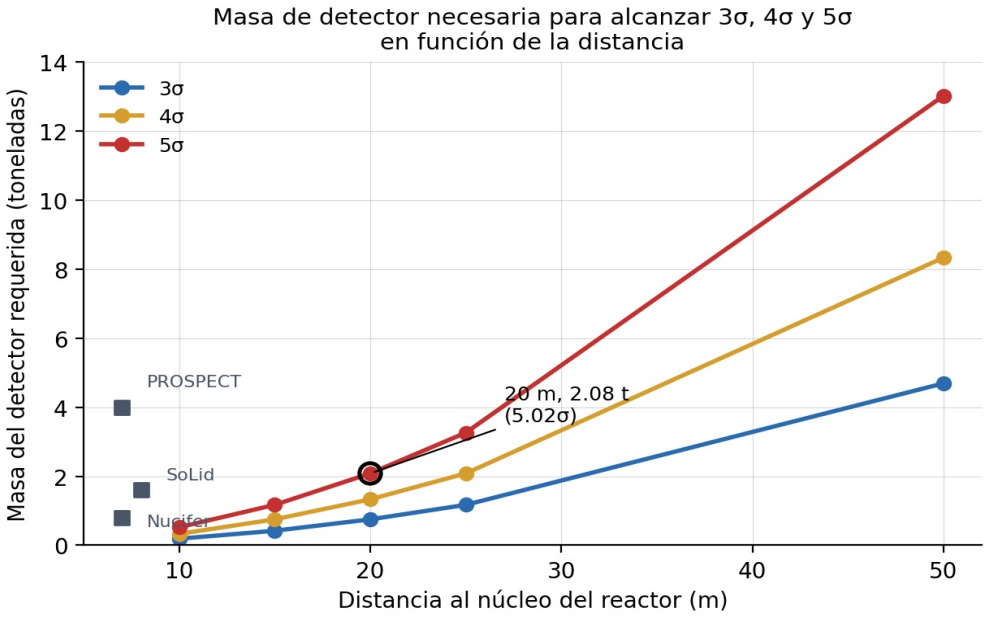


Figura 5. Masa de detector necesaria para alcanzar una señal de 3σ, 4σ y 5σ en función de la distancia al núcleo del reactor. El punto marcado, 20 m y 2.08 t, alcanza 5.02σ, una escala comparable a detectores de campo cercano ya construidos como PROSPECT, Nucifer y SoLid.

Mejoras de software: análisis estadístico

Del lado del análisis, dos mejoras concretas ayudan a desvíos menos extremos que el peor caso calibrado, aunque no cierran su límite de información. La primera reemplaza la aproximación gaussiana del z-score, ecuación de la sección 5, por la razón de verosimilitud de Poisson exacta de la sección 5, evaluada de forma secuencial día a día como una prueba de razón de probabilidad secuencial (SPRT; Wald, 1945), en su variante para conteos de Poisson (Lucas, 1985). La ventaja es que no se degrada cuando el conteo esperado es bajo, justamente el régimen del intervalo de mayor energía donde la aproximación gaussiana pierde precisión.

La segunda es un filtro acoplado continuo: en vez de fusionar el espectro en 3 o 4 intervalos anchos (sección 3), se usa el peso óptimo punto a punto en energía, $w_i = S_{^{239}Pu}(E_i) - S_{^{235}U}(E_i)$ evaluado en cada bin fino de 0.01 MeV en vez de promediado por intervalo. Esto evita cancelar señal real cuando el peso espectral cambia de signo dentro de un mismo intervalo ancho (figura 1, panel inferior), algo que ocurre justo alrededor de 4-5 MeV.

Ese bin de 0.01 MeV es una malla numérica de integración, no la resolución real de una medición: ningún detector reconstruye la energía del antineutrino con precisión arbitraria. La energía depositada por el positrón se infiere a partir de la luz de centelleo colectada, y esa reconstrucción tiene una resolución finita $\sigma_E(E)$, que se parametriza típicamente como la suma en cuadratura de un término estocástico, la fluctuación en el número de fotones detectados, dominante a baja energía, y un término constante, por no uniformidad en la colección de luz y calibración:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b$$

con a ≈ 3-5% MeV$^{1/2}$ y b ≈ 1-2%, valores típicos de un centelleador líquido dopado con gadolinio como los usados en Nucifer o PROSPECT. En términos absolutos, esto equivale a una resolución de varios cientos de

keV en el rango de 2 a 10 MeV, muy por encima de los 0.01 MeV de la malla numérica. En la práctica, el espectro que mide cualquier detector real es una convolución del espectro verdadero con esa función de resolución, así que dos bins más finos que $\sigma_E(E)$ no aportan información independiente entre sí: el ancho de bin útil está limitado por la resolución del detector, no por la resolución que se elija en la simulación. El filtro acoplado continúo descrito aquí representa entonces el límite ideal de resolución infinita; aplicarlo con un detector real exigiría, como paso adicional, convolucionar el peso w_i con la función de resolución antes de usarlo, lo que en la práctica ensancha y suaviza el cambio de signo del peso alrededor de 4-5 MeV.

La prueba de razón de verosimilitud generalizada (GLRT)

La prueba de Neyman-Pearson de la sección 5 necesita conocer el desplazamiento exacto $\mu_i(t)$ que produciría un ataque específico, una suposición poco realista en un despliegue real, donde no se sabe de antemano qué fracción de Pu-239 desviaría un adversario ni con qué precisión compensaría la potencia. La prueba de razón de verosimilitud generalizada (GLRT; Kay, 1998) releva esa exigencia construyendo un banco de C candidatos, cada uno una combinación (fracción de Pu-239 desviada, ajuste de potencia) en una rejilla, y calculando cada día la razón de verosimilitud contra todos ellos a la vez.

Para cada candidato c, el logaritmo de la razón de verosimilitud exacta (Poisson) por bin es:

$$LLR_i(t) = x_i(t)\ln\left(\frac{N_{c,\mathrm{i}}}{N_{\mathrm{decl},i}}\right) - \left(N_{c,\mathrm{i}} - N_{\mathrm{decl},i}\right)$$

Cuando el desplazamiento es pequeño frente al conteo declarado, esta expresión se reduce, por expansión de Taylor, a la versión gaussiana ya usada en el CUSUM espectral y la prueba de Neyman-Pearson, normalizada para tener varianza 1 bajo H_0:

$$z_c(t) = \frac{\sum_i \mu_{c,\mathrm{i}} \dfrac{x_i - N_{\mathrm{decl},i}}{N_{\mathrm{decl},i}}}{\sqrt{\sum_i \dfrac{\mu^2}{N_{c,\mathrm{i}}}}}$$

y cada día se toma el máximo sobre los C candidatos, sin necesitar saber cuál es el correcto de antemano:

$$z_{\mathrm{GLRT}}(t) = \max_c z_c(t)$$

El resultado, un solo número por día que ya combina información de los 36 candidatos, se acumula exactamente igual que cualquier otro canal en el motor de CUSUM de la sección 5, con el mismo rearmado y la misma lógica de persistencia. El diagrama de flujo de la figura 6 resume el procedimiento completo tal como se implementó en este trabajo, con el banco de 36 candidatos (6 valores de fracción de Pu-239 × 6 valores de ajuste de potencia, incluyendo el ataque real de la sección 4 sin marcarlo especial).

VOLUMEN 48
XXXI Verano De la Ciencia
ISSN 2395-9797
www.jovenesenlaciencia.ugto.mx

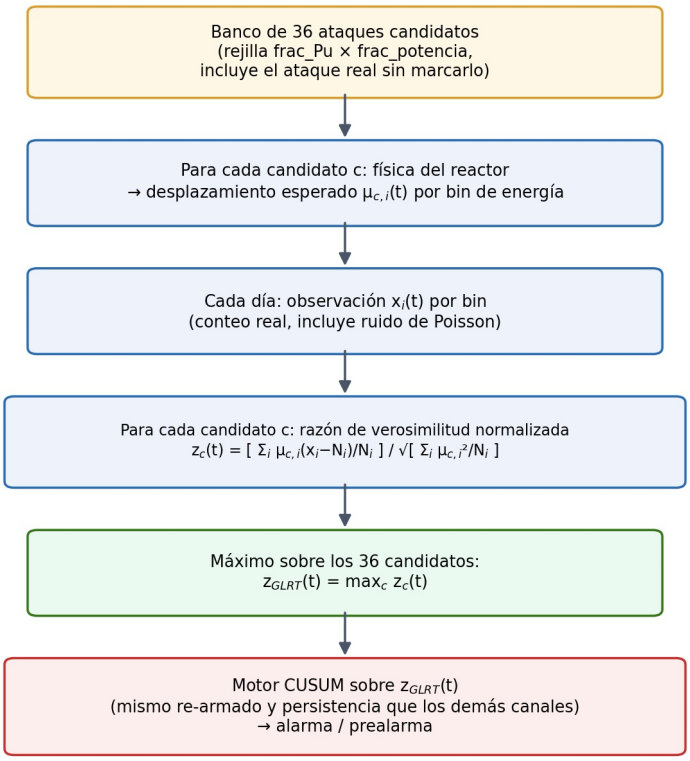


Figura 6. Diagrama de flujo de la prueba GLRT implementado. El banco de 36 ataques candidatos se construye una sola vez a partir de la física del reactor; cada día se calcula la razón de verosimilitud normalizada $z_c(t)$ contra los 36 candidatos y se toma el máximo, que alimenta el mismo motor de CUSUM que los demás canales.

Con la geometría actual (50 m) el GLRT casi duplica al filtro genérico que no conoce el ataque, pero ambos siguen lejos del límite teórico; con la geometría óptima (20 m, 2.08 t) el GLRT alcanza 85.5% de detección neta, a solo 4.5 puntos del techo de 90% (tabla 3). El costo es un umbral de alarma más alto para controlar la tasa de falsa alarma al escanear 36 hipótesis a la vez, el equivalente estadístico de buscar en muchos lugares al mismo tiempo, el ajuste habitual por comparaciones múltiples.

***Tabla 3**. Detección neta de la prueba GLRT (36 candidatos) frente a otros algoritmos, 200 semillas por configuración*

Configuración	Algoritmo	Detección neta
50 m, 1× masa (actual)	Filtro genérico (no conoce el ataque)	6.5%
50 m, 1× masa (actual)	GLRT (36 candidatos)	12.5%
50 m, 1× masa (actual)	Óptimo (conoce el ataque exacto)	7.5%
20 m, 2.08 t (óptima)	Filtro genérico (no conoce el ataque)	29.5%
20 m, 2.08 t (óptima)	GLRT (36 candidatos)	85.5%

20 m, 2.08 t (óptima)	Óptimo (conoce el ataque exacto)	90.0%

8. Descripción de los scripts de simulación

Todo el trabajo numérico de este artículo se implementó en cinco scripts de Python, todos construidos sobre el mismo módulo de física del reactor para que sus resultados sean directamente comparables entre sí. Esta sección resume qué hace cada uno.

reactor_cusum.py — módulo base: física, detector y motor CUSUM

Es el módulo del que dependen los otros cuatro scripts (se importa como "base"). Contiene la física del reactor completa: el espectro de Huber-Mueller por isótopo, la sección eficaz de IBD, la evolución temporal de f235(t) y f239(t) a lo largo del ciclo, y la construcción del detector por capas (núcleo, plomo, polietileno) descrita en la sección 3. A partir de ahí calcula el número de eventos esperado por día y por bin de energía, muestrea observaciones Poisson y calcula los z-scores del CUSUM simple y del CUSUM espectral (con la fusión automática de intervalos de baja estadística de la sección 3). Incluye también calibrar_frac_potencia_peor_caso(), la función que resuelve numéricamente la ecuación del peor caso de la sección 4, y correr_cusum(), el motor de CUSUM con rearmado tras cada evento que usan todos los demás scripts. Su función main() ejecuta una búsqueda en rejilla sobre (k, h, persistencia, factor de prealerta) calibrada contra el escenario de peor caso, y escribe grid_search_resultados.csv (336 combinaciones evaluadas), alarmas_prealarmas.csv (el registro día a día de cada alarma y prealarma) y las series diarias de ejemplo serie_diaria_baseline.csv y serie_diaria_con_robo.csv usadas en las figuras 2 y 3.

reactor_cusum_fusion.py — reglas de fusión de canales

Evalúa cinco formas distintas de combinar el CUSUM simple con los CUSUM espectrales en una sola alerta global: OR (cualquier canal dispara), AND con al menos 1 o 2 canales espectrales de acuerdo además del simple, y votación de 2 o 3 canales cualesquiera, cada una con distintas ventanas de coincidencia (0, 3, 5 o 10 días). Para cada combinación de parámetros extrae, por semilla, la lista de días en que cada canal dispara (reutilizando el rearmado de reactor_cusum.py) y agrupa eventos cercanos en el tiempo entre canales para decidir si cumplen la regla de fusión. Calcula la detección neta restando la tasa de alertas que aparecería por puro azar en corridas limpias dentro de la misma ventana, la corrección metodológica que evitó confundir falsas alarmas coincidentes con detección real. Escribe fusion_comparacion.csv, con 6,720 combinaciones de parámetros y regla de fusión evaluadas.

reactor_cusum_contraste.py — escenarios sin compensación cruzada

Prueba los dos ingredientes del peor caso por separado: una discrepancia de potencia sin ningún desvío de combustible, y un desvío de Pu-239 sin ningún ajuste de potencia (sección 4). Primero calibra los parámetros del CUSUM simple contra un objetivo de falsa alarma del 5% usando corridas limpias, y con esos mismos parámetros evalúa la detección neta y el retardo promedio, con 200 semillas por magnitud, para una rejilla de magnitudes de cada escenario. Escribe discrepancia_potencia_sola.csv y robo_pu_sin_compensar.csv, la fuente de los datos de la tabla 2 y la figura 4.

reactor_cusum_glrt.py — test de razón de verosimilitud generalizada

Implementa el GLRT de la sección 7: construye un banco de 36 ataques candidatos variando la fracción de Pu-239 desviada (6 valores) y el ajuste de potencia (6 valores) en una rejilla, sin marcar cuál de ellos es el ataque real. Para cada día calcula la razón de verosimilitud normalizada contra los 36 candidatos a la vez ($z_c(t)$ de la sección 7) y se queda con el máximo, que alimenta el mismo motor de CUSUM de reactor_cusum.py. El script permite reconstruir la física completa para una distancia y masa de detector arbitrarias, y en su bloque principal corre el GLRT dos veces: una con la geometría actual (50 m, 1×) y otra con la geometría óptima (20 m, 2.08 t), escribiendo resultados_glrt.csv con la comparación que resume la tabla 3.

reactor_cusum_algoritmos_mejorados.py — comparación acumulativa de mejoras de algoritmo

Implementa y compara, contra el mismo escenario de peor caso, las tres mejoras de análisis estadístico de la sección 7: el filtro acoplado continuo genérico (no necesita conocer el ataque), el CUSUM de razón de verosimilitud de Poisson exacta (sin la aproximación gaussiana, evaluado por intervalo espectral), y el estadístico de Neyman-Pearson gaussiano que ya fusiona de forma óptima tasa total y forma espectral (el límite teórico superior conociendo el ataque exacto). Cada mejora se evalúa de forma acumulativa frente a la línea base práctica ya implementada en reactor_cusum_fusion.py, y todo el procedimiento se repite para la geometría actual (50 m) y la óptima (20 m, 2.08 t), escribiendo comparacion_mejoras_algoritmo_actual.csv y comparacion_mejoras_algoritmo_optima.csv.

***Tabla 4**. Scripts de simulación y los archivos que generan*

Script	Función principal	Archivos generados
reactor_cusum.py	Física del reactor, detector, motor CUSUM, búsqueda en rejilla	grid_search_resultados.csv, alarmas_prealarmas.csv, serie_diaria_baseline.csv, serie_diaria_con_robo.csv
reactor_cusum_fusion.py	Reglas de fusión de canales simple + espectral	fusion_comparacion.csv
reactor_cusum_contraste.py	Escenarios sin compensación cruzada	discrepancia_potencia_sola.csv, robo_pu_sin_compensar.csv
reactor_cusum_glrt.py	Test de razón de verosimilitud generalizada (36 candidatos)	resultados_glrt.csv
reactor_cusum_algoritmos_mejorados.py	Comparación acumulativa de mejoras de algoritmo	comparacion_mejoras_algoritmo_actual.csv, comparacion_mejoras_algoritmo_optima.csv

9. Conclusión

El monitoreo de antineutrinos de reactor es una herramienta de salvaguardias físicamente sólida: aprovecha una señal que no puede bloquearse ni falsificarse, y que codifica directamente la composición del combustible que se está fisionando. Con la configuración de referencia (50 m, 1 tonelada), el sistema calibrado en este trabajo funciona bien contra desvíos de combustible o discrepancias de potencia que no estén deliberadamente coordinados entre sí, con una probabilidad de detección que crece de forma natural con la magnitud del desvío. Contra un adversario que sí coordina ambos ajustes con precisión, usando exactamente la ecuación de balance de la sección 4, el monitoreo de antineutrinos por sí solo no basta: la señal acumulada del peor caso, 1.3σ en 240 días, es un límite de información y no de algoritmo, y ninguna mejora puramente estadística lo cierra. Cerrar esa brecha exige combinar tres cosas: hardware más sensible, acercar y agrandar el detector según $SNR^2 \propto M/d^2$; un algoritmo que no dependa de conocer el ataque de antemano, el GLRT recupera 85.5% de un techo teórico de 90% con la geometría óptima; y fuentes de información independientes del propio flujo de antineutrinos, como el balance eléctrico o térmico de la planta, que obligarían a un eventual adversario a engañar dos sistemas de medición a la vez y no solo uno. Frente a la meta de 8 kg de plutonio que define una Cantidad Significativa, un sistema así, combinado con esas fuentes independientes, ofrece una capa adicional de verificación continua que ningún mecanismo de inspección periódica puede igualar por sí solo.

Bibliografía/Referencias

An, F. P., et al. (Daya Bay Collaboration). (2016). Measurement of the reactor antineutrino flux and spectrum at Daya Bay. Physical Review Letters, 116(6), 061801. https://doi.org/10.1103/PhysRevLett.116.061801

Ashenfelter, J., et al. (PROSPECT Collaboration). (2019). The PROSPECT reactor antineutrino experiment. Nuclear Instruments and Methods in Physics Research A, 922, 287–309. https://doi.org/10.1016/j.nima.2018.12.079

Bernstein, A., Bowden, N., Goldblum, B. L., Huber, P., Jovanovic, I., & Mattingly, J. (2020). Colloquium: Neutrino detectors as tools for nuclear security. Reviews of Modern Physics, 92(1), 011003. https://doi.org/10.1103/RevModPhys.92.011003

Boireau, G., et al. (Nucifer Collaboration). (2016). Online monitoring of the Osiris reactor with the Nucifer neutrino detector. Physical Review D, 93(11), 112006. https://doi.org/10.1103/PhysRevD.93.112006

Bowden, N. S., Bernstein, A., Dazeley, S., Svoboda, R., Misner, A., & Palmer, T. (2009). Observation of the isotopic evolution of pressurized water reactor fuel using an antineutrino detector. Journal of Applied Physics, 105(6), 064902. https://doi.org/10.1063/1.3080251

Huber, P. (2011). Determination of antineutrino spectra from nuclear reactors. Physical Review C, 84(2), 024617. https://doi.org/10.1103/PhysRevC.84.024617

Huber, P. (2017). NEOS data and the origin of the 5 MeV bump in the reactor antineutrino spectrum. Physical Review Letters, 118(4), 042502. https://doi.org/10.1103/PhysRevLett.118.042502

International Atomic Energy Agency. (2002). IAEA safeguards glossary (International Nuclear Verification Series No. 3). IAEA. https://www.iaea.org/sites/default/files/iaea_safeguards_glossary.pdf

Kay, S. M. (1998). Fundamentals of statistical signal processing, Volume II: Detection theory. Prentice Hall.

Lucas, J. M. (1985). Counted data CUSUM's. Technometrics, 27(2), 129–144. https://doi.org/10.1080/00401706.1985.10488030

Mueller, T. A., Lhuillier, D., Fallot, M., Letourneau, A., Cormon, S., Fechner, M., Giot, L., Lasserre, T., Martino, J., Mention, G., Porta, A., & Yermia, F. (2011). Improved predictions of reactor antineutrino spectra. Physical Review C, 83(5), 054615. https://doi.org/10.1103/PhysRevC.83.054615

Neyman, J., & Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. Philosophical Transactions of the Royal Society of London A, 231(694–706), 289–337. https://doi.org/10.1098/rsta.1933.0009

Page, E. S. (1954). Continuous inspection schemes. Biometrika, 41(1–2), 100–115. https://doi.org/10.1093/biomet/41.1-2.100

Vogel, P., & Beacom, J. F. (1999). Angular distribution of neutron inverse beta decay. Physical Review D, 60(5), 053003. https://doi.org/10.1103/PhysRevD.60.053003

Wald, A. (1945). Sequential tests of statistical hypotheses. Annals of Mathematical Statistics, 16(2), 117–186. https://doi.org/10.1214/aoms/1177731118