

# Filtro de Honestidad para Precursores Sísmicos

Protocolo E-Veto aplicado a sistemas críticos de infraestructura

Documento Operativo-Canónico

Versión cerrada

## Resumen (1 página)

Este constructo define un **Filtro de Honestidad** para sistemas críticos: un protocolo que **prioriza la credibilidad** y reduce afirmaciones no fundadas. El filtro se organiza en tres capas obligatorias:

**Detección causal** ( $t_C$ )  $\Rightarrow$  **Veto entrópico** ( $\Delta H$ )  $\Rightarrow$  **Persistencia coherencial** ( $\int \Delta H$ )

y se cierra con un **Estado Límite Epistemológico** cuando el ruido domina.

**Principio rector (infraestructura crítica):**

SILENCIO > FALSA ALARMA > FALSO NEGATIVO
--

Esto significa: el sistema **prefiere callar** antes que erosionar confianza.

## 1 I. Arquitectura del Filtro

### 1.1 I.1 Objetivo principal

**Minimizar desconfianza suprimiendo afirmaciones no fundadas**, no maximizar detecciones.

### 1.2 I.2 Alcance

El filtro aplica a cualquier pipeline multi-señal (sísmico y no sísmico) siempre que:

- existan señales primarias (candidatas a precursor),
- existan señales de control (ruido/regímenes conocidos),
- pueda definirse un conjunto de control para calibración regional.

### 1.3 I.3 Glosario mínimo

$t_C$	Tiempo causal (detección de cambio de régimen)
$t_\Sigma$	Tiempo coherencial (validación por persistencia)
$H(\cdot)$	Entropía de Shannon empírica sobre ventana
$\Delta H$	Cambio de entropía entre ventanas (post - pre)
$\epsilon_Q$	Umbral de empuje causal (derivadas)
$\epsilon_{\min}$	Umbral mínimo de reducción entrópica
$\Theta_\Sigma$	Umbral de persistencia (integral coherencial)
$\varphi$	Fricción informacional (ruido efectivo)
$\mathcal{L}_{\max}$	Estado Límite Epistemológico (ruido dominante)

## 2 II. Señales: primarias vs. control

### 2.1 II.1 Señales primarias (candidatas a precursor)

Ejemplos típicos (la lista es extensible por región y disponibilidad):

- $V_p/V_s$  (razón P/S)
- $b$ -value (pendiente Gutenberg–Richter)
- Radón (Rn-222) en pozos
- Strain / deformación (GNSS, inclinómetros)
- VLF/ULF (perturbaciones electromagnéticas)
- Ionosfera (anomalías TEC)

### 2.2 II.2 Señales de control (ruido/regímenes conocidos)

- variaciones térmicas diurnas,
- presión barométrica,
- mareas terrestres,
- actividad antropogénica (tráfico, industria).

### 2.3 II.3 Regla canónica de independencia

Toda confirmación multi-señal debe usar canales lo más independientes posible (en instrumentación, física, y fuente de datos). Esto reduce correlaciones espurias.

## 3 III. Protocolo E–Veto (triple compuerta)

### 3.1 III.1 Capa 1: Detección de empuje causal en $t_C$ (derivadas)

Para cada señal  $s(t)$  se busca un cambio de régimen en ventana causal. Criterio canónico:

$$\left| \frac{d\Sigma}{dt} \right| > \epsilon_Q \quad \text{o} \quad \frac{dH(t)}{dt} < -\epsilon_Q$$

donde  $H(t)$  es la entropía empírica estimada en una ventana centrada en  $t$ .

**Calibración de  $\epsilon_Q$**  Se define sobre control regional (sin eventos objetivo, o bajo régimen base):

$$\epsilon_Q = k_Q \cdot \sigma_{\text{null}} \quad (k_Q \approx 3)$$

donde  $\sigma_{\text{null}}$  es la dispersión de la derivada bajo hipótesis nula.

### 3.2 III.2 Capa 2: Veto entrópico (E-Veto) en $t_\Sigma$

Se define entropía de Shannon empírica:

$$H(X) = - \sum_i p_i \log_2(p_i)$$

y cambio entrópico:

$$\Delta H(s) = H_{\text{post}}(s) - H_{\text{pre}}(s)$$

**Condición E-Veto (negativa, obligatoria)**

$$\Delta H \geq -\epsilon_{\min} \Rightarrow \text{RECHAZAR}$$

Interpretación: si la señal no reduce incertidumbre más allá del mínimo significativo, se considera **ruido estructurado** o transitorio no concluyente.

**Calibración de  $\epsilon_{\min}$  (portable por dominio)** Se estima  $\Delta H$  bajo hipótesis nula con bootstrap sobre control:

$$\epsilon_{\min} = k \cdot \sigma_{H_{\text{null}}} \quad (k \in [2, 3])$$

**Regla:** la calibración es por región-síñal-canal-régimen (evita trasladar umbrales entre contextos distintos).

### 3.3 III.3 Capa 3: Persistencia coherencial (integral)

Un candidato no se confirma por un único  $\Delta H$ , sino por su persistencia en ventana  $t_\Sigma$ :

$$\int_{t_0}^{t_1} \Delta H(t) dt < -\Theta_\Sigma \Rightarrow \text{CANDIDATO VÁLIDO}$$

$\Theta_\Sigma$  se calibra por región y nivel de alerta.

## 4 IV. Matriz de Decisión Honesta (Estados)

### 4.1 IV.1 Definición de estados

Estado	Condición	Acción	Comunicación
$L_0$ Nominal	$Q < \epsilon_Q$	Monitoreo pasivo	Silencio operacional
$L_1$ Detección	$Q \geq \epsilon_Q$ , evaluando $\Delta H$	Buffer interno + registro	<b>No comunicar</b>
$L_2$ Candidato	$\Delta H < -\epsilon_{\min}$ , validando $\int \Delta H$	Análisis multi-síñal	Alerta interna (si aplica)

Estado	Condición	Acción	Comunicación
$L_3$ Confirmado	$\int \Delta H < -\Theta_\Sigma$ en múltiples señales	Protocolo de acción	Comunicar <b>con IC</b> y FPR
$\mathcal{L}_{\max}$ Límite	$\varphi > \varphi_{\text{crit}}$ (ruido domina)	Suspender inferencia	“Condiciones no concluyentes”

## 4.2 IV.2 Regla de comunicación

**Jamás** emitir lenguaje categórico. Solo: estado, IC, horizonte, y FPR histórico.

# 5 V. Anti-falsos positivos: regla multi-señal

## 5.1 V.1 Regla dura

Un precursor es comunicable solo si:

$$\boxed{\# \text{señales independientes confirmadas} \geq 3}$$

La cifra 3 es un umbral operativo conservador; puede elevarse en infraestructura crítica.

## 5.2 V.2 Administración explícita de falsos positivos

Se registra en cada reporte:

- `veto_reason` (motivo exacto),
- `thresholds` (umbrales usados),
- `control_signature` (identificador del conjunto de control),
- `FPR_expected` (histórico del patrón/estado).

# 6 VI. Calibración regional (plantilla)

## 6.1 VI.1 Tabla ejemplo de parámetros

Región	$\epsilon_Q$ (en $\sigma$ )	$\epsilon_{\min}$ (bits)	$\Theta_\Sigma$ (bits·día)	$t_\Sigma$ (días)
Baja sismicidad	3.5	0.8	12	7
Moderada	3.0	0.6	10	5
Alta (subducción)	2.5	0.5	8	3

## 6.2 VI.2 Protocolo mínimo de calibración

1. Recolectar  $\geq 2$  años de datos de control (sin eventos objetivo o bajo régimen base).
2. Estimar  $H_{\text{null}}$  y  $\sigma_{H_{\text{null}}}$  por señal/canal/régimen.
3. Ejecutar bootstrap ( $\geq 10,000$  simulaciones) para estabilizar percentiles.
4. Ajustar umbrales a un objetivo operativo de FPR (p.ej.  $< 0.5\%$  en  $L_3$ ).

## 7 VII. Comunicación con confianza (formato canónico)

### 7.1 VII.1 Ejemplo de mensaje correcto

“Anomalía confirmada en 4/6 señales monitoreadas. Reducción entrópica (media):  $\Delta H = -0.8 \pm 0.2$  bits. IC95%:  $[-1.1, -0.5]$ . Estado:  $L_2$  (candidato en validación). Horizonte operativo: 3–7 días. Acción: incrementar frecuencia de monitoreo. Nota: tasa histórica de falsos positivos para este patrón: 12%.”

### 7.2 VII.2 Mensajes prohibidos

- “Detectado precursor” (sin IC/FPR/estado).
- “Terremoto probable en 48h” (lenguaje categórico).
- “Confirmado” sin multi-señal y sin persistencia integral.

## 8 VIII. Implementación técnica (plantilla de referencia)

### 8.1 VIII.1 Nota

El siguiente código es **referencial**. La implementación real debe:

- registrar cada decisión con trazabilidad (parámetros/umbrales),
- separar detección (derivadas) de validación (integrales),
- soportar modo límite  $\mathcal{L}_{\max}$ .

### 8.2 VIII.2 Pseudocódigo mínimo

```
# --- Capa 1: Empuje causal (tC) ---
def detectar_empuje(x, kQ=3.0):
    dx = diff(x) # derivada discreta
    epsQ = kQ * std(dx_null) # calibrado con control
    return abs(dx) > epsQ, dx

# --- Capa 2: E-Veto (Delta H) ---
def delta_H(pre, post, bins=50):
    H_pre = shannon_empirical(pre, bins=bins)
    H_post = shannon_empirical(post, bins=bins)
    return H_post - H_pre, H_pre, H_post

def veto_entropico(dH, eps_min):
    # RECHAZAR si no hay reduccion suficiente
    return (dH < -eps_min)

# --- Capa 3: Persistencia (integral) ---
def persistencia(signal, t0, tSigma, eps_min):
    # integra dH en subventanas dentro de tSigma
    integral = 0.0
    for w in subwindows(signal[t0:t0+tSigma]):
        dH_,_ = delta_H(w[:len(w)//2], w[len(w)//2:])
        integral += dH
```

```

return integral

# --- Estado limite epistemolgico ---
def estado_limite(SNR_series, phi_crit):
    phi_total = sum(1.0/snr for snr in SNR_series)
    return phi_total > phi_crit

```

## 9 IX. Checklist de honestidad operacional

Antes de comunicar cualquier anomalía:

- ☐ Empuje detectado en  $t_C$  (derivadas) y calibración válida.
- ☐  $\Delta H < -\epsilon_{\min}$  con  $\epsilon_{\min}$  regional.
- ☐ Persistencia:  $\int \Delta H < -\Theta_{\Sigma}$ .
- ☐ Confirmación  $\geq 3$  señales independientes.
- ☐ Control de causas conocidas (antropogénicas/meteorológicas).
- ☐ Intervalos de confianza (IC) reportados.
- ☐ FPR histórico del patrón reportado.
- ☐ Lenguaje no categórico (estado + IC + horizonte + FPR).

## 10 X. Gestión del Estado Límite Epistemológico $\mathcal{L}_{\max}$

### 10.1 X.1 Criterio

Se declara  $\mathcal{L}_{\max}$  cuando el ruido domina la inferencia:

$$\varphi > \varphi_{\text{crit}}$$

Operacionalmente,  $\varphi$  puede estimarse como función de degradación SNR o deriva estadística multi-canal. En  $\mathcal{L}_{\max}$ :

- se suspende escalamiento inferencial,
- se continúa observación y recalibración,
- se comunica: “condiciones no concluyentes”.

## 11 XI. Métricas de credibilidad (KPIs)

1.  $\text{FPR}_{\text{real}} = \frac{\text{alarmas sin evento subsecuente}}{\text{total alarmas}}$
2.  $\text{TPR} = \frac{\text{eventos con precursor detectado}}{\text{total eventos objetivo}}$
3. Lead time:  $LT = t(\text{evento}) - t(\text{comunicación})$
4. Confianza operativa (encuesta/ejercicio): % que actuaría ante un  $L_3$ .

## 12 XII. Clausura

Nuestro trabajo no es “predecir”.  
Nuestro trabajo es no mentir sobre lo que sabemos.