

Nota técnica — Revisión de datos del tablero de riesgo (Tuxilo)

Fecha: 2026-06-20 · **Motivo:** observación de auditoría sobre el origen del "85.0%" en la alerta principal y revisión de la variable embalse. **Estado:** dos hallazgos corregidos; un punto abierto para la revisión de datos de la semana.

HALLAZGO 1 — EL "ESCENARIO DE REFERENCIA A 30 DÍAS = 85.0%" ERA UN VALOR ARTIFICIAL

Qué decía el tablero. En la alerta principal: "El sistema embalsado agregado se ubica en 74.4% y el escenario de referencia a 30 días apunta a 85.0%."

De dónde salía (código `_build_storage_series`).

- Leía el dataset SIMEM `8d3ccd` (proyección embalse agregado de corto plazo), que **solo cubre ~12 días** (no 30).
 - Para los 30 días NO usaba una proyección oficial: calculaba la **pendiente de los últimos 5 días y la extrapolaba linealmente**, con un **recorte duro** `max(0.45, min(0.85, ...))`.
 - El valor caía contra el techo, por lo que **"85.0%" era exactamente el tope del recorte (0.85)**, no un resultado del modelo.
- Riesgo.** Se presentaba como "escenario de referencia a 30 días" algo que (a) extrapolaba más allá del horizonte real del dato y (b) estaba acotado artificialmente, produciendo un número redondo engañoso.

Corrección aplicada.

- Se eliminó el recorte a 0.85/0.45 (queda solo la cota física [0,1]).
- La serie pasa a usar el **embalse observado de XM** (ver Hallazgo 2) y la proyección a 30 días se rotula explícitamente como **extrapolación de la pendiente observada**, no como proyección oficial a ese plazo.
- Texto nuevo: "El embalse agregado observado (XM) está en X%; la tendencia reciente extrapolada a 30 días lo ubica en Y% (extrapolación de la pendiente observada, no proyección oficial a ese plazo)."

HALLAZGO 2 — LA VARIABLE "EMBALSE" TENÍA CUATRO VALORES SEGÚN LA FUENTE

En una misma fecha coexistían cuatro cifras de "embalse agregado":

Fuente	Valor	Base
<code>data_snapshot.json</code> (lo usaba el índice)	~63.6%	base distinta / desactualizada (2026-06-10)
Dataset diario XM <code>VoluUtilDiarEner/CapaUtilDiarEner</code> (Sistema)	~73.4%	% útil energético nacional (métrica estándar XM)
Maestro de 23 embalses (suma de energía)	~78.2%	agregación por embalse
SIMEM <code>8d3ccd</code> (proyección corto plazo)	~75.6%	caso de planeación SIMEM

Comprobación clave: el mismo 2026-06-10, el dataset daba 73.75% y el snapshot 63.6% → la diferencia **no es por fecha sino por base de cálculo**.

Corrección aplicada.

- Se unificó la fuente del **hero/alerta** y del **índice integrado** a una sola: la serie observada XM `VoluUtilDiarEner/CapaUtilDiarEner` (Sistema), ~73%. Antes el índice leía el snapshot (63.6%) y el hero leía SIMEM (75.6%).
- Resultado: el tablero ya muestra **un solo número de embalse coherente** (~73%) en hero, alerta e índice.

Impacto en el índice. El índice integrado pasó de 68 a **66 (Naranja)**: con embalse ~73% el factor hídrico aporta 0 puntos (nivel sano) en vez de 2. **Robustez:** tanto con 73% como con 78% el factor da 0 puntos (umbral $\geq 70\%$), por lo que la cifra del índice **no depende** de cuál de esas dos bases se elija.

PUNTO ABIERTO PARA LA REVISIÓN DE DATOS (ESTA SEMANA)

Queda por **confirmar la base oficial del % de embalse** entre las dos métricas vigentes de XM:

- Sistema** `VoluUtilDiarEner/CapaUtilDiarEner` (~73.4%) — la que se dejó como fuente única.
- Suma de los 23 embalses** por energía del maestro (~78.2%).

Ambas son de XM y deberían coincidir; la diferencia (~5 pp) sugiere distinto universo de embalses o tratamiento de agregados. **Acción sugerida:** cotejar contra el reporte oficial de XM "% volumen útil del SIN" y fijar la métrica canónica. No afecta el nivel del índice (ver arriba), pero sí la cifra que se reporta como dato.

ARCHIVOS MODIFICADOS

- `scripts/tuxilo_render_dashboard.py` → `_build_storage_series` (fuente observada + sin recorte 0.85), texto de la alerta, escenarios sin recorte mágico.
- `tuxilo/sintesis/riesgo_integrado.py` → embalse desde la serie observada XM (no el snapshot), fuente declarada.

Todo es reproducible (`python3 scripts/tuxilo_sintesis.py` y el render). Cifras leídas en vivo.