

OPTIMIZACIÓN DE LA GESTIÓN DE DUCTOS MEDIANTE VALIDACIÓN ESTADÍSTICA AS-RUN DE HERRAMIENTAS ILI

Autor: Santiago Parodi - Alan Cornejo – Emanuel Benitez

Empresa: Transportadora de Gas del Norte S.A. (TGN)

Email: santiago.parodi@tgn.com.ar, alan.cornejo@tgn.com.ar, emanuel.benitez@tgn.com.ar

SINOPSIS

La norma API 1163, "In-line Inspection Systems Qualification", establece un marco riguroso para la calificación de sistemas de inspección en línea (ILI), siendo la validación de resultados un pilar fundamental para la toma de decisiones en los programas de integridad de cañerías. Dentro de esta norma, la transición de una validación Nivel 2 a una de Nivel 3 representa una evolución significativa, pasando de un enfoque de simple verificación a una caracterización estadística robusta del desempeño real de la herramienta.

Una validación de Nivel 2 consiste en una verificación que compara un número limitado de mediciones de campo con la especificación de desempeño previamente declarada por el proveedor del servicio. Su principal objetivo es determinar, con un alto grado de confianza, si el desempeño de la herramienta cumple con lo especificado, permitiendo así rechazar una corrida de inspección si no cumple con el estándar prometido. Sin embargo, este nivel no permite cuantificar el desempeño real de la corrida; simplemente la contrasta contra un umbral predefinido.

En contraste, la validación de Nivel 3 introduce un análisis estadístico más profundo que utiliza mediciones de campo extensivas para estimar el desempeño real "as-run" de la herramienta. La principal ventaja de este enfoque es que, en lugar de solo aceptar o rechazar una especificación genérica, genera una especificación de desempeño propia y específica para esa inspección en particular. Esto permite al operador conocer con precisión la exactitud y la incertidumbre de los datos reportados, como el error medio y la desviación estándar en el dimensionamiento de las anomalías.

En este trabajo se detallará su implementación y resultados para el sistema de gasoductos de TGN, su capacidad para refinar significativamente los cálculos de integridad y optimizar las decisiones de gestión de ductos. Al estimar estadísticamente el desempeño real de la herramienta en una corrida específica ("as-run"), esta validación proporciona una medida cuantitativa y defendible de la incertidumbre de los datos. Este conocimiento es fundamental para las evaluaciones de ingeniería, como los cálculos de presión de falla y aptitud para el servicio, ya que permite análisis más precisos al utilizar la exactitud demostrada en lugar de las especificaciones genéricas del proveedor. Esta precisión se traduce directamente en una toma de decisiones más eficiente. Su aplicación es especialmente crucial en cañerías ubicadas en áreas sensibles, donde un entendimiento preciso del riesgo es imperativo para garantizar la integridad del activo.

INTRODUCCIÓN

La integridad de los sistemas de transporte de hidrocarburos por ductos es una disciplina que ha evolucionado desde el mantenimiento correctivo hacia modelos predictivos de gestión de riesgos. En este

paradigma, el uso de herramientas de inspección interna (ILI, por sus siglas en inglés) es uno de los principales métodos que permiten conocer el estado de integridad mecánica del activo. Sin embargo, la información provista por estas no es infalible; está sujeta a incertidumbres derivadas de la física de medición, la velocidad de la herramienta, y las propiedades electromagnéticas o acústicas del material base. La industria, tradicionalmente, ha gestionado esta incertidumbre mediante la aceptación de las especificaciones de desempeño declaradas por el proveedor, asumiendo que reflejan fielmente la realidad de la inspección. Esta práctica, aunque administrativamente conveniente, introduce un sesgo que puede distorsionar los cálculos de Aptitud para el Servicio (Fitness-for-Service o FFS) y la probabilidad de falla en análisis de riesgos.

El estándar **API 1163 (3ra Edición, 2021)**, junto con las guías de implementación desarrolladas por el Pipeline Research Council International (PRCI), específicamente el documento **PR-719-223803-R01**, proporcionan la arquitectura metodológica para validar objetivamente el desempeño de las herramientas ILI. A pesar de la disponibilidad de estos estándares, gran parte de los operadores limitan sus análisis a lo que la norma define como Validación de Nivel 2: una prueba binaria de cumplimiento. Si bien este nivel asegura que la herramienta no ha fallado catastróficamente en cumplir sus especificaciones objetivo, falla en capturar el matiz crítico del comportamiento específico de la herramienta bajo las condiciones en las cuales fue realizado el pasaje.

El presente trabajo técnico tiene como objetivo desglosar la metodología de los niveles 2 y 3, y la comparación de la aplicación de ambos para distintas corridas ILI, puntualizado en las tolerancias, sesgos y niveles de confianza de cada método. Mediante la caracterización estadística avanzada del error por Nivel 3, es posible mejorar los márgenes de seguridad conservadores en variables de ingeniería precisas como lo es el cálculo de probabilidad de falla de defectos (metodología POE) y optimizar los planes de gestión de integridad reduciendo la cantidad de verificaciones, reparaciones y asignando los recursos de forma más eficiente para aquellas amenazas que realmente lo ameritan.

DESARROLLO

Fundamentos Normativos y la Insuficiencia del Enfoque Binomial

La norma API 1163 estructura la calificación de los sistemas ILI basándose en el control de las "Variables Esenciales". Estas variables, detalladas en la Sección 5 del estándar, incluyen parámetros operativos como la velocidad de excursión, la limpieza del ducto, y características físicas como el espesor de pared y el tipo de costura. El proceso de validación, descrito en la Sección 8, busca confirmar que, bajo las condiciones específicas de la corrida, las variables esenciales se mantuvieron dentro de los rangos calificados y que el desempeño resultante es aceptable.

Para lograr esta confirmación empírica, la Sección 8 del estándar establece tres niveles jerárquicos de validación. La selección del nivel adecuado recae en el operador y depende directamente de la criticidad del ducto, la disponibilidad de datos de referencia y el grado de certidumbre requerido para los análisis de ingeniería posteriores.

Validación de Nivel 1

Este constituye el enfoque cualitativo y de menor complejidad estadística. Se fundamenta en la premisa de equivalencia operativa: se emplea cuando se puede demostrar formalmente que la tecnología de inspección, los procedimientos analíticos y las condiciones físicas del ducto (las variables esenciales) son análogas a corridas previas que ya han sido validadas de forma exitosa. Su característica definitoria es que no exige la realización de nuevas excavaciones exploratorias ni mediciones directas en campo (END) para aprobar el desempeño de la corrida actual.

Si bien su principal ventaja es la economía de recursos operativos y logísticos, conlleva la mayor carga de incertidumbre de los tres niveles. Al asumir que el éxito histórico garantiza el desempeño presente, el Nivel 1 resulta desaconsejable para líneas que enfrentan amenazas progresivas, escenarios de corrosión activa o modificaciones recientes en sus regímenes de flujo.

Validación de Nivel 2

Representa el estándar de práctica predominante en la industria, utilizado fundamentalmente para el cierre y aceptación de contratos de servicio. Este nivel de validación consiste en comparar los datos reportados por la herramienta ILI contra las anomalías inspeccionadas en campo mediante un análisis cuantitativo para “rechazar” o “no rechazar” la especificación de la herramienta.

A partir de “X” mediciones de campo que caen dentro de la tolerancia de la especificación comercial de la herramienta (típicamente $\pm 10\%$ del espesor de pared) sobre un total de “n” mediciones, $\hat{p}$ determina la proporción mediciones que caen dentro de dicha tolerancia. Para limitar la incertidumbre de la proporción estimada, se determina la proporción poblacional $\tilde{p}$ ajustada de mediciones que caen dentro de dicha tolerancia, y $\tilde{n}$ el tamaño de muestra ajustada.

Para determinar cuál es la probabilidad de que las mediciones caigan dentro de la especificación ILI dado un determinado número de mediciones, se determinan los intervalos de confianza binomial $[\hat{p}_{lower}, \hat{p}_{upper}]$ mediante el método de “Agresti-Coull” que representa la probabilidad mínima y máxima con un nivel de confianza $(1-\alpha)$, con $Z_{\alpha/2}$ como el percentil $\alpha/2$ de la distribución normal estándar.

$$\hat{p}_{lower} = \tilde{p} - Z_{\alpha/2} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}$$

$$\hat{p}_{upper} = \tilde{p} + Z_{\alpha/2} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}$$

De esta manera, se obtienen 3 resultados posibles en relación a la certeza p especificada por la herramienta:

- **Resultado 1:** $\hat{p}_{upper} < p$. El limite superior del intervalo de confianza observado es menor a la especificada por la herramienta. La especificación comercial es rechazada. El operador debe coleccionar más datos de campo para determinar si la “performance” actual es mejor que la calculada,

o debe pasarse a un nivel de evaluación 3.

- **Resultado 2:** $\hat{p}_{lower} < p < \hat{p}_{upper}$. El límite inferior del intervalo de confianza observado contiene la especificación comercial de la herramienta. La especificación del proveedor puede caracterizar con precisión el rendimiento real, pero otros rendimientos son plausibles. Antes de utilizar la especificación del proveedor, se debe considerar la amplitud del intervalo de confianza y el punto final del intervalo de confianza inferior. Un intervalo amplio o un punto final del intervalo de confianza bajo indican un alto grado de incertidumbre en la estimación.
- **Resultado 3:** $p < \hat{p}_{lower}$. El punto final del intervalo de confianza inferior supera la certeza especificada por el proveedor. La especificación del proveedor puede utilizarse con seguridad. Es probable que el rendimiento real sea mejor que el especificado por el proveedor.

API 1163 no establece los límites de confianza inferior y superior y el número de verificaciones de campo mínimas que influyen en la incerteza y errores asociados en la aceptación la validación, los cuales pueden llevar, por ejemplo, no rechazar una herramienta que tiene una performance menor a la especificada. El reporte PRCI PR-719 recomienda un nivel de confianza del 90%, con un límite inferior del 60% (que representa una tolerancia equivalente de $\pm 15,2\%$ del espesor de pared) y un número mínimo de verificaciones de 6.

Aunque el Nivel 2 proporciona una resolución administrativa clara, adolece de una limitación ingenieril severa: no define la magnitud real del error. Al conformarse con verificar si la herramienta "cumplió", el modelo es ciego ante la posible existencia de un sesgo sistemático (bias) o ante el hecho de que la dispersión real de los datos pueda ser significativamente menor que la tolerancia de catálogo, forzando al operador a utilizar valores conservadores genéricos en sus modelos de integridad.

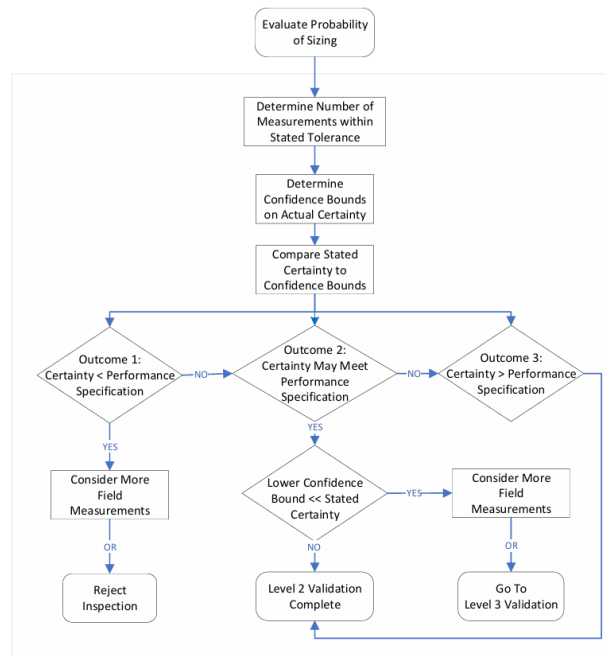


Fig. 1. Diagrama de flujo. Validación Nivel 2.

Validación de Nivel 3

Constituye el estadio de mayor madurez analítica y representa el núcleo metodológico de este trabajo técnico. El Nivel 3 abandona la verificación de la validación netamente contractual para enfocarse en la determinación de la “performance” real de la herramienta en base a las mediciones de campo mediante métodos estadísticos. Requiere un muestreo de campo de alta calidad y la aplicación de intervalos de tolerancia estadística avanzados (típicamente modelos basados en la distribución *t* no central o métodos de Howe-Guenther). A través de este procesamiento, se genera una nueva hoja de especificaciones de desempeño construida exclusivamente con los datos de esa inspección en particular (“As-Run”). Su superioridad técnica radica en su capacidad para desagregar y cuantificar tanto la exactitud —identificando y corrigiendo matemáticamente cualquier sesgo sistemático de la herramienta— como la precisión, mediante el cálculo de la desviación estándar real del error. Al proveer esta caracterización profunda, el Nivel 3 mitiga las limitaciones conservadoras del Nivel 2, entregando a los especialistas de integridad datos calibrados que reducen drásticamente la incertidumbre epistémica en las evaluaciones de Aptitud para el Servicio (FFS) y la probabilidad de falla en análisis de riesgos.

La industria ha adoptado masivamente el Nivel 2 de validación debido a su simplicidad matemática. Este nivel utiliza la distribución binomial para evaluar una hipótesis nula. Esencialmente, se plantea la pregunta: “¿Es la cantidad de errores fuera de tolerancia lo suficientemente baja como para afirmar con un X% de confianza que la herramienta cumple su especificación?”. Si la respuesta es afirmativa, se acepta el reporte completo. No obstante, la guía PRCI PR-719 advierte sobre las limitaciones de este enfoque, particularmente en el “resultado 2”. En este escenario, la herramienta cumple técnicamente con la especificación, pero el análisis binario oculta si existe un sesgo sistemático (*bias*) o si la dispersión real de los datos es significativamente menor a la contratada. Al quedarse en el Nivel 2, el operador se ve obligado a utilizar la tolerancia genérica del proveedor (típicamente $\pm 10\%$ del espesor de pared) para todos sus cálculos de presión de falla, desperdiciando la precisión potencial de herramientas de alta resolución.

La Figura 2 ilustra conceptualmente el cambio de paradigma estadístico y la consecuente reducción de incertidumbre que ofrece la transición hacia una validación de Nivel 3. El área sombreada rectangular representa el enfoque normativo del Nivel 2; al carecer de una caracterización del error, el analista de integridad se ve forzado a asumir una distribución uniforme, donde cualquier error de dimensionamiento dentro de la banda genérica del proveedor (por ejemplo, $\pm 10\%$ del espesor de pared) se considera igualmente probable. Esta amplitud refleja un alto grado de incertidumbre epistémica y conservadurismo.

Por el contrario, la curva acampanada representa el desempeño “As-Run” caracterizado mediante el Nivel 3. Al procesar las mediciones reales obtenidas en la zanja, el error deja de ser un bloque de incertidumbre para convertirse en una distribución normal parametrizada. Visualmente, esta curva revela dos factores críticos que el Nivel 2 oculta: primero, el sesgo sistemático o *bias* (evidenciado por el desplazamiento del pico de la campana respecto al eje de error cero) y segundo, la precisión real de la herramienta (representada por la estrechez de la campana, dictada por su desviación estándar empírica, σ). Esta condensación visual de la incertidumbre demuestra cómo el Nivel 3 entrega a los modelos de Aptitud para el Servicio (FFS) datos calibrados, permitiendo recuperar márgenes de seguridad que antes se perdían bajo tolerancias comerciales genéricas.

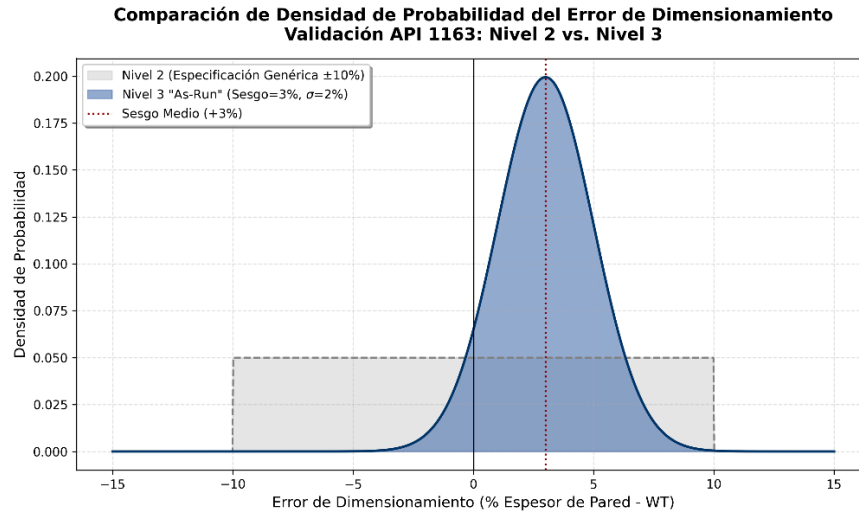


Fig. 2: Gráfico comparativo de densidades de probabilidad. Se muestra una distribución 'rectangular' asumida en el Nivel 2 frente a una distribución normal ajustada en el Nivel 3, ilustrando la reducción de incertidumbre.

Metodología de Validación Nivel 3: El Paradigma "As-Run"

La validación de Nivel 3 constituye un salto cualitativo hacia la ingeniería estadística. En lugar de verificar la especificación comercial, este nivel busca caracterizar el error de medición como una variable aleatoria continua. El fundamento matemático reside en el cálculo de Intervalos de Tolerancia Estadística. A diferencia de un intervalo de confianza simple que acota la media, un intervalo de tolerancia acota una proporción específica de la población (certeza) con un nivel de confianza dado.

En la validación de Nivel 3 el objetivo final es obtener la verdadera "especificación" de la herramienta ("As-Run"). Para lograr esto, no basta con promediar los errores encontrados en campo; necesitamos un margen que nos garantice, por ejemplo, que el 80% de los defectos reales estarán contenidos dentro de la banda, con un 95% de nivel de confianza.

Para resolver esto de manera directa, la industria y la guía PRCI PR-719 recomiendan utilizar la aproximación de Howe-Guenther. En términos prácticos, este método es una herramienta que nos entrega un factor multiplicador (k), el cual define qué tan ancha debe ser nuestra banda de tolerancia en función de los datos obtenidos.

La nueva especificación operativa de la herramienta se calcula entonces con una ecuación muy sencilla:

$$E_{AsRun} = Sesgo \pm (k \cdot \sigma)$$

Donde el Sesgo es el error promedio de la herramienta (cuánto sobrestima o subestima en general) y σ es la desviación estándar (la dispersión de los errores medidos en campo).

El gran valor del método de Howe-Guenther para la gestión de ductos radica en cómo define ese factor k . Este número no es estático, sino que depende directamente de la cantidad de excavaciones (n):

1. **Penalización por falta de datos:** Si se validan muy pocos defectos en campo (por ejemplo, $n=3$), el método de Howe-Guenther arroja un factor k muy grande. Esto ensancha la banda de tolerancia, obligando al operador a ser conservador en sus cálculos de presión de falla ante la falta de evidencia empírica.
2. **Recompensa por muestreo robusto:** A medida que se incrementan las mediciones de campo de alta calidad (idealmente $n > 10$), el factor k se reduce drásticamente.

La figura 3 ilustra el diagrama de flujo general del método de validación de corridas ILI para los niveles 1, 2 y 3, destacando el nivel 3.

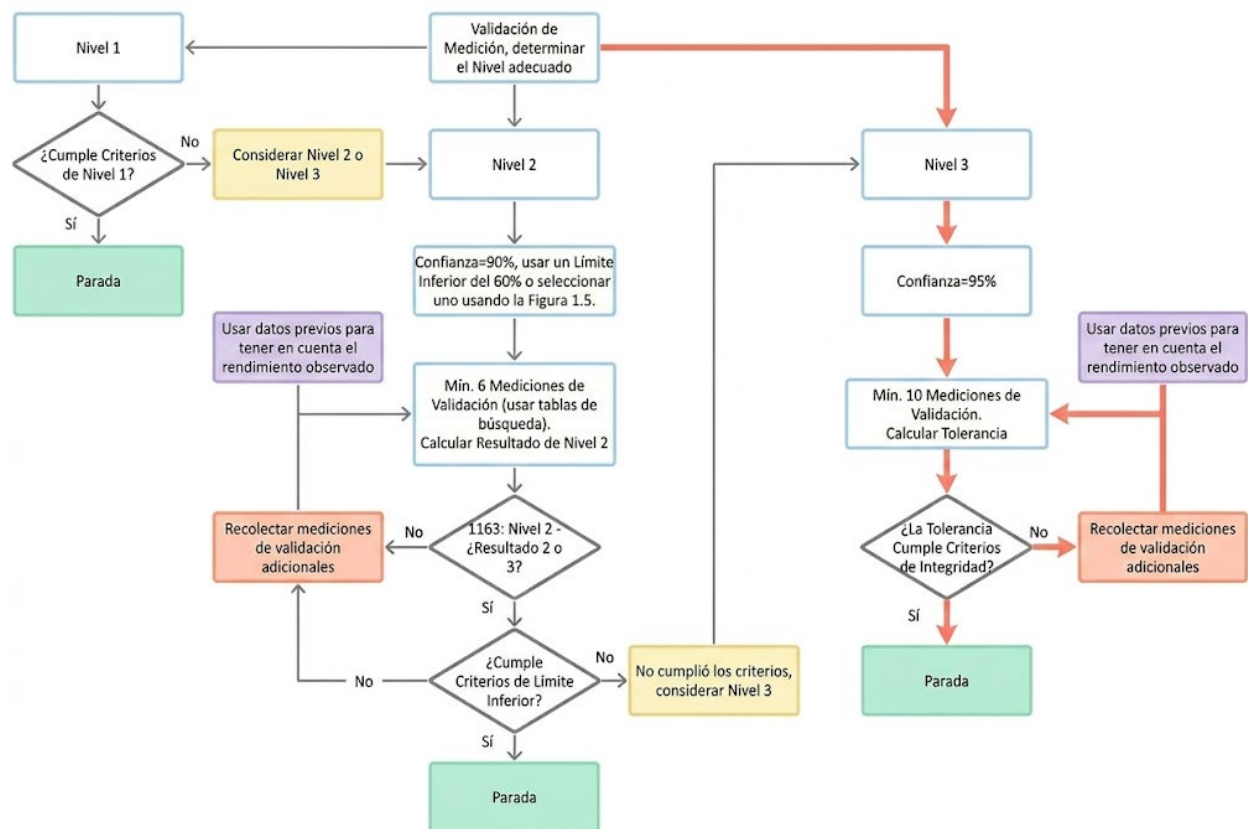


Fig. 3. Diagrama de flujo de validación según API 1163, destacando la validación Nivel 3.

La siguiente tabla muestra las diferencias entre los niveles de validación 2 y 3 planteados en este trabajo.

Nivel 2

- * **Familiaridad dentro de la industria (API 1163):** Es un método conocido y estandarizado.
- * **Resultado simple de aprobado/reprobado:** Se determina si se cumple o no con la especificación de la Inspección en Línea (ILI).
- * **No considera el grado de desviación:** Solo detecta que hubo un error de medición (mal reporte), sin evaluar cuánto se sobreestimó o subestimó la anomalía.
- * **El conservadurismo disminuye con la confianza:** A mayor confianza en los datos, los márgenes de seguridad suelen ser menos estrictos.
- * **Aumento del error:** El error en el resultado aumenta a medida que disminuye el umbral del punto final inferior (ver Figura 16).
- * **Similitud con el Nivel 3:** Puede producir resultados muy similares al Nivel 3 si el umbral del punto final inferior se ajusta a la certeza de la especificación (Resultado 3 del Nivel 2).

Nivel 3

- * **Menos común en la industria:** Su uso no es tan extendido como el Nivel 2.
- * **Intervalo de tolerancia calculado:** Toma en cuenta la dispersión de los datos y la incertidumbre derivada de un tamaño de muestra pequeño.
- * **Considera el grado de desviación:** Evalúa cuánto se desvía la lectura de la ILI con respecto a la media.
- * **El conservadurismo aumenta con la confianza:** Los márgenes de seguridad se vuelven más estrictos conforme aumenta la confianza estadística.
- * **Requerimiento de datos:** Potencialmente requiere más puntos de datos para ofrecer una interpretación significativa del desempeño de la ILI.

Tabla 3. Comparación del Nivel 2 y 3 de validación según API 1163.

Aplicación en Campo: Protocolos de Adquisición de Datos de campo

La validez de cualquier modelo estadístico depende intrínsecamente de la calidad de los datos de entrada. Por ello, la aplicación en campo para una validación Nivel 3 exige un rigor superior al de las excavaciones de mantenimiento rutinario. En TGN, la estrategia de campo se ha redefinido para cumplir con los requisitos del Anexo C de la norma API 1163.

El primer desafío es la selección de los sitios de verificación. No se deben seleccionar únicamente los defectos más severos, ya que esto sesgaría la muestra hacia los extremos de la distribución ("sesgo de cola"). Se implementa una selección estratificada que incluye una gama representativa de profundidades, longitudes y orientaciones horarias. Además, se priorizan los tramos con alta densidad de defectos que permiten obtener una muestra "n" estadísticamente significativa (idealmente $n > 10$ para Nivel 3).

En campo, la premisa es que el error de la medición de referencia debe ser despreciable frente al error del sistema ILI. Las técnicas manuales con "pit gauge" son a menudo insuficientes debido a la dependencia del operador. Por ello, se ha estandarizado el uso de escáneres láser 3D (Handyscan o similares) o mapeo de corrosión por ultrasonido automatizado (AUT). Estas tecnologías generan una nube de puntos digital de la superficie corroída con una resolución sub-milimétrica.

Impacto en la Gestión de Integridad y Toma de Decisiones

La integración de los resultados de la validación Nivel 3 en los modelos de cálculo de la probabilidad de falla de defectos (Probabilidad de Excedencia POE) genera un impacto directo en el Plan de Integridad a través de los resultados de los análisis de riesgo. Al reemplazar la tolerancia genérica del proveedor por la tolerancia validada y corregida por sesgo, se altera la Probabilidad de Falla (POF) calculada para cada defecto.

Ante la incertidumbre de la herramienta ILI (bajo una validación Nivel 2), la práctica estándar obliga a adoptar el escenario más desfavorable: al valor de profundidad reportado por la herramienta se le suma directamente la tolerancia máxima especificada por el proveedor (por ejemplo, +10% del espesor de pared). El resultado es el cálculo de una Presión de Falla única y penalizada, sobre la cual se aplica un factor de seguridad. Si la Presión Máxima Operativa Admisible (MAPO) supera este límite penalizado, el defecto se considera crítico.

La validación estadística "As-Run" (Nivel 3) disrumpe este apilamiento de conservadurismos. Al no depender de una tolerancia máxima teórica, sino de una función de densidad de probabilidad real caracterizada por un sesgo y una desviación estándar, se habilita la transición hacia metodologías de evaluación probabilística, como el Análisis de Confiabilidad Estructural (SRA) o simulaciones de Monte Carlo.

Bajo este nuevo paradigma, la variable de entrada para el modelo de presión de falla deja de ser un número fijo (ej. Profundidad + 10%) para convertirse en una distribución estadística. Esto permite calcular ya no un "pasa/no pasa" determinístico, sino la Probabilidad de Excedencia (POE) y, consecuentemente, la Probabilidad de Falla (POF). La POF cuantifica matemáticamente la probabilidad de que la presión de falla sea menor a la presión de operación MAPO de la cañería.

Esto permite una optimización de los recursos económicos y humanos en las anomalías que verdaderamente comprometen la seguridad estructural del ducto, permitiendo la permanencia en servicio de anomalías que el método determinístico condenaría prematuramente, traduciendo la precisión estadística en eficiencia de capital.

Resultados de validación de corridas MFL del sistema de ductos de TGN

Para el sistema de ductos de TGN, se presentan los siguientes resultados de validación de nivel 2 y 3 de corridas MFL para 5 gasoductos particularmente.

Tramo	Nro defectos medidos	Validación de corrida				
		Nivel 2		Nivel 3		
		Intervalo de confianza [%WT]	Outcome	Sesgo [%WT]	Tolerancia [%WT]	Intervalo de confianza [%WT]
A	17	[56,4 ; 89,3]	1 - Fuera de especificación	1	+/- 19,69	[-18,69 ; 20,69]
B	11	[61 ; 100]	2 - No Rechazado	-3,97	+/- 18,16	[-22,13 ; 14,19]
C	23	[76,03 ; 97,88]	2 - No Rechazado	-0,58	+/- 16,09	[-16,68, 15,51]
D	10	[75,12 ; 100]	2 - No Rechazado	0,6	+/- 13,62	[-13,02 ; 14,22]
E	35	[80,4 ; 99,4]	3 - No Rechazado	5.74	+/- 4,33	[1,41 ; 10,08]

Tabla 1. Resultados de validación de corridas MFL. Nivel 2 y 3.

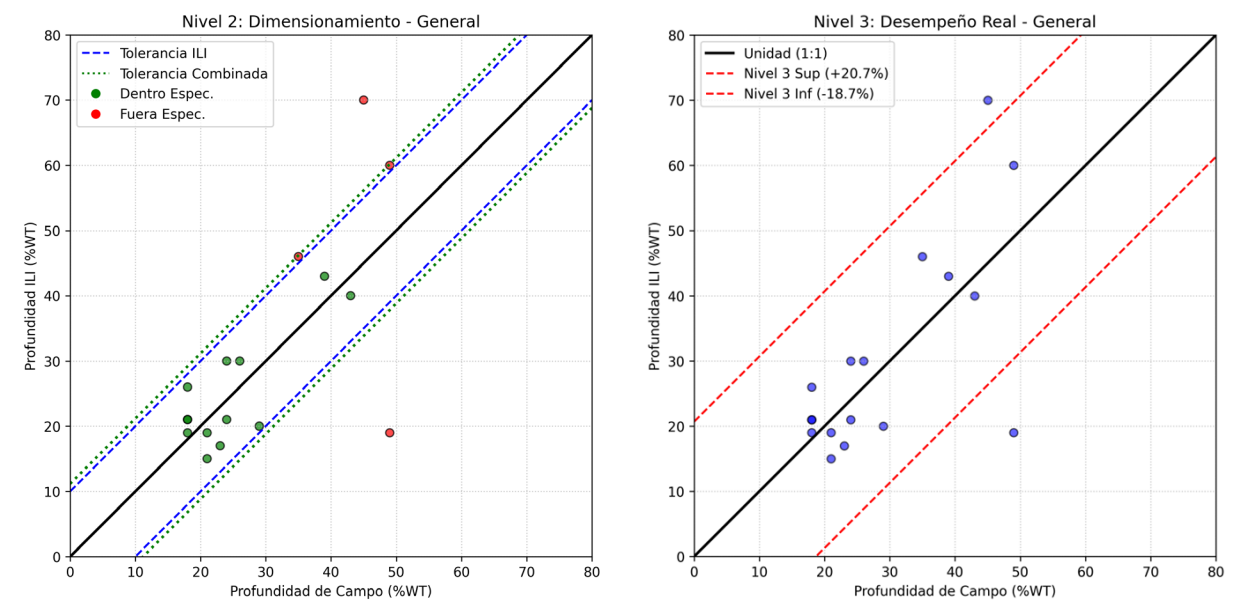


Fig. 4. Gráficos de validación Nivel 2 y 3. Tramo A.

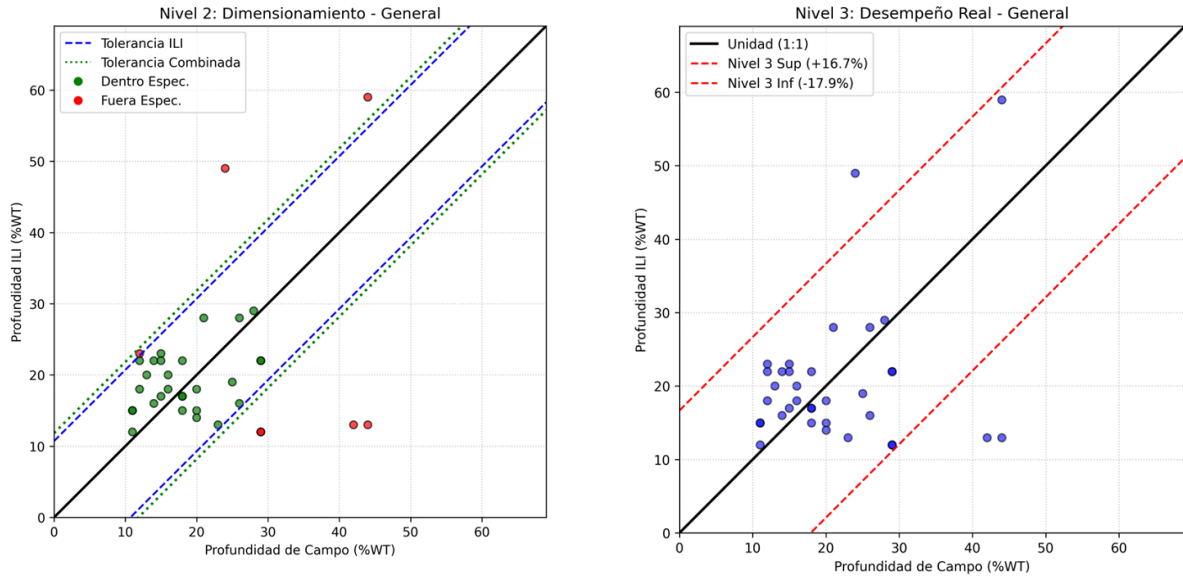


Fig. 5. Gráficos de validación Nivel 2 y 3. Tramo C.

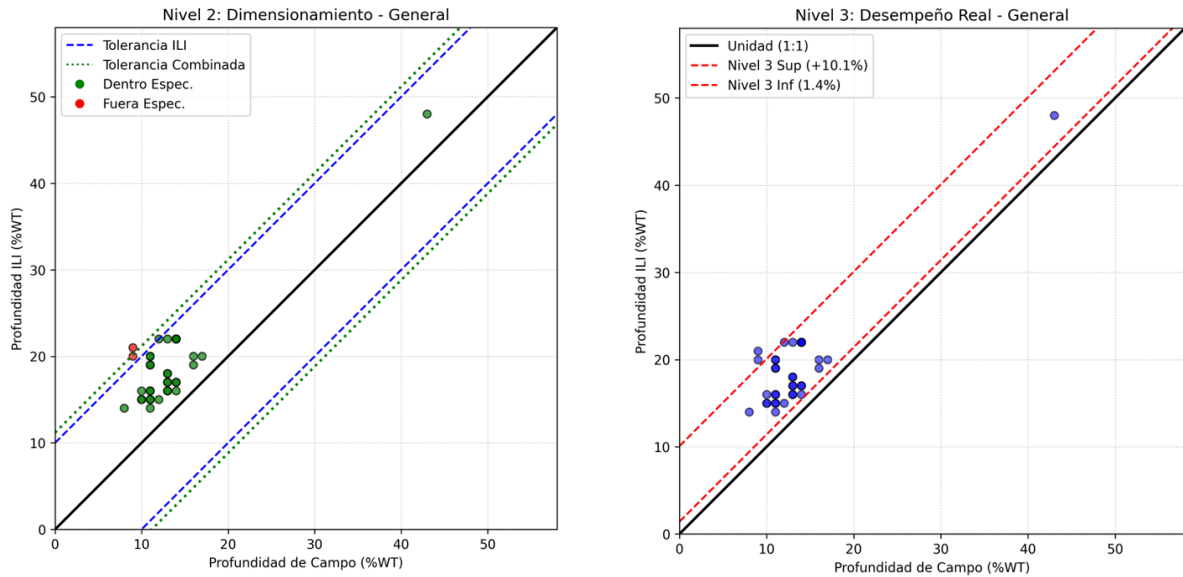


Fig. 6. Gráficos de validación Nivel 2 y 3. Tramo E.

El análisis detallado de estos resultados empíricos expone las falencias del enfoque de cumplimiento binario (Nivel 2) y subraya la superioridad técnica de la caracterización "As-Run" (Nivel 3) para la toma de decisiones en integridad. Las ventajas del Nivel 3 se evidencian en tres escenarios operativos distintos, delineados a continuación.

El peligro más significativo del Nivel 2 es el escenario de "Outcome 2" (No Rechazado). En los Tramos B, C y D, la prueba binomial del Nivel 2 no encontró evidencia estadística suficiente para rechazar la corrida.

Bajo la práctica estándar de la industria, un operador aceptaría estos reportes y utilizaría la especificación genérica del proveedor (típicamente $\pm 10\%$ del espesor de pared) para cálculos de probabilidad de falla por POE.

Sin embargo, el análisis de Nivel 3 revela un escenario más complejo: el desempeño real de la herramienta en estos tramos fue considerablemente inferior al estándar comercial. Las tolerancias "As-Run" calculadas arrojaron dispersiones de $\pm 18,16\%$ WT (Tramo B), $\pm 16,09\%$ WT (Tramo C) y $\pm 13,62\%$ WT (Tramo D). Si el operador se hubiese detenido en el Nivel 2, habría asumido una incertidumbre del $\pm 10\%$, subestimando gravemente el error real y aceptando un riesgo latente inaceptable en su cálculo de Aptitud para el Servicio (FFS).

Aún más crítico es el caso del Tramo B, donde el Nivel 3 identificó un sesgo sistemático de $-3,97\%$ WT. Este valor negativo indica que la herramienta subestimó consistentemente la profundidad de los defectos. El Nivel 2 es completamente ciego ante este sesgo direccional. Al aplicar el Nivel 3, el operador de TGN puede penalizar matemáticamente las mediciones reportadas, sumando este casi un 4% de subestimación antes de evaluar la integridad, garantizando así la seguridad física del activo.

El Tramo E (basado en 35 mediciones) ejemplifica la capacidad del Nivel 3 para obtener una especificación más realista y optimizar recursos. En este caso, el Nivel 2 arrojó un "Outcome 3", indicando que el desempeño fue estadísticamente mejor que la especificación.

Al aplicar el Nivel 3, se descubrió una precisión superior al comercial de la herramienta; es decir una tolerancia más estrecha del $\pm 4,33\%$ WT. Simultáneamente, se identificó un sesgo positivo pronunciado de $+5,74\%$ WT (sobrestimación sistemática). Esta caracterización precisa incrementa dramáticamente la Presión de Falla calculada, y disminuyendo la probabilidad de falla en análisis de riesgos.

Finalmente, en el Tramo A, el Nivel 2 cumplió su función básica de protección arrojando un "Outcome 1" (Fuera de especificación / Rechazado). Según el flujograma de validación de corridas, se debería optar por aumentar la muestra de defectos en campo y/o un análisis de validación Nivel 3.

CONCLUSIONES

La gestión de integridad moderna no admite la incertidumbre no gestionada. La transición de la validación de cumplimiento (Nivel 2) a la validación de caracterización (Nivel 3) propuesta en este trabajo, alineada con API 1163 y PRCI PR-719, ofrece resultados más ajustados a la realidad, lo que se traduce en cálculos más seguros.

Los análisis demuestran que el comportamiento real de una herramienta ILI rara vez coincide exactamente con su hoja de especificaciones comerciales; a menudo es mejor en precisión, pero puede contener sesgos sistemáticos que deben ser corregidos. La metodología Nivel 3 permite capturar esta realidad, "limpiando" los datos de ruido y sesgo antes de que ingresen a los modelos de evaluación crítica.

Para TGN, la adopción de este enfoque significa pasar de una metodología determinística a una probabilística, con un mayor grado de confianza. Permite extender la vida útil de los activos de manera segura, evitar intervenciones innecesarias y, fundamentalmente, basar la seguridad pública y ambiental en datos empíricos validados y no en promesas contractuales genéricas.

BIBLIOGRAFÍA

1. **American Petroleum Institute (API).** (2021). *API Standard 1163: In-line Inspection Systems Qualification* (3rd Edition). Washington, D.C.
2. **Pipeline Research Council International (PRCI).** (2023). *Performance Validation Guidelines for API 1163*. Catalog No. PR-719-223803-R01. Prepared by Integral Engineering.
3. **American Society of Mechanical Engineers (ASME).** (2014). *ASME B31.8S: Managing System Integrity of Gas Pipelines*. New York, NY.
4. **American Petroleum Institute (API).** (2016). *API Recommended Practice 1160: Managing System Integrity for Hazardous Liquid Pipelines*.
5. **American Petroleum Institute (API) / ASME.** (2016). *API 579-1/ASME FFS-1: Fitness-For-Service*.
6. **Howe, W. G.** (1969). *Two-sided Tolerance Limits for Normal Populations - Some Improvements*. Journal of the American Statistical Association, 64, 610-620.
7. **Desjardins, G.** (2014). *Advanced ILI Validation: Moving Beyond Pass/Fail*. Proceedings of the 10th International Pipeline Conference (IPC2014). Calgary, AB.

BREVE CV

Alan Cornejo

Ingeniero Mecánico, graduado de la Universidad Tecnológica Nacional (UTN FRBA), con más de 11 años de experiencia en la industria del petróleo y gas. Desde 2015, me he dedicado a tareas relacionadas con la gestión de inspecciones internas, la definición de verificaciones en campo, el análisis de aptitud para el servicio y la implementación de medidas de mitigación.

Emmanuel Benitez

Ingeniero Electromecánico, graduado en la Universidad Tecnológica Nacional (UTN FRP). Nivel 2 en Partículas Magnetizables y Nivel 2 en Líquidos Penetrantes, según IRAM-NM-ISO 9712. Ingreso a TGN en 2018 como Mecánico de Gasoductos. Desde 2022 a la actualidad se desempeña como Analista de Integridad en el Departamento de Integridad de TGN.

Santiago Parodi

Ingeniero en Materiales, graduado de la Universidad Nacional de San Martín (UNSAM), con más de 8 años de experiencia en la industria del gas y petróleo. Desde 2017, me he dedicado a tareas de gestión de integridad de ductos, inspección basada en riesgo (RBI) y aptitud para el servicio y mitigación de defectos.