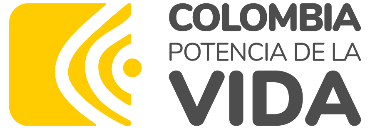


Informe Técnico: Validaciones



Contenido

[1. Objetivos 3](#_Toc207904943)

[Objetivo general 3](#_Toc207904944)

[Objetivos específicos 3](#_Toc207904945)

[2. Metodología 4](#_Toc207904946)

[3. Resultados 5](#_Toc207904947)

[CONCLUSIÓN 6](#_Toc207904948)

Índice de tablas

[Tabla 1: Registros de terrenos Urbanos y rurales con sus informalidades](#_Toc199374285)

[Tabla 2: Depuración inicial de terrenos](#_Toc199374286)

Índice de Ilustraciones

[Ilustración 1: Diagrama de Metodología](#_Toc199374306)

[Ilustración 2: Caso 1 Emparejamiento directo por número predial y código de Terrenos](#_Toc199374307)

# Objetivos

## Objetivo general

Garantizar la calidad, coherencia y trazabilidad de los datos LADM-COL en PREPROD y PROD mediante un paquete de reglas SQL reproducibles, que detecten inconsistencias críticas y generen salidas estandarizadas para su corrección y seguimiento.

## Objetivos específicos

* Inventariar las reglas cubiertas
* Precisar entradas y llaves de enlace usadas: id\_operacion ↔ id\_operacion\_predio, predio\_guid ↔ globalid, y NPN/código.
* Definir la metodología estándar (staging/normalización, joins, SRID 9377, criterios de incumplimiento).
* Establecer la configuración de ejecución por ambiente:
* Publicar las salidas esperadas por regla (estructura reglas.regla\_XXX: regla, objeto, tabla, id\_operacion, npn, descripcion, valor, cumple, timestamps).
* Definir criterios de severidad y priorización de hallazgos para orientar la remediación.
* Entregar lineamientos de rendimiento (índices mínimos, ANALYZE, EXPLAIN(ANALYZE)) y orden de ejecución.

# Metodología

La metodología consiste en validar la consistencia y coherencia del modelo LADM-COL mediante reglas SQL reproducibles que leen desde PREPROD/PROD, normalizan datos (btrim/lower/NULLIF) y unen entidades por claves operacionales (id\_operacion ↔ id\_operacion\_predio) o por GUID (predio\_guid ↔ globalid), aplicando cuando existe el filtro transversal COALESCE(gdb\_branch\_id,0)=0.

Cada regla sigue un patrón simple:

Tabla : Diseño de reglas

|  |
| --- |
| * **staging con CTE base para limpiar y estandarizar** |
| * **eval para calcular banderas y métricas (cardinalidades, patrones NPN, formatos registrales, sumas de áreas, duplicados de interesados/derechos)** |
| * **out para componer el mensaje de incumplimiento; las áreas geométricas se calculan en SRID 9377 transformando cuando haga falta.** |

Las salidas se materializan en tablas reglas.regla\_XXX con esquema homogéneo (regla, objeto, tabla, objectid, globalid, id\_operacion, npn, descripcion, valor, cumple=false, created\_at, updated\_at) para trazabilidad y explotación. Para rendimiento se evitan funciones en cláusulas JOIN (se normaliza en base o se usan índices de expresión), se mantienen índices en id\_operacion, id\_operacion\_predio, globalid y predio\_guid; La ejecución es idempotente (DROP/CREATE TABLE AS), con verificación previa de fuentes y posterior consolidación opcional de indicadores.

Para manejar id\_operacion se obtuvo **propagándolo desde la fuente canónica preprod.t\_ilc\_predio.id\_operacion** hacia las demás tablas. El orden de resolución fue: **(1)** cuando la tabla traía **predio\_guid**, se hizo *join* exacto con preprod.t\_ilc\_predio.globalid y se asignó id\_operacion\_predio := p.id\_operacion (siempre con COALESCE(gdb\_branch\_id,0)=0 donde aplique, en este caso el mismo query pero se colocaba un filtro para datos de producción); **(2)** si ya venía un **id\_operacion\_predio** válido, solo se validó contra p.id\_operacion; **(3)** si la tabla tenía **numero\_predial\_nacional (NPN)** o un campo equivalente (p. ej., codigo en UC), se enlazó con p.numero\_predial\_nacional y se tomó p.id\_operacion (controlando ambigüedades: branch=0, misma autoridad/ciclo y, si había más de un match, se eligió con DISTINCT ON o por fecha más reciente); **(4)** si la tabla se relacionaba vía otra entidad (p. ej., Derecho o DatosPHCondominio) que sí tenía predio\_guid o id\_operacion\_predio, se encadenó el *join* y se heredó el id\_operacion de Predio; **(5)** como último recurso, en clases puramente espaciales se usó **emparejamiento geográfico** (intersección/contención con tolerancia de área) para identificar el predio único y, de allí, su id\_operacion. Con esto, en cada staging CTE se normalizó (btrim/lower/NULLIF), se filtró **branch=0** en todas las tablas que lo tuvieran, y se materializó el resultado; por ejemplo, para DatosAdicionales y Derecho se hizo: JOIN preprod.t\_ilc\_predio p ON p.globalid = d.predio\_guid → **id\_operacion\_predio := p.id\_operacion**; para Interesado, cuando faltaba, se resolvió por **NPN** o por el **Derecho** asociado y se validó unicidad antes de escribir.

Diagrama

El contenido generado por IA puede ser incorrecto.

Ilustración : Metodología de validación

# Resultados

Para estos resultados se ejecutaron las siguientes reglas que se pueden evidenciar en la base de datos de test, tanto en reglas\_preprod como en reglas\_prod.

Tabla : Reglas resumen

|  |  |
| --- | --- |
| **Etiquetas de fila** | **Cuenta de resumen** |
| Geográfica | 17 |
| OK | 110 |
| OK\* (Sin datos) | 97 |
| (en blanco) |  |
| **Total reglas** | **224** |

Tabla 3: Reglas por implementar

|  |  |
| --- | --- |
| **Etiquetas de fila** | **Cuenta de Estado** |
| Falta | 94 |
| OK | 130 |
| (en blanco) |  |
| **Total general** | **224** |

Para ver información completa de resultado por cada tabla dirigirse al Excel anexo a este documento en donde se encuentra el semáforo de reglas en donde se encuentran reglas hechas pero sin información o con información y reglas por implementar

# CONCLUSIÓN

En síntesis, el documento deja estandarizada una metodología reproducible para validar LADM-COL: se normalizan fuentes, se enlazan entidades propagando **id\_operacion\_predio** desde **ILC\_Predio** (prioridad: *predio\_guid ⇄ globalid*, luego *id\_operacion* y, si hace falta, NPN/código con desambiguación), y se evalúan reglas que materializan resultados en tablas **reglas.regla\_XXX** con campos homogéneos (regla, objeto, tabla, id\_operacion, npn, descripcion, valor, timestamps). Se cubrió un conjunto amplio de reglas (series 6xx–8xx), enfocadas en inconsistencias críticas (duplicados de interesados y documentos, NPN, cardinalidades/áreas, y cruces Derecho–Predio como *Privado ≠ Ocupación*), con lineamientos de rendimiento (índices, ANALYZE/EXPLAIN) que mantienen tiempos de cómputo estables. Persisten limitaciones por tablas con datos faltantes, que se listan para seguimiento y no como fallas de la lógica.