

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 1 de 14



IDEAM

**INSTITUTO DE HIDROLOGIA,
METEOROLOGIA Y
ESTUDIOS AMBIENTALES**

**REGLAS DE VALIDACIÓN
OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS - OEGRDP**

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 2 de 14

1 INTRODUCCIÓN

El IDEAM cuenta entre sus funciones, con la de entregar la información ambiental disponible a las entidades pertenecientes al Sistema Nacional Ambiental – SINA (MAVDT, 1994), labor que realiza a través del Sistema de Información Ambiental de Colombia – SIAC.

La operación estadística del Registro de Generadores de Residuos o Desechos Peligrosos – RESPEL se desarrolló para contar con información normalizada, homogénea y sistemática sobre la generación y manejo de residuos o desechos peligrosos originados por las diferentes actividades productivas y sectoriales del país.

La Resolución 1362 del 2 de agosto de 2007 expedida por el antes Ministerio de Ambiente, Vivienda y Desarrollo Territorial, (hoy Ministerio de Ambiente y Desarrollo Sostenible - MADS), reglamentó el Registro de Generadores de Residuos o Desechos Peligrosos con base en lo estipulado en el Artículo 27 del Decreto 4741 del 30 de Diciembre de 2005 (Compilado en el Decreto 1076 de 2015), de acuerdo con los estándares para el acopio de datos, procesamiento, transmisión y difusión de la información que estableció el IDEAM para tal fin (Resolución 043 de 2007).

El Registro de Generadores de Residuos o Desechos Peligrosos es un instrumento de gestión mediante el cual se captura información de manera sistemática y continúa, sobre la generación y el manejo de residuos o desechos peligrosos, originados por las diferentes actividades productivas y sectoriales del país. La información que permite consultar es la generación de Residuos Peligrosos en un año determinado, por tipo de RESPEL, por actividad económica clasificada por código CIU (Clasificación Industrial Internacional Uniforme de todas las actividades económicas), por Autoridad Ambiental, por municipio, por departamento, por estado de la materia, así como las cantidades gestionadas (aprovechamiento, tratamiento y disposición final).

Los datos son reportados anualmente por los generadores de residuos peligrosos que están inscritos de manera previa ante la Autoridad Ambiental. Una vez se surte el procedimiento de revisión y verificación para garantizar la calidad del dato, la Autoridad Ambiental transmite la información de cada establecimiento generador al IDEAM, entidad que se encarga de administrar el Sistema de Información y consolidar, analizar y publicar los informes nacionales anuales e indicadores ambientales sobre el tema.

Por tratarse de registros administrativos oficiales utilizados para diseño, formulación, seguimiento y evaluación de planes, programas y políticas públicas, los datos administrados por el RESPEL hacen parte del Sistema Estadístico Nacional – SEN, que se encuentra bajo la coordinación del Departamento Administrativo Nacional de Estadísticas – DANE.

Con el fin de certificar el cumplimiento de las buenas prácticas en la operación estadística “OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS - OEGRDP”, el IDEAM realiza la identificación de las reglas de consistencia y validación necesarias para especificar las condiciones que deben satisfacer los datos de entrada para ser considerados aceptables para su procesamiento.

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 3 de 14

Este documento tiene el propósito de presentar las reglas de consistencia y validación que aplican para los datos capturados dentro de la operación estadística de generadores de residuos o desechos peligrosos - OEGRDP que analiza el IDEAM en la etapa de crítica y que le permite determinar la calidad de la información. Además, se tienen en cuenta la calidad de la información y el cumplimiento de los criterios de validación para depurar y mejorar la calidad de los datos finales que llegan al público objetivo.

2 ALCANCE

En el presente documento se describen las reglas de validación y consistencia que aplica el IDEAM durante el proceso de análisis estadístico preliminar para los ítems relacionados con la generación y gestión de los residuos o desechos peligrosos y datos geográficos en la operación estadística de RESPEL, relacionada con la información alimentada al Sistema.

3 MARCO CONCEPTUAL

El manejo estadístico de la operación estadística de RESPEL se realiza con el software libre R que, mediante el uso de sus herramientas de programación, genera salidas útiles para el análisis de los datos de acuerdo con criterios de validación y gráficas previamente definidos. R se compila y ejecuta en una amplia variedad de plataformas; en el IDEAM se trabaja bajo Windows.

Este software, multiplataforma, lo usa el IDEAM sin requerir una licencia comercial y, por ser de código abierto, es moldeado y dirigido para el análisis específico de las operaciones estadísticas del IDEAM, tal como la OEGRDP.

En esta plataforma se programa la identificación de datos atípicos e inconsistentes en la población de datos recolectados en los diferentes registros administrativos, lo cual permite obtener una base de datos depurada que aumenta la confiabilidad y calidad de los resultados finales.

3.1 Dato atípico

En estadística un valor atípico corresponde a un dato que es considerablemente diferente a los otros datos de la muestra. Con frecuencia, los valores atípicos en un conjunto de datos pueden alertar a los estadísticos sobre las anomalías experimentales o los errores en las mediciones tomadas, y debido a esto pueden ser descartados del conjunto de datos. Si los valores atípicos del conjunto se ignoran, puede haber cambios importantes en las conclusiones obtenidas del estudio.

Para determinar los datos atípicos, el algoritmo elaborado en el lenguaje R realiza internamente el cálculo de cuartiles (Q1, Q2, Q3) y del rango intercuartílico ($RIC = Q3 - Q1$), tal como se ilustra en el diagrama de cajas y bigotes. El procedimiento identifica los datos que se encuentran por fuera de los límites superior ($Li = Q1 - 1.5RIC$) e inferior ($Ls = Q3 + 1.5RIC$) como datos atípicos y genera una consolidado para ser revisado por el líder temático que facilite la gestión con las Autoridades

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 4 de 14

Ambientales con el objetivo de mejorar la calidad de la información. (Ver

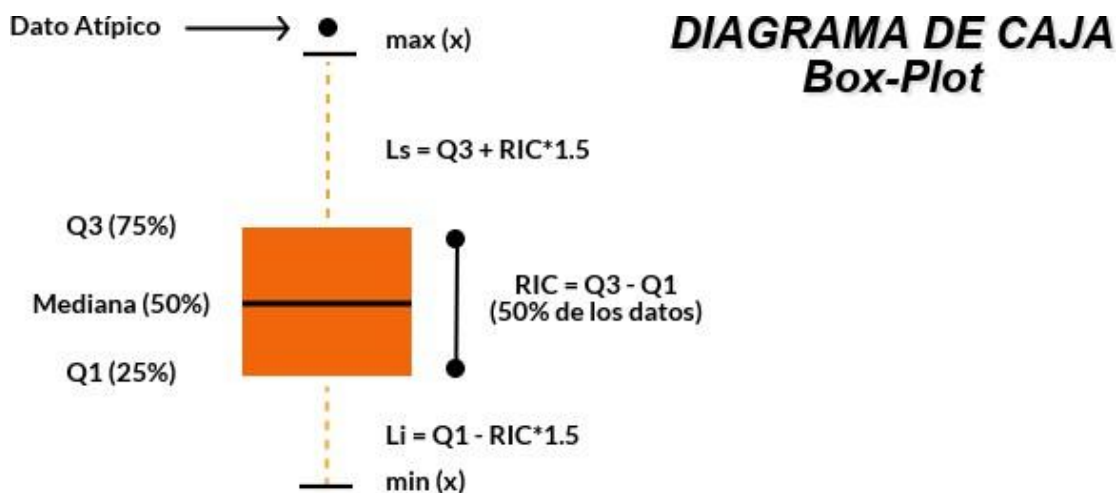


Figura 1)

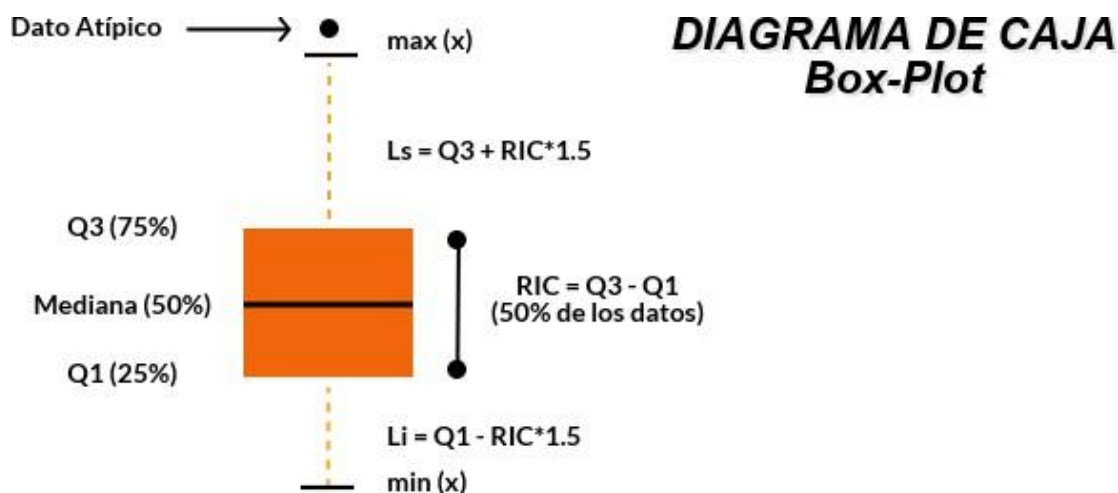


Figura 1. Ejemplo de Diagrama de caja (Box –plot)

Fuente: Elaboración propia

3.2 Dato inconsistente

Para este documento el término dato inconsistente se relaciona al incumplimiento de las reglas de validación que se han establecido para diferentes variables ambientales cuantificables. Las reglas

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 5 de 14

de validación nos permiten establecer si la operación estadística arroja información que esté de acuerdo con las leyes estadísticas y físicas relacionadas con las variables ambientales medidas.

El DANE en el documento Lineamientos Generales para el Diseño de la Operación Estadística define el proceso de validación como los valores que en forma individual pueden asumir las variables y el proceso de consistencia como las relaciones que deberían existir entre unas variables y otras (DANE, 2014).

4 OBJETIVOS

4.1 General

Establecer las reglas de validación y consistencia de los datos de RGDP que permita aumentar la confiabilidad y validez de los resultados de la operación estadística.

4.2 Específicos

- Establecer las reglas de validación previas al análisis estadístico final de los datos del RGDP.
- Identificar los datos inconsistentes durante el análisis estadístico inicial de los reportes transmitidos por la autoridad ambiental y remitidos por los establecimientos en la plataforma.
- Aumentar la confiabilidad de los datos estadísticos resultantes de la validación y detección de consistencia en el producto de difusión de la operación estadística del RGDP.

5 METODOLOGÍA DE VALIDACIÓN Y PROCESAMIENTO

El Anexo 1 presenta un esquema con la metodología para la validación y el procesamiento de los datos obtenidos mediante la plataforma del RESPEL.

6 REGLAS DE VALIDACIÓN DE LOS DATOS DEL RGDP

El proceso de validación permite al IDEAM garantizar que los resultados las operaciones estadísticas entregadas generen confianza en el público receptor y que se conviertan en instrumento de toma de decisiones para los responsables de la gestión ambiental en el país. Para la operación estadística del Registro de Generadores de Residuos o Desechos Peligrosos se establecieron reglas de validación e inconsistencias conforme a las características de las variables a determinar. A continuación, se presentan las reglas de toma de decisiones en la validación de los datos.

6.1 Datos atípicos

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 6 de 14

Para encontrar los casos atípicos se analizó la generación de establecimientos con la misma categoría de empleados (dato relacionado con el tamaño de la empresa) y por cada código CIIU (clasificación industrial internacional uniforme de todas las actividades).

Las categorías que se definieron por número de empleados, fueron las siguientes:

- 1 a 10 empleados
- 11 a 50 empleados
- 51 a 100 empleados
- 501 a 1000 empleados
- 1000 en adelante empleados

Para encontrar el dato atípico, se utilizó el test de Tukey que toma como referencia la diferencia entre el primer cuartil y el tercer cuartil o rango intercuartílico. En un diagrama de caja se considera un valor atípico el que se encuentra 1,5 veces esa distancia de uno de esos cuartiles.

Estos se calculan así:

$$Q_1 = (n + 1)/4$$

$$Q_3 = 3(n + 1)/4$$

$$RI = Q_3 - Q_1$$

Siendo:

n: Número de datos

Q₁: El primer cuartil – (percentil 25)

Q₃: El tercer cuartil – (percentil 75)

RI: El rango intercuartílico – (diferencia entre el tercer y el primer cuartil)

$$\text{Límite inferior} = Q_1 - 1,5 * RI$$

$$\text{Límite superior} = Q_3 + 1,5 * RI$$

Con los límites establecidos se define como dato atípico o anormal, aquel que se encuentra por fuera de los límites inferior y superior calculados (IDEAM, 2016).

Según la anterior clasificación, el algoritmo elaborado en R realiza la identificación de datos atípicos para cada grupo de establecimientos teniendo en cuenta la División Económica del CIIU Rev. 4 A.C y el rango del número de empleados, lo que disminuye el sesgo en las comparaciones entre datos de variables cuantitativas y aumenta la certidumbre en el proceso de identificación de valores atípicos. Las variables a las cuales se les realiza el análisis de datos atípicos son:

- **Capítulo III-Sección 3. Generación y Manejo de Residuos o Desechos Peligrosos**
 - ***Identificación de datos atípicos en la generación de residuos***

Los datos identificados como atípicos, se remiten mediante un informe a las autoridades ambientales para que sean confirmados con los establecimientos que generaron el dato, de no tener

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 7 de 14

respuesta de las autoridades sobre los datos consultados la atipicidad se excluye para obtener una base depurada por datos atípicos. Si se confirma el dato por parte de la Autoridad Ambiental entonces se tiene en cuenta dentro de los microdatos.

6.2 Datos Inconsistentes

Las inconsistencias se analizan para las sábanas del Capítulo III – Sección 3 y del Capítulo III – Sección 1, las cuales se descargan de la plataforma de RESPEL y están incluidas para su procesamiento en el código del software R para su identificación. Estas inconsistencias son inicialmente consultadas con la autoridad ambiental para su verificación y posteriormente tomar la decisión de excluir, o no, el dato del análisis estadístico.

Las inconsistencias que se verifican dentro del análisis estadístico adelantado por la entidad son:

- Eliminación de datos con dos o más separadores decimales.
- Generación total de residuos negativa.
- Identificación de establecimientos con el mismo NIT y dirección, pero diferente nombre.
- Establecimientos que reportan diferente número de empleados en diferentes registros.
- Ajuste de los nombres de los departamentos y de los municipios a aquellos aceptados oficialmente por el DANE en la DIVIPOLA. El proceso de corrección de la información errónea o diferente se realiza de forma automática.

7 REGLAS DE CONSISTENCIA DE LOS DATOS DEL RGRDP

En la siguiente tabla se resumen las consistencias de las variables que repercuten en la operación estadística del RGRDP sobre la generación y manejo de los residuos o desechos peligrosos en el país:

Tabla 1. Especificaciones de las variables que repercuten en la OE

Nombre de la variable	Tipo de dato	Longitud	Descripción/ Regla de validación
Departamento	Listado	100	Campo cerrado ¹ ; alfabético; listado desplegable; variable obligatoria.
Municipio	Listado	12	Campo cerrado; alfabético; listado desplegable; variable obligatoria.
Jurisdicción Autoridad Ambiental	Dato generado por el sistema	20	Campo cerrado; alfabético.

¹ **Campo cerrado:** definido por el sistema a través de un listado.

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 8 de 14

Nombre de la variable	Tipo de dato	Longitud	Descripción/ Regla de validación
Código CIU/Descripción	Listado	20	Campo cerrado; alfanumérico; listado desplegable; variable obligatoria.
Código CPC/ Descripción	Listado	20	Campo cerrado; alfanumérico; listado desplegable; variable obligatoria.
Cantidad total Consumida/ Unidad de medida	Number	Por definir (Sistemas)	Campo abierto ² ; numérico; variable obligatoria.
Código de la corriente del residuo o desechos peligroso	Listado	20	Campo cerrado; alfanumérico; listado desplegable; variable obligatoria.
Unidad de medida estado	Listado	Por definir (Sistemas)	Campo cerrado; alfanumérico; listado desplegable; variable obligatoria.
Cantidad almacenada por el generador al inicio	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad almacenada por el generador al final	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad almacenada por el generador consolidada	Dato generado por el sistema	Por definir (Sistemas)	Campo cerrado; numérico.
Cantidad almacenada a través de un tercero al inicio	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad almacenada a través de un tercero al final	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad almacenada por el tercero consolidada	Dato generado por el sistema	Por definir (Sistemas)	Campo cerrado; numérico.
Cantidad aprovechada y/o valorizada por el generador al inicio	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad aprovechada y/o valorizada por el generador al final	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad aprovechada y/o valorizada por el generador consolidada	Dato generado por el sistema	Por definir (Sistemas)	Campo cerrado; numérico.

² **Campo abierto:** diligenciado por el usuario según las especificaciones del sistema.

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 9 de 14

Nombre de la variable	Tipo de dato	Longitud	Descripción/ Regla de validación
Cantidad aprovechada y/o valorizada a través de un tercero al inicio	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad aprovechada y/o valorizada a través de un tercero al final	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad aprovechada y/o valorizada a través de un tercero consolidada	Dato generado por el sistema	Por definir (Sistemas)	Campo cerrado; numérico.
Tipo de aprovechamiento	Listado	Por definir (Sistemas)	Campo cerrado; alfabético; listado desplegable; variable obligatoria.
Cantidad tratada por el generador al inicio	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad tratada por el generador al final	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad tratada por el generador consolidada	Dato generado por el sistema	Por definir (Sistemas)	Campo cerrado; numérico.
Cantidad tratada a través de un tercero al inicio	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad tratada a través de un tercero al final	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad tratada a través de un tercero consolidada	Dato generado por el sistema	Por definir (Sistemas)	Campo cerrado; numérico.
Tipo de tratamiento	Listado	Por definir (Sistemas)	Campo cerrado; alfabético; listado desplegable; variable obligatoria.
Cantidad dispuesta por el generador al inicio	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad dispuesta por el generador al final	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad dispuesta por el generador consolidada	Dato generado por el sistema	Por definir (Sistemas)	Campo cerrado; numérico.
Cantidad dispuesta a través de un tercero al inicio	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 10 de 14

Nombre de la variable	Tipo de dato	Longitud	Descripción/ Regla de validación
Cantidad dispuesta a través de un tercero al final	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.
Cantidad dispuesta a través de un tercero consolidada	Dato generado por el sistema	Por definir (Sistemas)	Campo cerrado; numérico.
Tipo de disposición	Listado	Por definir (Sistemas)	Campo cerrado; alfabético; listado desplegable; variable obligatoria.
Cantidad total de residuos o desechos peligrosos generados durante el período de balance	Number	Por definir (Sistemas)	Campo abierto; numérico; variable obligatoria. Criterio de Validación: Mayor o igual a cero.

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 11 de 14

GLOSARIO

Crítica estadística: es un proceso en el que se depura el conjunto de datos durante su recolección y procesamiento, aplicando unas reglas que permitan garantizar que dichos datos estén completos, sean correctos y consistentes (DANE, 2014). Las reglas se aplican para los siguientes tres procesos.

Proceso de validación: define los valores que en forma individual pueden asumir las variables (DANE, 2014).

Proceso de consistencia: establece las relaciones que deberían existir entre unas variables y otras (DANE, 2014).

Autoridad Ambiental - AA. Se refiere a las Corporaciones Autónomas Regionales y las Autoridades Ambientales de los grandes centros urbanos y a las que se refiere el artículo 13 de la Ley 768 del 2002. (Congreso, 1993).

Operación estadística: Aplicación de un proceso estadístico sobre un objeto de estudio que conduce a la producción de información estadística.

Proceso estadístico: Conjunto sistemático de actividades encaminadas a la producción de estadísticas que comprende, entre otras, la detección de necesidades, el diseño, la recolección, el procesamiento, el análisis y la difusión.

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 12 de 14

BIBLIOGRAFÍA

Congreso. (1993). *Ley 99 de 1993*. Por la cual se crea el Ministerio del Medio Ambiente, se reordena el Sector Público encargado de la gestión y conservación del medio ambiente y los recursos naturales renovables, organiza el Sistema Nacional Ambiental –SINA, dictan otras disposiciones, Bogotá D.C.

DANE. (2014). *Lineamientos generales para el diseño de la operación estadística*. Departamento Administrativo Nacional de Estadísticas, Bogotá D. C.

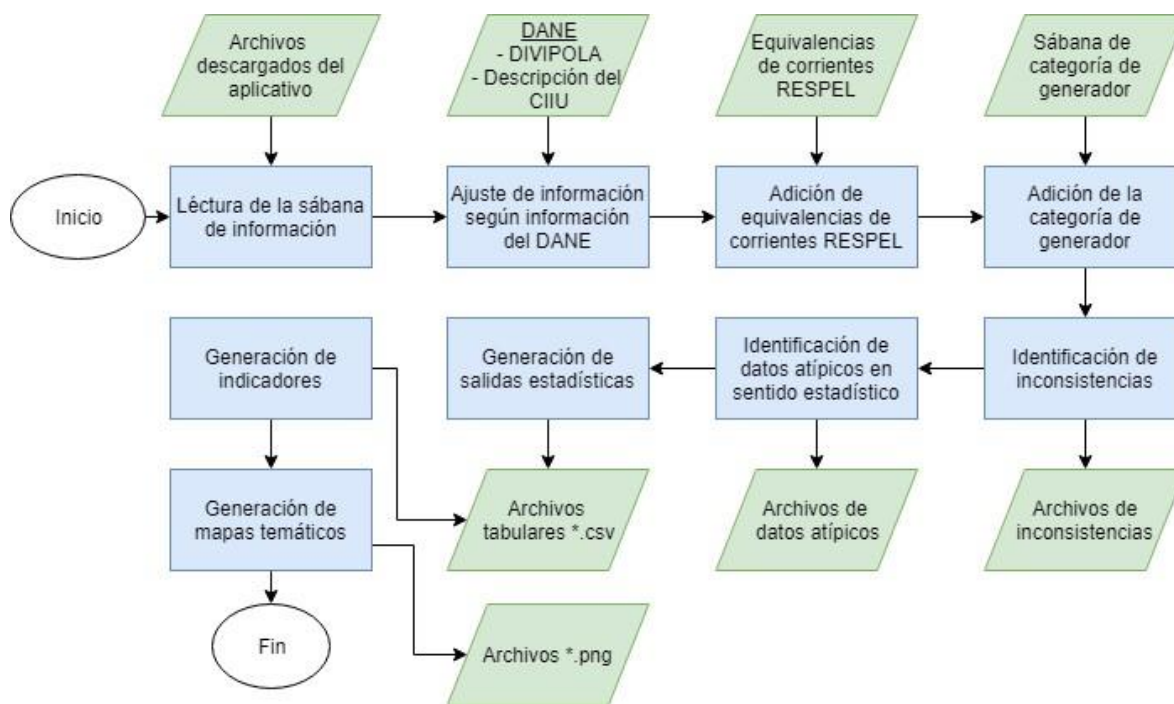
IDEAM. (2016). *Informe Nacional de Residuos o Desechos*. Bogotá.

MAVDT. (1994). *Decreto 1600 de 1994*. Ministerio de Ambiente, Vivienda y Desarrollo Territorial.

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 13 de 14

ANEXOS

Anexo 1. Esquema de la metodología de procesamiento y validación de los datos en la operación estadística de RESPEL



	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICA DE GENERADORES DE RESIDUOS O DESECHOS PELIGROSOS	Código: GCI-RP-M001
		Versión: 02
		Fecha: 11/07/2025
		Página 14 de 14

HISTORIAL DE CAMBIOS

VERSIÓN	FECHA	DESCRIPCIÓN
01	24/10/2018	Creación del documento
02	11/07/2025	Se actualiza el Formato de acuerdo con el memorando enviado por la OAP memorando 20251100097283 lineamientos para la actualización documental en el marco de la implementación del aplicativo suite visión. Pasa de M-GCI-EA-M021 a GCI-RP-M001.

ELABORÓ: Juan Carlos Ariza Porras Contratistas de Estudios Ambientales	REVISÓ: Ana M. Hernández Coordinadora del Grupo de Seguimiento a la Sostenibilidad del Desarrollo Adriana María Zapata Profesional Especializado Subdirección de Estudios Ambientales	APROBÓ: Diana Marcela Vargas Subdirectora de Estudios Ambientales
--	--	---