




IDEAM

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

**REGLAS DE VALIDACIÓN
ESTADÍSTICA DE MONITOREO Y SEGUIMIENTO DEL RUA MANUFACTURERO – EMSRUAM**

Diciembre de 2017

 <p>IDEAM Instituto de Hidrología, Meteorología y Estudios Ambientales</p>	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICAS DEL MONITOREO Y SEGUIMIENTO DEL RUA MANUFACTURERO - EMSRUAM	Código: M-GCI-EA-M011
		Versión: 01
		Fecha: 14/12/2017
		Página 2 de 10

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.

El Registro Único Ambiental para el sector manufacturero - RUAM, captura información a través de registros administrativos de los diferentes temas ambientales que lo conforman tales como; recurso agua, energía, emisiones a la atmósfera, recursos Naturales que son sujetos a salvoconducto o remisión de movilización y residuos o desechos no peligrosos que diligencian las empresas manufactureras que se encuentran dentro del alcance del ámbito de estudio.


El RUAM, es un instrumento de captura para el Subsistema de Información sobre Uso de Recursos Naturales Renovables – SIUR que hace parte del Sistema de Información Ambiental para Colombia –SIAC. Este Subsistema permite contar con información normalizada homogénea y sistemática sobre el uso y/o aprovechamiento de los recursos naturales renovables, originado por las actividades manufactureras.

El Ministerio de Ambiente adoptó el protocolo para el monitoreo y seguimiento del Subsistema de Información sobre Uso de Recursos Naturales Renovables-SIUR para el sector manufacturero mediante la Resolución 1023 de 2010, en donde se estableció como ámbito de aplicación a los establecimientos cuya actividad productiva principal se encuentre incluida en la Sección C – Industrias Manufactureras, divisiones 10 a 33 de la Clasificación Industrial Internacional Uniforme – CIU, Revisión 4.0 adaptada para Colombia por el Departamento Administrativo Nacional de Estadística – DANE, o aquella que la modifique o sustituya, que de acuerdo a la normativa ambiental vigente, requieran de licencia ambiental, plan de manejo ambiental, permisos, concesiones, y demás autorizaciones ambientales, así como aquellas actividades que requieran de registros de carácter ambiental¹.

El aplicativo web del RUAM permite el acopio, almacenamiento, procesamiento, análisis y consulta de indicadores e información sobre el uso y/o aprovechamiento de los recursos naturales renovables, que reportan las actividades manufactureras. El subsistema procesa información por periodos de balance, que comprende el periodo entre el 1 de enero al 31 de diciembre del año inmediatamente anterior al año de diligenciamiento, sobre la información de las entradas (agua, energía, materias primas) y las salidas (energía, vertimientos, emisiones a la atmósfera productos y residuos) de cada actividad inscrita que reporta al RUAM.

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

¹ Se debe tener en cuenta que en el Art 3. de la Resolución 1023 de 2010 se indica que el alcance corresponde a Sección D – Industrias Manufactureras, divisiones 15 a 37, Revisión 3.0 adaptada para Colombia.

 <p>IDEAM Instituto de Hidrología, Meteorología y Estudios Ambientales</p>	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICAS DEL MONITOREO Y SEGUIMIENTO DEL RUA MANUFACTURERO - EMSRUAM	Código: M-GCI-EA-M011
		Versión: 01
		Fecha: 14/12/2017
		Página 3 de 10

Con el fin de certificar el cumplimiento de las buenas prácticas en la operación estadística “OPERACIÓN ESTADÍSTICAS DEL MONITOREO Y SEGUIMIENTO DEL RUA MANUFACTURERO - EMSRUAM”, 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.

Este documento tiene el propósito de presentar las reglas de inconsistencia y validación que aplican para los datos de cada matriz ambiental de los registros de cada usuario 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 ítem relacionados con calidad de recursos naturales y datos geográficos en la operación estadística del RUAM, relacionada con la información alimentada al Sistema.

3. MARCO CONCEPTUAL

El manejo estadístico del RUAM se realiza con el software libre R Project que genera el análisis de los datos conforme las validaciones que se le programen y los gráficos estadísticos requeridos. R se compila y ejecuta con una amplia variedad de plataformas, en el IDEAM se trabaja bajo Windows.

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

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 software R realiza internamente el cálculo de cuartiles (Q1, Q2, Q3) y del rango intercuartílico ($RIC = Q3 - Q1$) con base en el diagrama de cajas y bigotes. R identifica los datos que se encuentran por fuera de los límites superior ($Ls = Q3 + 1.5RIC$) e inferior ($Li = Q1 - 1.5RIC$) como datos atípicos y suprimiéndolos de la población de datos. (Ver

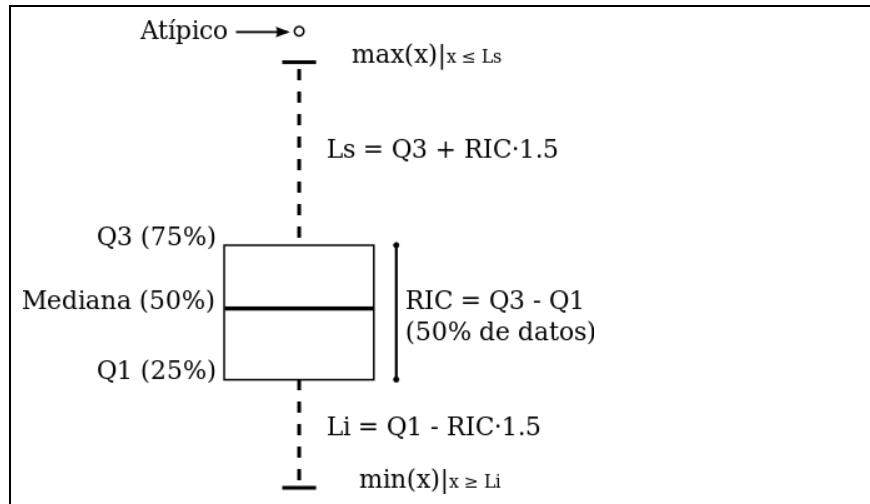


Figura 1. Ejemplo de Diagrama de caja (Box –plot)

Fuente: Fundación Wikimedia, Inc.

https://es.wikipedia.org/wiki/Diagrama_de_caja

)

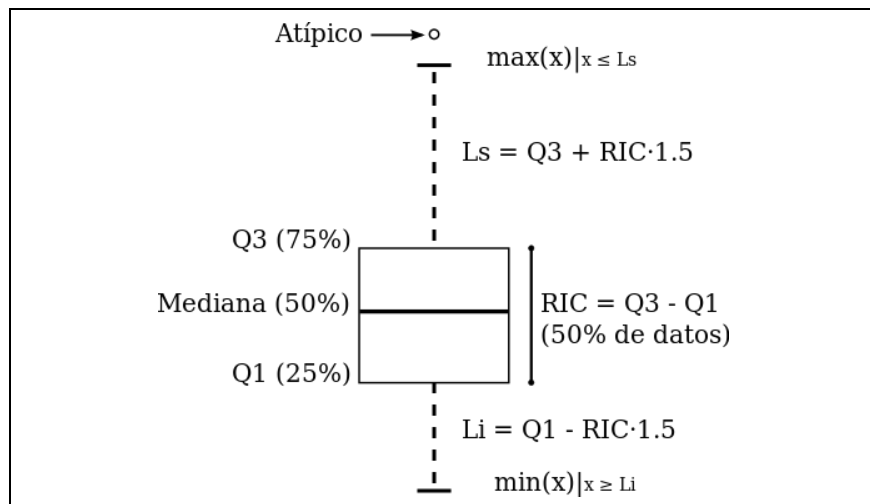



Figura 1. Ejemplo de Diagrama de caja (Box –plot)

Fuente: Fundación Wikimedia, Inc.

https://es.wikipedia.org/wiki/Diagrama_de_caja

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

 <p> IDEAM Instituto de Hidrología, Meteorología y Estudios Ambientales </p>	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICAS DEL MONITOREO Y SEGUIMIENTO DEL RUA MANUFACTURERO - EMSRUAM	Código: M-GCI-EA-M011
		Versión: 01
		Fecha: 14/12/2017
		Página 5 de 10

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 del RUAMF 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 RUAM, de acuerdo a las características de cada componente ambiental.
- Identificar los datos inconsistentes durante el análisis estadístico inicial de los datos transmitidos por la autoridad ambiental y remitidos por los establecimientos al registro.
- 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 RUAM.

5. REGLAS DE VALIDACIÓN Y CONSISTENCIA DE DATOS DEL RUAM

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

5.1 Datos atípicos


Para establecer datos atípicos en la serie de datos por periodo de balance se tuvo en cuenta las características semejantes de grupos de actividades manufactureras tales como el tamaño de la empresa y la división económica según el CIU.

Para el análisis estadístico por cuartiles, para determinar los datos atípicos según el análisis de caja y bigotes, el software R realiza inicialmente una distribución de las actividades por grupos según el tamaño de empresa, por número de empleados, y por división económica del CIU.

Para el RUAM se determinaron inicialmente 6 grupos de establecimientos teniendo la siguiente distribución:

Tabla 1. Distribución de empresas por número de empleados.

Grupo	Número de empleados
-------	---------------------

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICAS DEL MONITOREO Y SEGUIMIENTO DEL RUA MANUFACTURERO - EMSRUAM	Código: M-GCI-EA-M011
		Versión: 01
		Fecha: 14/12/2017
		Página 6 de 10

1	0-10
2	11-50
3	51 – 100
4	101 -500
5	501 – 1000
6	Mayor de 1000

Para continuar la subdivisión de los grupos y aumentar la especificidad de las características a comparar se tiene en cuenta la división de la sección C del CIU.

Según la anterior clasificación el software R realiza la identificación de datos atípicos para cada grupo de establecimientos según la División Económica de la sección C del CIU Rev. 4 A.C, lo que permite establecer comparaciones con un bajo error de magnitud y de características de datos atípicos para variable cuantitativas a analizar. Las variables a las cuales se les realiza el análisis de datos atípicos son:

Capítulo III. Recurso Agua

- Atípicos para el volumen captado
- Atípicos para el volumen vertido
- Atípicos para DBO
- Atípicos para SST

Capítulo IV. Energía

- Atípicos para consumo de energía

Capítulo V. Emisiones a La Atmósfera

- Atípicos para flujo másico

Capítulo VIII Residuos o Desechos

- Atípicos para residuos no peligrosos


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 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.

5.2 Datos Inconsistentes

Las inconsistencias determinadas aplican para los capítulos relacionados con los recursos agua, energía y residuos o desechos no peligrosos, la cuales fueron incluidas en el código del software R para su identificación. Estas inconsistencias son inicialmente consultadas con la autoridad ambiental para su verificación para posteriormente tomar la decisión de excluir, o no, el dato del análisis estadístico.

Capítulo III. Recurso Agua

Este capítulo permite recolectar, en la sección mediciones, los resultados de las caracterizaciones fisicoquímicas del recurso agua que han sido solicitado por las autoridades ambientales, según

	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICAS DEL MONITOREO Y SEGUIMIENTO DEL RUA MANUFACTURERO - EMSRUAM	Código: M-GCI-EA-M011
		Versión: 01
		Fecha: 14/12/2017
		Página 7 de 10

corresponda (Entrada y/o vertimiento), que permiten realizar la verificación de la coherencia de la variable Volumen de Agua a través de la siguiente regla:

- Identificación de registros en los cuales se encuentra que el agua vertida excede en más del 10% al agua captada y que el volumen de agua tratada es 0 m³

Resultados que cumplan con este criterio de validación son excluidos del análisis estadístico.

Capítulo IV. Energía.

En este capítulo el establecimiento reporta los consumos de energía según su origen de forma que se establece la cantidad total de Consumo de Energía Eléctrica. Esta variable corresponde a la sumatoria de:

$$\text{Total Consumo Energía Eléctrica} = \text{EEC} + \text{EERT} + \text{EEG} - \text{EEV} - \text{EECT}$$

Donde,

- **EEC:** Energía Eléctrica Comprada
- **EERT:** Energía Eléctrica Recibida en Transferencia
- **EEG:** Energía Eléctrica Generada
- **EEV:** Energía Eléctrica Vendida
- **EECT:** Cedida en Transferencia

De acuerdo a esto se establecieron las siguientes posibles inconsistencias para la población de datos:

- Identificación de registros en los cuales se encuentran los mismos valores para cantidad de energía, recibida, generada, consumida y transferida
- Energía comprada menor que la energía vendida

Capítulo VIII A. Información sobre generación, manejo y existencias de residuos o desechos no peligrosos

En este capítulo se involucran diferentes variables con el fin de generar el balance de materia y determinar la cantidad total de residuos o desechos no peligrosos generada en el Periodo de Balance. Teniendo en cuenta los criterios de balance se establecieron los siguientes criterios técnicos de inconsistencias:

- Identificación de registros de la sábana de residuos no peligrosos para los cuales se encuentran valores iguales en distintos tipos de generación (Almacenamiento, aprovechamiento, tratamiento y disposición final).




**REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA
OPERACIÓN ESTADÍSTICAS DEL MONITOREO Y
SEGUIMIENTO DEL RUA MANUFACTURERO -
EMSRUAM**

Código: M-GCI-EA-M011

Versión: 01

Fecha: 14/12/2017

Página **8** de **10**

 <p>IDEAM Instituto de Hidrología, Meteorología y Estudios Ambientales</p>	REGLAS DE VALIDACIÓN Y CONSISTENCIA DE LA OPERACIÓN ESTADÍSTICAS DEL MONITOREO Y SEGUIMIENTO DEL RUA MANUFACTURERO - EMSRUAM	Código: M-GCI-EA-M011
		Versión: 01
		Fecha: 14/12/2017
		Página 9 de 10

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 ESTADÍSTICAS DEL MONITOREO Y SEGUIMIENTO DEL RUA MANUFACTURERO - EMSRUAM	Código: M-GCI-EA-M011
		Versión: 01
		Fecha: 14/12/2017
		Página 10 de 10

BIBLIOGRAFIA

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*. Departamenteo Administrativo Nacional de Estadísticas, Bogotá D. C.

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

HISTORIAL DE CAMBIOS

VERSIÓN	FECHA	DESCRIPCIÓN
01	14/12/2017	Creación del documento

ELABORÓ: Tania M. Carpio Galván Contratistas de Estudios Ambientales	REVISÓ: German D. García / Ana M. Hernández Profesional / Coordinadora del Grupo de Seguimiento a la Sostenibilidad del Desarrollo	APROBÓ: Diana Marcela Vargas Subdirectora de Estudios Ambientales
--	--	---