



bluewater

CONTROL, TRATAMIENTO Y REDUCCIÓN DE MICRO PLÁSTICOS
Y CONTAMINANTES EMERGENTES EN AGUAS RESIDUALES
URBANAS Y EN EL MEDIO COSTERO FRONTERIZO

INFORME FINAL DEL EJERCICIO INTERLABORATORIO

ACTIVIDAD 1: Desarrollo de herramientas para la
evaluación de microplásticos y CEC en las EDAR y
el medio acuático.

Interreg



Cofinanciado por
la Unión Europea
Cofinanciado pela
União Europeia

España – Portugal





Actividad	1. Desarrollo de herramientas para la evaluación de microplásticos y CEC en las EDAR y el medio acuático
Acción	1.3 Ejercicio interlaboratorio para la validación de metodología analítica para la determinación de compuestos indicadores en aguas residuales
Fecha de actualización	15/06/2026
Versión	1
Autores	INTECMAR
Participantes	USC, AEdP, IEO





Índice

1. INTRODUCCIÓN	4
2. OBJETIVOS	4
3. ASEGURAMIENTO DE CALIDAD.....	5
4. ORGANIZACION DEL PROGRAMA	5
4.1 Gestión y planificación	5
4.2 Confidencialidad	5
4.3 Etapas	6
4.4 Sistema de gestión.....	7
4.5 Subcontratistas	7
5. MATERIALES DE ENSAYO	7
5.1. Preparación de las muestras de ensayo	7
5.1.a. Identificación de las muestras	9
5.1.b. Manipulación y almacenamiento	9
5.1.c. Embalaje y etiquetado	9
5.2 Envío y recepción de los materiales de ensayo.....	10
5.3. Análisis de los materiales de ensayo	10
5.4. Envío de resultados de los participantes.....	11
5.4.a. Colusión y falsificación de resultados	11
5.5. Emisión de informes	12
6. ESTUDIO DE HOMOGENEIDAD Y ESTABILIDAD.....	12
6.1. Ensayos de homogeneidad.....	12
6.2. Ensayos de estabilidad	13
7. TRATAMIENTO ESTADÍSTICO DE LOS RESULTADOS	17
7.1. Importantes consideraciones previas	17
7.2. Tratamiento de los datos inconsistentes o aberrantes.....	17
7.2.a. Test de Grubbs.	18
7.2.b. Test estadísticos robustos basados en la mediana.....	18
7.3. Normalidad estadística de los datos	19
7.3.a. Test de Kernel.....	19
7.4. Establecimiento del valor asignado.....	20
7.5. Incertidumbre estándar del valor asignado	21
7.6. Desviación estándar del ensayo de aptitud	22



8. RESUMEN ESTADISTICO DE LOS RESULTADOS.....	23
8.1. Resumen de estadísticas	23
8.2. Clasificación general de puntuación z-score de los laboratorios	24
9. EVALUACIÓN DEL DESEMPEÑO	26
10. VARIABLES METODOLÓGICAS Y DESEMPEÑO.....	27
10.1. Evaluación de significancia estadística	28
11. DISCUSIÓN DE RESULTADOS.....	32
12. CONCLUSIONES	33
13. DEFINICIONES	36
14. REFERENCIAS	37
15. ANEXOS	38
ANEXO I. DATOS PRELIMINARES. RESUMEN	38
ANEXO II. GRAFICAS DE PRECISIÓN (RSD%)	41
ANEXO III. GRÁFICAS DE RESULTADOS Z-SCORE DE LOS LABORATORIOS	43
ANEXO IV. RESUMEN DATOS TÉCNICOS DE ENSAYO POR LABORATORIO	79





1. INTRODUCCIÓN

El proyecto BlueWWater tiene como objetivo la mejora de la calidad de las masas de agua fluviales, de transición y costeras mediante el control, seguimiento y evaluación de las emisiones de microplásticos y contaminantes de preocupación emergente (CEC) en el medio acuático. Para ello, se estudia la eficiencia de las estaciones de tratamiento de aguas residuales urbanas (EDAR) en las regiones de Galicia y el norte de Portugal, así como el riesgo ambiental asociado a estos contaminantes, con el fin de favorecer el uso sostenible de los recursos hídricos y contribuir a la implementación de la normativa comunitaria.

En este contexto, el proyecto desarrollará metodologías analíticas automatizadas de alto rendimiento (Actividad 1) que permitan mejorar la monitorización de estos contaminantes en las masas de agua.

Para evaluar la eficacia de eliminación de compuestos indicadores en las EDAR, tal como establece la Directiva sobre el tratamiento de aguas residuales urbanas, es necesario disponer de metodologías analíticas validadas para estas determinaciones. En este sentido, los laboratorios deben implantar sistemas de control de calidad en los que la participación en ensayos de aptitud constituye un requisito altamente recomendable y, en algunos casos, obligatorio.

Con este objetivo, se ha organizado una comparación interlaboratorios para la determinación de compuestos indicadores en aguas residuales, con la participación de 12 laboratorios.

El protocolo de desarrollo del ejercicio se ha basado en guías y documentos internacionalmente reconocidos, entre ellos la norma UNE-EN ISO/IEC 17043:2023 sobre requisitos generales para ensayos de aptitud, la norma ISO 13528:2022 sobre métodos estadísticos aplicables a comparaciones interlaboratorios, y el *International Harmonised Protocol for the Proficiency Testing of Analytical Chemistry Laboratories* (IUPAC Technical Report, Pure and Applied Chemistry, 78(1), 145–196, 2006)

2. OBJETIVOS

El siguiente Informe tiene como objetivo presentar la sistemática establecida por INTECMAR, como organizador del ejercicio, para el correcto desarrollo del ensayo de aptitud entre aquellos laboratorios participantes, las muestras analizadas en las distintas matrices y los resultados obtenidos.

El ejercicio interlaboratorio BlueWWater, es una iniciativa cuyo objetivo es servir de herramienta para evaluar y validar los métodos analíticos utilizados por los laboratorios para detectar contaminantes emergentes en aguas residuales y superficiales. Está diseñado para ayudar a los participantes a mejorar sus técnicas, identificar áreas de mejora y reforzar la confianza en los resultados obtenidos.

Realizado conforme a las directrices de la norma ISO 13528:2022 y del Protocolo internacional armonizado para las pruebas de aptitud de los laboratorios de química



analítica de la IUPAC, el ejercicio tiene como objetivo la determinación de ciertas sustancias farmacéuticas y cosméticas en aguas residuales y del medio receptor. Estas sustancias, incluso en concentraciones muy bajas, pueden representar un riesgo para el medio ambiente y la salud pública. A través de este ejercicio, buscamos proporcionar a los laboratorios las herramientas necesarias para cumplir con los nuevos requisitos establecidos por la Directiva sobre el tratamiento de aguas residuales urbanas (DARU).

3. ASEGURAMIENTO DE CALIDAD

Este ensayo de aptitud está basado fundamentalmente en los requisitos detallados en la norma internacional UNE-EN ISO 17043.

El programa cuenta con un grupo de trabajo dirigido por los coordinadores del proyecto en INTECMAR y asesorado por profesionales del ámbito público y privado, seleccionados según su experiencia profesional previamente acreditada.

Este grupo se reunió, como mínimo una vez al mes, con el fin de colaborar en el control y desarrollo del Protocolo general y de esta parte del proyecto BlueWWater.

4. ORGANIZACION DEL PROGRAMA

4.1 Gestión y planificación

La gestión de este ensayo de aptitud/interlaboratorios se realiza de acuerdo con principios internacionalmente aceptados, en particular los establecidos en el *International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry* de IUPAC.

El programa, calendario y solicitud de los ítems de ensayo se planifican con suficiente antelación (último trimestre de 2024), garantizando que dichos ítems se preparan conforme a lo establecido.

La planificación incluye, como mínimo, información relativa a la obtención, procesamiento y distribución de los ítems, las fechas previstas de inicio y finalización, la información que se remitirá a los participantes, el grado de difusión de los resultados y su origen.

La organización cuenta con los conocimientos técnicos y la experiencia necesarios para la planificación y desarrollo de ensayos interlaboratorio.

El seguimiento rutinario del programa y la responsabilidad final del ensayo de aptitud corresponden a los coordinadores del proyecto. En caso de recurrir a expertos externos, no se revelará información relativa a los participantes.

4.2 Confidencialidad

Toda la información obtenida sobre los participantes, incluidas sus valoraciones de desempeño y z-score, es confidencial y no se revelará a terceros, salvo autorización expresa del participante para los fines concretos del Proyecto BlueWWater.

Para preservar dicha confidencialidad, los participantes reciben los informes de resultados de este ejercicio interlaboratorio sin identificación de los laboratorios individuales.



Cualquier uso de datos contenidos en los informes en publicaciones externas requerirá autorización previa. Asimismo, la distribución de publicaciones que exploten dichos datos deberá contar con consentimiento, conocimiento o aprobación previos.

4.3 Etapas

El proceso típico de realización de un ensayo de aptitud es resumidamente el siguiente:

- a) Preparación de las muestras de ensayos
- b) Ensayos de homogeneidad en laboratorios de referencia.
- c) Envío de los materiales de ensayo en la fecha anunciada.
- d) Los laboratorios participantes analizan las muestras de ensayo y envían los resultados en una fecha determinada.
- e) Evaluación estadística de los resultados.
- f) Envío del informe final a todos participantes.

Si existe un retraso en alguna de estas etapas se informa por correo electrónico a los participantes.

Esquemáticamente, el programa de ensayos de aptitud/interlaboratorios se organiza de acuerdo con las siguientes etapas:





Etapa 6 Análisis de resultados:

Evaluación y comparación de resultados de los participantes.

Evaluación estadística y desempeño

Etapa 7 Informe de resultados:

Envío del Informe de resultados a los participantes.

4.4 Sistema de gestión

Este ensayo de aptitud/interlaboratorios está gestionado de acuerdo con la sistemática y requisitos establecidos en la norma UNE-EN ISO 17043:2023 y el estudio de los resultados está basado en la estadística robusta definida en la norma ISO 13528:2022.

4.5 Subcontratistas

La planificación del programa de ensayo de aptitud/interlaboratorios, la evaluación del desempeño y la autorización del informe final no son objeto de subcontratación. No obstante, para este ejercicio se cuenta con el asesoramiento de una consultoría contratada específicamente en el marco del Proyecto BlueWWater, conforme a lo aprobado en el formulario de candidatura.

Cuando se subcontraten servicios analíticos a entidades externas con fines de verificación o evaluación de los ítems de ensayo, estas deberán disponer de la competencia técnica adecuada y, cuando corresponda, estar acreditadas conforme a la norma ISO 17025 o a las especificaciones aplicables en cada caso.

La preparación de los materiales de ensayo se realiza internamente. Los ensayos de homogeneidad y estabilidad son efectuados por un laboratorio designado distinto del encargado de la preparación y distribución del material de ensayo.

El laboratorio designado podrá participar igualmente en el ensayo de aptitud/interlaboratorios. En tal caso, su participación será tratada en igualdad de condiciones respecto al resto de participantes y quedará sujeta a las mismas normas de confidencialidad.

5. MATERIALES DE ENSAYO

5.1. Preparación de las muestras de ensayo

Los siguientes materiales de ensayo distribuidos para su análisis entre los participantes fueron preparados por INTECMAR:

- 1) Agua de medio receptor: agua superficial de río. Recogida en el río Ulla, provincia de Pontevedra, España.





- 2) Agua residual influente: agua recogida en una estación depuradora después del desbaste, antes del tratamiento primario.
- 3) Agua residual efluente: agua recogida en una estación depuradora a la salida del tratamiento UV, antes de su vertido al medio.

Las muestras fueron recogidas en bidones de polietileno de 30 litros de capacidad. A su llegada al laboratorio fueron inmediatamente procesadas. Se filtraron empleando filtros de fibra de vidrio de 90 mm de diámetro y 2.0 µm de tamaño de poro, recogiendo el filtrado en otro bidón provisto de llave en su parte inferior. Se homogeneizaron durante 40 min mediante un agitador de varilla a 1600 rpm. Transcurrido ese tiempo y sin dejar de agitar, se rellenaron botellas de 1 litro, previamente etiquetadas, mediante dispensación a través de la llave inferior. Se conservaron en congelador a -20 °C hasta el momento de su distribución.

En el caso de la muestra de agua de río se añadió una mezcla de los compuestos a determinar. A partir de las sustancias sólidas se preparó en metanol una disolución concentrada, de la cual se tomó 1 ml que se adicionó al inicio de la etapa de homogeneización. Las concentraciones resultantes en el agua son las siguientes:

TABLA 1: Compuestos adicionados al agua de río

Compuesto adicionado	Concentración en agua de medio receptor (ng/l)
5-Methylbenzotriazole (98%)	119,4
Amisulpride	81,3
Benzotriazole	99,4
Candesartan cilexetil*	59,4
Carbamazepine	26,3
Citalopram hydrobromide	139,4
Clarithromycin	93,8
Diclofenac sodium	255,0
Hydrochlorothiazide	230,6
Irbesartan	151,9
Metoprolol tartrate	96,9
Venlafaxine hydrochloride	96,9

**Observación:* en la preparación se utilizó inadvertidamente candesartán cilexetil en lugar de candesartán, que era el analito objetivo. Esta desviación implica que la fortificación no fue estrictamente equivalente a una adición directa del compuesto libre, ya que el cilexetil puede hidrolizarse a candesartán y presentar además diferencias de solubilidad, estabilidad y recuperación. En consecuencia, el resultado obtenido para este compuesto en esta muestra debe considerarse afectado por una desviación de preparación y valorarse con cautela.





5.1.a. Identificación de las muestras

Se adjudica un código de identificación para cada muestra preparada. Se genera una etiqueta de identificación de los ítems destinados al ensayo de aptitud, que servirán para garantizar su trazabilidad durante el desarrollo de la ronda. Cada botella es identificada con un código único.



5.1.b. Manipulación y almacenamiento

Se ha garantizado el control de acceso a los materiales de ensayo mediante su correcta identificación y la aplicación de condiciones adecuadas de almacenamiento y seguridad, con el fin de preservar su integridad y evitar contaminaciones o degradaciones.

La manipulación y conservación de los ítems de ensayo se realizó conforme a lo establecido en el protocolo general.

Asimismo, el estado de los ítems fue evaluado periódicamente durante el periodo de almacenamiento para detectar posibles deterioros.

Todas estas actividades se llevaron a cabo bajo la supervisión y control del responsable designado para dichas tareas.

5.1.c. Embalaje y etiquetado

El proceso de embalaje y etiquetado de los ítems de ensayo ha sido realizado por personal autorizado atendiendo a lo descrito en el protocolo general.

La secuencia de tareas principales han sido las siguientes:

- ✓ Estimar la cantidad de muestra necesaria y suficiente para la ronda, teniendo en cuenta:
 - *Número de laboratorios participantes*
 - *Ensayos de homogeneidad y estabilidad*
 - *Ensayos de control de estabilidad a largo plazo*
 - *Cantidad de muestras sobrantes para reenviar en caso necesario*
 - *Posibles pérdidas, contaminación o deterioros de los ítems*
- ✓ Muestreo previo para asegurar su fiabilidad y representatividad
- ✓ Embotellado de las muestras en cantidad suficiente
- ✓ Etiquetado de identificación de los ítems de ensayo



- ✓ Segregación al azar de 10 muestras ya preparadas para ensayos de homogeneidad y estabilidad a realizar en el laboratorio designado.
- ✓ Segregación al azar de los ítems de ensayo a enviar a cada participante, tras la confirmación de la homogeneidad.
- ✓ Preparación de los envíos correspondientes.

Se controlan los procesos de embalaje y etiquetado, cuando es necesario, para asegurar la conformidad con los requisitos de seguridad y transporte pertinentes tanto nacionales, regionales o internacionales.

Se ha establecido un número mínimo de 7 participantes por parámetro para desarrollar el ensayo de aptitud con tratamiento estadístico. Por debajo de 7 resultados no se realizará ninguna evaluación estadística ni de desempeño.

Se ha solicitado a los laboratorios participantes analizar todos los parámetros en todos los materiales de ensayo que integran la ronda correspondiente.

5.2 Envío y recepción de los materiales de ensayo

Se realizaron los procesos de embalaje y etiquetado de los materiales de ensayo, así como la gestión del transporte nacional e internacional. En todos los casos se controlaron las condiciones de transporte, la sistemática de envío y la correcta entrega dentro de los plazos previstos.

El embalaje y etiquetado de los materiales de ensayo se realizó conforme a lo descrito en el apartado 4.2.c del Protocolo General. Una vez completado este proceso, se procedió al envío de las muestras.

Los participantes fueron responsables de comunicar la correcta recepción de las muestras solicitadas y, en caso de incidencias, retrasos o deficiencias en su estado, informar de dichas circunstancias a la organización.

Cuando, por cualquier motivo, se produjo un retraso en el envío de las muestras, los participantes fueron informados por correo electrónico con la mayor antelación posible y siempre antes de la fecha inicialmente programada.

5.3. Análisis de los materiales de ensayo

Los participantes utilizaron sus métodos de análisis de rutina para los analitos y matrices incluidos en el ejercicio. Cuando los participantes declararon métodos diferentes, los resultados obtenidos de la evaluación estadística se compararon respecto del valor asignado por consenso y del desempeño obtenidos de acuerdo con el método.

- ✓ Número de matrices: 2, agua residual y agua superficial
- ✓ Número de muestras: 3
 - Agua de medio receptor: río
 - Aguas residuales procedentes de EDAR: influente (no tratada) y efluente (tratada).





- ✓ Parámetros para analizar: sustancias de preocupación emergente contenidas en la Directiva (UE) 2024/3019 sobre el tratamiento de las aguas residuales urbanas:

Categoría 1: Sustancias que pueden tratarse con mucha facilidad	Nº CAS
Amisulpride	71675-85-9
Carbamazepine	298-46-4
Citalopram	59729-33-8
Clarithromycin	81103-11-9
Diclofenac	15307-86-5
Hydrochlorothiazide	58-93-5
Metoprolol	37350-58-6
Venlafaxine	93413-69-5

Categoría 2: Sustancias que pueden eliminarse con facilidad	Nº CAS
Benzotriazole	95-14-7
Candesartan	139481-59-7
Irbesartan	138402-11-6
Mezcla de 4-Methylbenzotriazole y 5-Methylbenzotriazole: 4(5)Me-BT	29878-31-7 / 136-85-6

- ✓ Método analítico propuesto: cromatografía líquida de alto rendimiento (UHPLC) acoplada a espectrometría de masas en tándem (MS/MS).

5.4. Envío de resultados de los participantes

Los participantes fueron informados previamente de los plazos establecidos para el envío de resultados, tanto por correo electrónico como a través del sitio web [BlueWWater Interlab](#)

El cumplimiento de dichos plazos fue responsabilidad de los laboratorios participantes. Para facilitar su seguimiento, se establecieron las siguientes medidas informativas:

- ✓ *Aviso de fin de plazo 5 días antes del límite*
- ✓ *Fecha límite de envío de resultados*

5.4.a. Colusión y falsificación de resultados

No se aceptaron situaciones en las que existiesen sospechas de colusión, connivencia o falsificación de resultados, ya fuese entre participantes o por parte de alguno de ellos de





forma individual. Corresponde a los laboratorios participantes evitar cualquier práctica que pueda comprometer la calidad del ensayo de aptitud y el cumplimiento de las buenas prácticas profesionales.

5.5. Emisión de informes

Los resultados obtenidos se presentan a los participantes en el presente Informe Final del ensayo de aptitud, elaborado conforme a lo establecido en el Protocolo General.

La reproducción total o parcial del informe requerirá autorización previa por escrito.

6. ESTUDIO DE HOMOGENEIDAD Y ESTABILIDAD

6.1. Ensayos de homogeneidad

La homogeneidad de los ítems de ensayo se garantizó mediante la aplicación de los criterios estadísticos establecidos en la norma ISO 13528:2022, el Protocolo Armonizado de la IUPAC y otros documentos de referencia. La evaluación de la homogeneidad se realizó una vez envasados los ítems, en las mismas condiciones en las que fueron suministrados a los participantes y antes de su distribución.

Los materiales no fueron enviados hasta comprobar que las submuestras individuales presentaban un grado de homogeneidad adecuado. Asimismo, la evaluación de la homogeneidad y estabilidad requirió, siempre que fue posible, un muestreo previo destinado a asegurar la fiabilidad de los resultados.

Los materiales de ensayo debían cumplir los requisitos de homogeneidad y estabilidad definidos en el protocolo *Evaluación de la homogeneidad y estabilidad de muestras del intercomparativo*.

La verificación de la homogeneidad consistió en seleccionar aleatoriamente al menos 10 porciones del material de ensayo, analizar por duplicado cada una de ellas en condiciones de repetibilidad y evaluar los resultados mediante un método cuya desviación estándar en repetibilidad (S_{muestra}) no superase el 30 % de la desviación estándar objetivo (σ_{PT}), correspondiente a la reproducibilidad esperada del ensayo de aptitud.

Por tanto, el material se consideró homogéneo cuando se cumplió el siguiente criterio:

$$S_{\text{muestra}} \leq 0,3 * \sigma_{\text{PT}} (\text{desv. std. objetivo})$$

Cuando la homogeneidad resultó satisfactoria, el análisis de varianza de un factor mostró que la varianza entre muestras no era significativamente superior a la varianza intramuestras.

Dado que se trataba del primer estudio interlaboratorio realizado para estos analitos, la desviación estándar objetivo (σ_{PT}) era desconocida. Por ello, se adoptó como criterio el 25 % del valor asignado correspondiente, aproximación habitualmente aceptada en ensayos de aptitud de pesticidas en rangos de concentración del orden de ppt.





Para verificar la homogeneidad del material preparado, se enviaron 10 muestras al laboratorio de referencia, donde fueron analizadas por duplicado en condiciones de repetibilidad. Los resultados obtenidos fueron sometidos a evaluación estadística conforme a los criterios establecidos en la norma ISO 13528 y en el Protocolo Armonizado de la IUPAC.

Los resultados de homogeneidad fueron satisfactorios y se presentan de forma resumida en la Tabla 2, correspondiente a los analitos evaluados en las tres matrices estudiadas.

TABLA 2: Resumen de los resultados de homogeneidad, todas las matrices.

Compuesto	Agua superficial				Agua residual efluente				Agua residual influente			
	15%	15% + margen	25%	25% + margen	15%	15% + margen	25%	25% + margen	15%	15% + margen	25%	25% + margen
4(5)Me-BT	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
amisulpride	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
benzotriazole	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
candesartan	—	—	—	—	X	✓	✓	✓	✓	✓	✓	✓
carbamazepine	✓	✓	✓	✓	X	✓	✓	✓	✓	✓	✓	✓
citalopram	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
clarithromycin	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
diclofenac	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
hydrochlorothiazide	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
irbesartan	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
metoprolol	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
venlafaxine	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

✓ Cumple criterio de homogeneidad X No cumple criterio — Sin dato disponible

6.2. Ensayos de estabilidad

Los ítems de ensayo deben mantener una estabilidad adecuada durante todo el desarrollo del ensayo de aptitud, incluyendo el periodo comprendido entre su preparación y el inicio





del ejercicio, el transporte de las muestras y el tiempo establecido para que los participantes realicen los análisis.

La evaluación de la estabilidad se realizó conforme a los criterios definidos en el protocolo *Evaluación de la homogeneidad y estabilidad de muestras del intercomparativo*.

Para las comprobaciones de estabilidad se consideraron los siguientes aspectos:

- ✓ Todas las propiedades incluidas en la ronda de ensayo debían analizarse o, alternativamente, verificarse mediante experiencia previa y justificación técnica basada en el conocimiento de la matriz y del método de medida.
- ✓ Se analizaron al menos tres ítems de ensayo por duplicado. En aquellos casos en los que la repetibilidad resultó insuficiente, se incrementó el número de muestras o réplicas.
- ✓ Las determinaciones de estabilidad se realizaron con el mismo número de repeticiones empleado en el estudio previo de homogeneidad.
- ✓ El criterio considerado para identificar problemas de repetibilidad fue:

$$Sr \text{ o } Sw > 0,5 \sigma_{PT}$$

La estabilidad se evaluó mediante el análisis de una muestra representativa de ítems al final de la ronda de ensayo y su comparación estadística con los resultados obtenidos antes del inicio del ejercicio, con objeto de verificar la ausencia de cambios significativos durante el periodo de estudio.

Para ello, se analizaron por duplicado tres muestras antes del inicio del ejercicio y otras tres una vez finalizada la ronda y recibidos los resultados de todos los participantes. Posteriormente, se realizó el estudio estadístico correspondiente de acuerdo con los criterios establecidos en la norma ISO 13528 y en el Protocolo Armonizado de la IUPAC.

Se compara el promedio general de las mediciones obtenidas en el control de homogeneidad \bar{y}_1 con el promedio general de los resultados obtenidos en el control de estabilidad \bar{y}_2 . Las muestras pueden considerarse adecuadamente estables si:

$$|\bar{y}_1 - \bar{y}_2| \leq 0,3 \cdot \sigma_{PT}$$

Es probable que la precisión intermedia del método de medición (o la incertidumbre de medición del ítem) pueda contribuir a la imposibilidad de cumplir dicho criterio. En este caso el criterio puede expandirse agregando la incertidumbre de la diferencia a σ_{PT} objetivo y luego verificar la diferencia según lo siguiente:

$$|\bar{y}_1 - \bar{y}_2| \leq 0,3 \cdot \sqrt{\sigma_{PT}^2 + 2 \cdot u(x_i)^2}$$





En caso de no cumplirse los criterios de estabilidad, podían adoptarse distintas medidas, tales como:

- ✓ cuantificar el efecto de la inestabilidad e incorporarlo a la evaluación del desempeño (por ejemplo, mediante z'-score);
- ✓ revisar los procedimientos de preparación y almacenamiento de las muestras;
- ✓ limitar la evaluación del desempeño a fines exclusivamente informativos.

Los resultados de estabilidad fueron satisfactorios en la mayoría de los casos, aunque algunos analitos no presentaron estabilidad adecuada. Los resultados resumidos se muestran en las Tabla 3, correspondientes a las tres matrices estudiadas.

TABLA 3: Resumen de los resultados de estabilidad, todas las matrices.

Compuesto	Agua superficial				Agua residual efluente				Agua residual influente			
	15%	15% + margen	25%	25% + margen	15%	15% + margen	25%	25% + margen	15%	15% + margen	25%	25% + margen
4(5)Me-BT	X	X	X	X	X	X	X	X	X	X	X	✓
amisulpride	✓	✓	✓	✓	X	✓	X	✓	X	X	X	X
benzotriazole	X	X	X	X	✓	✓	✓	✓	X	✓	✓	✓
candesartan	—	—	—	—	✓	✓	✓	✓	X	✓	✓	✓
carbamazepine	X	✓	X	✓	X	X	X	X	X	X	X	X
citalopram	X	✓	✓	✓	X	X	X	X	X	✓	✓	✓
clarithromycin	X	✓	✓	✓	✓	✓	✓	✓	X	X	X	✓
diclofenac	X	X	X	X	X	✓	✓	✓	✓	✓	✓	✓
hydrochlorothiazide	✓	✓	✓	✓	X	✓	X	✓	X	X	X	✓
irbesartan	X	✓	X	✓	X	✓	✓	✓	X	✓	X	✓
metoprolol	X	X	X	X	X	X	X	X	X	✓	X	✓
venlafaxine	✓	✓	✓	✓	X	X	X	X	X	X	X	X

✓ Cumple criterio de homogeneidad X No cumple criterio — Sin dato disponible

Los resultados del estudio de estabilidad muestran un comportamiento heterogéneo en función del analito y de la matriz. Algunos compuestos presentaron una estabilidad satisfactoria en todas las matrices y criterios evaluados (amisulpride, hydrochlorothiazide y venlafaxine en agua superficial; benzotriazole, candesartan y clarithromycin en agua





residual efluente; diclofenac en agua residual influente). Sin embargo, un número significativo de analitos no superó los criterios de estabilidad en alguna de las matrices, incluso aplicando el margen adicional o el criterio del 25 %:

- ✓ Agua superficial: 4(5)Me-BT, benzotriazole, diclofenac y metoprolol no cumplieron en ningún criterio.
- ✓ Agua residual efluente: 4(5)Me-BT, carbamazepine, citalopram, metoprolol y venlafaxine no cumplieron en ningún criterio.
- ✓ Agua residual influente: amisulpride, carbamazepine y venlafaxine no cumplieron en ningún criterio; 4(5)Me-BT únicamente cumplió con el criterio más amplio (25 % con margen adicional).

Este comportamiento selectivo analito-matriz es compatible con fenómenos de adsorción, degradación diferencial o efectos de la congelación/descongelación, más que con una alteración global de las muestras durante su conservación. En consecuencia, la estabilidad previa a la distribución fue adecuada para un subconjunto de analitos, pero debe considerarse comprometida para los compuestos indicados, lo que deberá tenerse en cuenta en la interpretación de los resultados del ejercicio intercomparativo.

Se ha tenido en consideración cualquier efecto conocido del transporte al evaluar el desempeño. Cualquier aumento en la incertidumbre debido al transporte se incluirá en la incertidumbre del valor asignado. Como el valor asignado y la desviación estándar para la evaluación de competencia se determinan a partir de los resultados de los participantes, estos estadísticos podrían reflejar cualquier sesgo y efectos en la incertidumbre causados por las condiciones de transporte.

TABLA 4: Clasificación de compuestos según su estabilidad

Matriz agua de medio receptor	Matriz agua residual influente	Matriz agua residual efluente
<p>Alta estabilidad</p> <ul style="list-style-type: none"> • amisulpride • hydrochlorothiazide • venlafaxine 	<p>Alta estabilidad</p> <ul style="list-style-type: none"> • diclofenac • benzotriazole • candesartan 	<p>Alta estabilidad</p> <ul style="list-style-type: none"> • diclofenac • benzotriazole • candesartan • clarithromycin • irbesartan
<p>Estabilidad media</p> <ul style="list-style-type: none"> • citalopram • clarithromycin • irbesartan • carbamazepine 	<p>Estabilidad media</p> <ul style="list-style-type: none"> • citalopram • irbesartan • metoprolol • hydrochlorothiazide 	<p>Estabilidad media</p> <ul style="list-style-type: none"> • amisulpride • hydrochlorothiazide
<p>Baja estabilidad</p> <ul style="list-style-type: none"> • 4(5)-MeBT • benzotriazole • diclofenac • metoprolol 	<p>Baja estabilidad</p> <ul style="list-style-type: none"> • amisulpride • carbamazepine • venlafaxin 	<p>Baja estabilidad</p> <ul style="list-style-type: none"> • 4(5)-MeBT • citalopram • carbamazepine • metoprolol • venlafaxin





7. TRATAMIENTO ESTADÍSTICO DE LOS RESULTADOS

7.1. Importantes consideraciones previas

La norma ISO 13528 y el Protocolo Armonizado de la IUPAC se aplicaron estrictamente en el tratamiento estadístico de los datos aportados por los participantes, especialmente en los siguientes aspectos:

- ✓ *Las sistemáticas posibles de asignación de valores al valor central incluyendo la utilización de estadística robusta, y al valor de variabilidad, fundamentándose generalmente en la utilización de una σ objetivo.*
- ✓ *La conveniencia de la comparación de la σ realmente obtenida con respecto a la σ objetivo considerando que no debe superar ésta en un factor crítico de 1,2.*
- ✓ *Una característica fundamental de la norma es que utiliza todos los valores obtenidos por los participantes, sin descartar ninguno, pero modificando aquellos que considera atípicos.*

Dado que la σ_{PT} objetivo era desconocida, se adoptó como tal el 25% del valor asignado respectivo, que es un criterio habitualmente aceptado en los ensayos de aptitud de pesticidas en niveles de concentración del orden de ppt.

No obstante, para la valoración del desempeño se utilizó la σ_{PT} robusta calculada según la norma ISO 13528, aun sabiendo de los altos valores de variabilidad detectados en el interlaboratorio y de la presencia de escenarios de bimodalidad.

La razón fundamental fue determinar una precisión de reproducibilidad lo más realista posible, para el binomio analito/matriz, de cara a establecer esa figura de mérito de precisión y unos límites máximos de la misma en términos de S_R de reproducibilidad o RSD%, en cuanto a la validación del método en el futuro por los laboratorios y no tanto la evaluación del desempeño de los laboratorios participantes.

Adicionalmente, se ha realizado el tratamiento estadístico previo de los datos siguiendo las recomendaciones indicadas en la norma ISO 5725-2, basada en estadística clásica, de manera que se pudieron detectar valores discrepantes tanto en sesgo como en repetibilidad, y así ayudase a una mejor interpretación de los resultados de precisión perseguidos y del desempeño de los participantes.

7.2. Tratamiento de los datos inconsistentes o aberrantes

Los datos aportados por los laboratorios participantes, a menudo y según corroboran las pruebas experimentales, aparecen distribuidos según modelos similares a distribuciones normales o gaussianas, a las que, con frecuencia, se añaden observaciones de carácter “anómalo” respecto del conjunto.

El tratamiento previo de estos datos supone minimizar o eliminar el peso estadístico de estas observaciones anómalas, con un objetivo: establecer parámetros adecuados para evaluar los resultados de los laboratorios participantes.

En primer lugar, se han de descartar aquellos valores considerados aberrantes utilizando las indicaciones del Protocolo Armonizado de la IUPAC, según el cual se consideran como



tales los resultados que se diferencian en más de un $\pm 50\%$ de la mediana de todos los resultados informados por los participantes.

Se procede, si es el caso, de la siguiente manera:

1º *Test de Cochran* (en el caso de realizar 2 o más medidas por parámetro)

2º *Test de Grubbs*

3º *Test estadísticos robustos basados en la mediana*

7.2.a. *Test de Grubbs.*

El tratamiento estadístico y análisis de los datos inconsistentes se realiza según el método test de Grubbs sólo cuando hay un mínimo de 7 resultados en el ejercicio, con el fin de establecer los valores anómalos.

El Test de Grubbs simple y doble se utiliza para la eliminación de resultados con un sesgo significativo.

Se realiza la prueba de Grubbs simple para determinar si el resultado más alejado por presentar un sesgo significativo es incompatible con el resto.

Si el valor más alejado es el mayor del conjunto de resultados, el estadístico de Grubbs se calcula mediante la siguiente fórmula y se compara con el valor crítico de tablas:

$$G_h = \left(x_h - \bar{x} \right) / s$$

En el caso de que se trate del valor más pequeño, se calcula el estadístico:

$$G_l = \left(\bar{x} - x_l \right) / s$$

Y se procede a comparar dicho valor con el valor crítico de tablas que corresponda, en nuestro caso, para un 95 % de nivel de significación.

Se debe verificar previamente la hipótesis de normalidad en la distribución de datos.

En cualquier caso, siempre debe prevalecer el juicio crítico del organizador de la intercomparación para decidir qué valores son realmente discrepantes.

7.2.b. *Test estadísticos robustos basados en la mediana.*

Si se presenta una cantidad considerable de valores anómalos (del orden del 10% de los valores de la serie), la prueba de detección de datos discrepantes puede arrojar conclusiones estadísticamente erróneas.

En este caso, existen criterios más útiles que se fundamentan en *estadísticas robustas basadas en propiedades de la mediana*, que no se ven tan afectados por el tipo de población existente.

La Norma ISO 13528 y el Protocolo de la IUPAC establecen, en lo que se refiere al tratamiento de los datos aportados por los participantes lo siguiente:



- ✓ *Las sistemáticas posibles de asignación de valores al valor central incluyendo la utilización de estadística robusta, y al valor de variabilidad, fundamentándose generalmente en la utilización de una σ objetivo.*
- ✓ *La conveniencia de la comparación de la σ realmente obtenida con respecto a la σ objetivo considerando que no debe superar ésta en un factor crítico de 1,2.*
- ✓ *Una característica fundamental de la norma es que utiliza todos los valores obtenidos por los participantes, sin descartar ninguno, pero modificando aquellos que considera atípicos.*

La sistemática se desarrolla en los siguientes pasos:

- ✓ *Se listan todos los datos válidos.*
- ✓ *Ordenar los datos de menor a mayor.*
- ✓ *Calcular la mediana de los datos.*
- ✓ *Determinar la Desviación Absoluta de la mediana (MAD): Es la mediana del valor absoluto de todas las desviaciones de cada resultado de ensayo respecto a la mediana. Es decir:*

$$MAD = \text{mediana} (| xi - \text{mediana} |)$$

donde xi es el resultado de ensayo de cada participante

Y se procede a comparar si existen valores anómalos/discrepantes mediante el criterio siguiente:

Mediana \pm 1,5 SMAD

Para que la MAD resulte equivalente a la desviación estándar de una distribución normal, se multiplica por 1,483. Al valor resultante se le denomina SMAD.

Una vez determinados los datos anómalos, se procede a realizar el análisis estadístico del índice de desempeño z-score, *sin que sean excluidos los valores anómalos detectados.*

7.3. Normalidad estadística de los datos

Una vez recibidos todos los resultados, se evalúa la unimodalidad mediante el *test de Kernel* y otros test estadísticos adecuados, explicándose el procedimiento a seguir en el caso de que se aprecie más de una distribución o moda estadística.

En general las distribuciones normales de datos ofrecen valores de mediana y moda iguales o muy próximos. A medida que la distribución se aleja de la normalidad estadística estos valores difieren significativamente.

7.3.a. Test de Kernel

Para verificar la posible multimodalidad del ensayo, se calcula una estimación de Kernel utilizando como valor $h = 0,75 \sigma_{PT}$ (75% de la desviación estándar robusta). En caso de



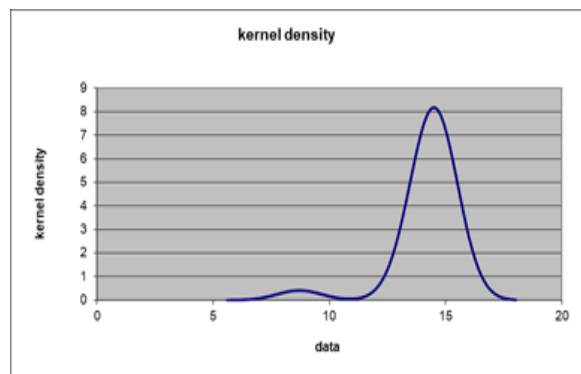


aparecer una distribución multimodal se valorarán las posibles causas de esta, y de no poderse obtener una explicación razonable solo se indicarán valores asignados y desviación típica a título informativo.

Según la recomendación del Protocolo de la IUPAC, *no se determina un valor de consenso y no se informa ningún índice de desempeño de laboratorio individual para la ronda*. Sin embargo, aún puede ser útil proporcionar a las participantes estadísticas resumidas sobre el conjunto de datos.

En el test de densidad de Kernel, cada dato es reemplazado por una distribución normal con centro en dicho punto y la desviación estándar de h . Las distribuciones normales son sumadas para obtener una distribución final conocida como la Densidad de Kernel.

La representación gráfica permite visualizar numéricamente el valor de consenso más representativo en el centro de esta distribución normal. Ejemplos gráficos de Kernel:



7.4. Establecimiento del valor asignado

Para cada una de las determinaciones analíticas realizadas en el ensayo de aptitud, el correspondiente valor asignado (valor de referencia x_{PT}) se establece por consenso, a partir de la media robusta del conjunto de resultados iniciales, de acuerdo con los criterios definidos en la norma ISO 13528:2022.

Se determina la media robusta $\hat{\mu}_{rob}$ y la desviación estándar $\hat{\sigma}_{rob}$ de los n resultados. Si $\hat{\sigma}_{rob}$ es menor que aproximadamente $1.2 \sigma_{PT}$, *entonces se usa $\hat{\mu}_{rob}$ como el valor asignado x_{PT}*

Previamente se realiza una estimación de densidad de kernel de la distribución de los resultados utilizando un ancho de banda $h = 0,75 \sigma_{PT}$. Si esto resulta en una densidad de kernel unimodal y aproximadamente simétrica, y la moda y la mediana son casi coincidentes, *entonces se confirma el uso de $\hat{\mu}_{rob}$ como el valor asignado x_{PT}* .

Deben cumplirse ambas condiciones en el caso de valor asignado por consenso.

Los valores asignados y sus incertidumbres se han determinado por consenso de acuerdo con lo establecido en la norma ISO 13528:2022.



7.5. Incertidumbre estándar del valor asignado

La incertidumbre del valor asignado, por consenso, se determina mediante estadística robusta, siendo la fórmula de cálculo la siguiente:

$$\mu (x_{PT}) = 1,25 \times s^* / \sqrt{n}$$

donde n es el número de participantes en la ronda. (Norma ISO 13528:2015, punto 7.7.3)

Se considera apropiado comparar $\mu (x_{PT})$ con $\sigma_{objetivo}$ para verificar que la incertidumbre no tenga un efecto adverso en las puntuaciones z -score.

Para lograr la mejor decisión para la evaluación del desempeño de los laboratorios se tiene en cuenta algunas pautas. El elemento clave sería la cuantía de la incertidumbre del valor asignado $\mu (x_{PT})$ en comparación con el de la desviación estándar para la evaluación de competencia σ_{PT} (objetivo).

El método para la determinación de estos dos parámetros también es crucial. Puede ser conveniente reconsiderar (no aplicar) el uso de cualquier puntuación en el caso de una incertidumbre $\mu (x_{PT})$ demasiado "grande". Esto evitaría cualquier falsa impresión de buen desempeño por parte de los laboratorios.

Según lo anterior, se pueden producir tres situaciones:

1. Si $\mu (x_{PT}) < 0.3 * \sigma_{objetivo} \Rightarrow$ la incertidumbre del valor asignado es despreciable y no necesita ser incluida en la interpretación de los resultados del ejercicio de intercomparación
2. Si $0.3 * \sigma_{objetivo} < \mu (x_{PT}) < 0.7 * \sigma_{objetivo} \Rightarrow$ los índices de evaluación deben considerar la influencia de la incertidumbre valor asignado en la calidad ejercicio (z' -score)
3. Si $\mu (x_{PT}) > 0.7 * \sigma_{objetivo} \Rightarrow$ la calidad del ejercicio de intercomparación es limitada y su evaluación queda a título meramente informativo

En situaciones donde $\mu (x_{PT}) > 0.3 * \sigma_{objetivo}$, la norma ISO 13528 estipula que se debe:

- ✓ *buscar un método de determinación del valor asignado x_{PT} en el que su incertidumbre cumpla con el criterio anterior.*
- ✓ *aplicar índices de evaluación que incluyan $\mu (x_{PT})$ en su expresión*
- ✓ *informar a los participantes que $\mu (x_{PT})$ no es despreciable*
- ✓ *investigar si la falta de consenso revela bimodalidad (test de kernel)*
- ✓ *considerar que la elevada $\mu (x_{PT})$ es inherente a la distribución de los resultados de los participantes*





La norma ISO 13528:2022 describe el uso de puntuaciones alternativas a z-score cuando se presentan estos escenarios como el índice z' -score.

En este ejercicio no se ha aplicado el uso de z' -score. El valor de σ_{objetivo} se define en sí mismo por la adecuación al uso previsto. Por lo tanto, el uso de z' -score se desvía de esa adecuación al uso previsto predefinida y esconde con eficacia la alta incertidumbre del valor asignado. Por tanto, se prefiere emitir z-scores “solo a título informativo”, cuando la incertidumbre es mayor que la ideal o simplemente no publicar las puntuaciones z en absoluto cuando la incertidumbre es inaceptable.

7.6. Desviación estándar del ensayo de aptitud

La desviación estándar para la aptitud (informalmente desviación estándar “objetivo”, σ_{PT}) determina los límites para el rendimiento satisfactorio en un ensayo de aptitud. Se establece como un valor que refleja la aptitud al fin previsto del análisis en cuestión.

Las desviaciones estándar del rendimiento adecuadas para el fin previsto pueden obtenerse a partir de:

- ✓ *Definición legal*
- ✓ *Reglamentos aplicables*
- ✓ *Estudios colaborativos / Estudios de rendimiento de métodos*
- ✓ *Modelos predictivos, por ejemplo, la ecuación Horwitz o sus modificaciones*
- ✓ *Opinión de expertos*
- ✓ *Otras opciones: Histórico de rondas anteriores*

Para cada una de las determinaciones analíticas realizadas en el ensayo de aptitud, la correspondiente desviación estándar $\hat{\sigma}_{rob}(s^*)$ se ha establecido por consenso, a partir de la desviación estándar robusta del conjunto de resultados iniciales, de acuerdo con los criterios definidos en la norma ISO 13528:2022.

Se determina por tanto la desviación estándar $\hat{\sigma}_{rob}$ de los n resultados. Si $\hat{\sigma}_{rob}$ es menor que aproximadamente $1.2 \sigma_{PT}$, entonces se toma $\hat{\sigma}_{rob}(s^*)$ como el valor de desviación estándar para el ensayo de aptitud.

La desviación estándar para el ensayo de aptitud se ha determinado de acuerdo con la norma ISO 13528:2022.

Al ser el primer estudio interlaboratorio sobre estos analitos, la σ_{PT} objetivo es desconocida, por lo que se adoptó como tal el 25% del valor asignado respectivo, que es un criterio habitualmente aceptado en los ensayos de aptitud de pesticidas en entornos de concentración de ppt.





8. RESUMEN ESTADISTICO DE LOS RESULTADOS

Se detalla a continuación un resumen estadístico de los resultados obtenidos en cada una de las matrices estudiadas y los analitos definidos previamente.

8.1. Resumen de estadísticas

El resumen de estadísticas se detalla en las tablas 5, 6 y 7, correspondientes a los analitos evaluados en las 3 matrices estudiadas.

En dichas tablas, se muestran, entre otros, los valores de la precisión obtenida en el ensayo de intercomparación, en términos de reproducibilidad (SR_{PT}), junto con el valor asignado para cada analito ($VA_{x_{PT}}$), objeto de esta intercomparación. Asimismo, se indica el número de participantes (valores n en cada parámetro estudiado).

TABLA 5. Resumen de resultados en matriz agua superficial (medio receptor)

Analito	$VA(x_{PT})$	Unidades	SR_{PT}	$\sigma_{objetivo}$	CV-RSD %	n	Mediana	MAD	$u_{VA}(x_{PT})$
Amisulpride	66,99	ng/l	25,22	16,75	37,64	10	64,27	17,32	9,97
Benzotriazole	104,11	ng/l	53,77	26,03	51,65	10	110,39	37,12	21,26
Candesartan	14,94	ng/l	12,54	3,73	83,94	6	11,96	7,29	6,40
Carbamazepine	23,55	ng/l	6,69	5,89	28,41	10	22,96	6,01	2,65
Citalopram	94,19	ng/l	53,31	23,55	56,60	10	93,31	37,01	21,07
Clarithromycin	72,10	ng/l	33,44	18,02	46,39	11	70,08	23,33	12,61
Diclofenac	217,76	ng/l	42,48	54,44	19,51	10	223,74	29,79	16,79
Hydrochlorothiazide	247,98	ng/l	81,15	61,99	32,72	8	224,09	47,55	35,86
Irbesartan	95,81	ng/l	25,77	23,95	26,90	11	90,18	16,44	9,71
Metoprolol	65,25	ng/l	21,94	16,31	33,62	11	64,42	14,91	8,27
Venlafaxine	88,52	ng/l	39,00	22,13	44,05	11	91,00	24,24	14,70
4(5)-MeBT	105,79	ng/l	45,46	26,45	42,97	9	92,45	30,96	18,94

TABLA 6. Resumen de resultados en matriz agua residual efluente

Analito	$VA(x_{PT})$	Unidades	SR_{PT}	$\sigma_{objetivo}$	CV-RSD %	n	Mediana	MAD	$u_{VA}(x_{PT})$
Amisulpride	187,25	ng/l	52,38	46,81	27,97	10	185,64	28,65	20,70
Benzotriazole	216,77	ng/l	134,42	54,19	62,01	10	202,69	81,62	53,13
Candesartan	209,02	ng/l	90,41	52,25	43,26	8	231,09	55,70	39,96
Carbamazepine	29,30	ng/l	11,52	7,32	39,33	11	26,05	7,95	4,34
Citalopram	59,51	ng/l	43,27	14,88	72,71	10	57,75	37,82	17,10
Clarithromycin	43,84	ng/l	15,31	10,96	34,92	10	38,05	9,54	6,05
Diclofenac	553,21	ng/l	258,74	138,30	46,77	12	584,65	149,47	93,36
Hydrochlorothiazide	1302,96	ng/l	437,61	325,74	33,59	8	1191,23	286,29	193,40
Irbesartan	478,76	ng/l	206,85	119,69	43,21	11	482,96	125,01	77,96





Analito	VA(x _{pt})	Unidades	SR _{PT}	σ _{objetivo}	CV-RSD %	n	Mediana	MAD	u _{VA} (x _{pt})
Metoprolol	19,78	ng/l	6,74	4,94	34,06	7	17,71	4,06	2,98
Venlafaxine	265,73	ng/l	96,72	66,43	36,40	11	255,90	54,64	36,45
4(5)-MeBT	469,43	ng/l	254,81	117,36	54,28	10	461,55	152,53	100,72

TABLA 7. Resumen de resultados en matriz agua residual influente

Analito	VA(x _{pt})	Unidades	SR _{PT}	σ _{objetivo}	CV-RSD %	n	Mediana	MAD	u _{VA} (x _{pt})
Amisulpride	124,26	ng/l	42,43	31,06	34,15	10	131,72	26,32	16,77
Benzotriazole	176,74	ng/l	123,71	44,18	69,99	10	162,15	92,21	48,90
Candesartan	181,25	ng/l	83,63	45,31	46,14	8	200,99	52,50	39,96
Carbamazepine	30,84	ng/l	13,89	7,71	45,04	11	36,49	12,13	5,24
Citalopram	78,67	ng/l	45,56	19,67	57,91	10	94,50	31,22	18,01
Clarithromycin	70,87	ng/l	41,83	17,72	59,02	10	64,96	25,56	16,53
Diclofenac	255,26	ng/l	178,79	63,82	70,04	12	295,53	133,49	64,52
Hydrochlorothiazide	969,65	ng/l	227,67	242,41	23,48	8	938,42	135,03	100,62
Irbesartan	435,37	ng/l	150,32	108,84	34,53	11	480,17	121,83	56,65
Metoprolol	12,29	ng/l	9,60	3,07	78,15	7	10,26	5,74	4,54
Venlafaxine	330,50	ng/l	129,03	82,63	39,04	11	343,05	81,24	48,63
4(5)-MeBT	330,60	ng/l	169,59	82,65	51,30	10	338,25	134,43	63,92

Donde:

- VA(x_{pt}): Valor asignado (consenso)
- SR_{PT}: Desviación estándar de reproducibilidad (consenso)
- σ_{objetivo}: Desviación estándar (Sigma objetivo) (25% de VA)
- CV-RSD %: Coeficiente de variación – Desviación estándar relativa (%)
- n: Número de valores (participantes)
- Mediana: Mediana de los valores iniciales reportados por los laboratorios
- MAD: Desviación absoluta de la mediana
- u_{VA}(x_{pt}): Incertidumbre estándar del valor asignado

8.2. Clasificación general de puntuación z-score de los laboratorios

Se estableció una clasificación general, para cada laboratorio participante, calculándose previamente sus puntuaciones z ordenadas.

Dicha clasificación se basó en dos criterios: el % Real de $|z| < 2$ y el % Posible de $|z| < 2$

El porcentaje Real de $|z| < 2$ se calcula a partir del número total de $|z| < 2$ obtenidos por un laboratorio, expresado como porcentaje en función del número de valores numéricos enviados por el propio laboratorio.



El porcentaje Posible de $|z| < 2$ se calcula a partir del número total de $|z| < 2$ obtenidos por un laboratorio, expresado en porcentaje, sobre el número total de valores asignados obtenidos en el estudio.

Se detallan estos porcentajes obtenidos por cada laboratorio en cada matriz estudiada, en las siguientes tablas 8, 9 y 10.

TABLA 8. Resumen de clasificación en matriz agua superficial (medio receptor)

Laboratorio	n $ z < 2$	n enviados	% Real $ z < 2$	Laboratorio	% Posible $ z < 2$
Lab.02	11	11	100,0	Lab.02	91,7
Lab.04	10	10	100,0	Lab.03	91,7
Lab.08	10	10	100,0	Lab.05	91,7
Lab.09	10	10	100,0	Lab.04	83,3
Lab.11	10	10	100,0	Lab.08	83,3
Lab.12	10	10	100,0	Lab.09	83,3
Lab.01	1	1	100,0	Lab.11	83,3
Lab.03	11	12	91,7	Lab.12	83,3
Lab.05	11	12	91,7	Lab.10	83,3
Lab.10	10	12	83,3	Lab.06	66,7
Lab.06	8	10	80,0	Lab.07	16,7
Lab.07	2	4	50,0	Lab.01	8,3

TABLA 9. Resumen de clasificación en matriz agua residual efluente

Laboratorio	n $ z < 2$	n enviados	% Real $ z < 2$	Laboratorio	% Posible $ z < 2$
Lab.09	11	11	100	Lab.09	91,7
Lab.08	10	10	100	Lab.02	91,7
Lab.12	10	10	100	Lab.05	91,7
Lab.01	1	1	100	Lab.10	91,7
Lab.02	11	12	91,7	Lab.08	83,3
Lab.05	11	12	91,7	Lab.12	83,3
Lab.10	11	12	91,7	Lab.06	83,3
Lab.06	10	11	90,9	Lab.04	75,0
Lab.04	9	10	90,0	Lab.11	66,7
Lab.11	8	9	88,9	Lab.03	58,3
Lab.07	4	6	66,7	Lab.07	33,3
Lab.03	7	12	58,3	Lab.01	8,3





TABLA 10. Resumen de clasificación en matriz agua residual influente

Laboratorio	n $ z < 2$	n enviados	% Real $ z < 2$	Laboratorio	% Posible $ z < 2$
Lab.02	12	12	100,0	Lab.02	100
Lab.09	10	10	100,0	Lab.10	91,7
Lab.12	9	9	100,0	Lab.09	83,3
Lab.07	5	5	100,0	Lab.06	83,3
Lab.01	1	1	100,0	Lab.03	83,3
Lab.10	11	12	91,7	Lab.05	83,3
Lab.06	10	11	90,9	Lab.12	75,0
Lab.08	9	10	90,0	Lab.08	75,0
Lab.11	8	9	88,9	Lab.11	66,7
Lab.03	10	12	83,3	Lab.04	66,7
Lab.05	10	12	83,3	Lab.07	41,7
Lab.04	8	10	80,0	Lab.01	8,3

9. EVALUACIÓN DEL DESEMPEÑO

Para la evaluación del desempeño se han tenido en cuenta diferentes criterios como los resultados obtenidos frente a los esperados, variación intra e interlaboratorios, variación entre métodos, posibles errores (valores atípicos), causas (normalmente inusuales) que imposibilitan la evaluación de los resultados, asesoramiento a los participantes, otras sugerencias o recomendaciones y conclusiones.

No se consideran en la evaluación los datos recibidos e informados como “menor que” o por debajo del límite de cuantificación o detección reportado por el laboratorio participante.

Si la evaluación estadística y de desempeño han presentado características inusuales, tales como: ciertas tendencias hacia sesgo positivo o negativo, distribuciones multimodales, etc. se han evaluado de forma especial.

Siguiendo el Protocolo General, en este Informe Final se presenta solamente una de las diferentes opciones de indicadores o puntuaciones de desempeño sin excluir que se puedan adoptar otras alternativas que ofrece las guías y normas aplicadas.

El análisis estadístico *z-score* es utilizado para este ensayo de aptitud cuantitativo. Una puntuación-*z* combina una estimación del error de un resultado con una desviación estándar.

Los resultados de los diferentes laboratorios para cada uno de los parámetros analizados se evalúan la puntuación *z-score*, que es un índice de desempeño el cual compara las diferencias entre los resultados de los participantes y el valor asignado en términos de la dispersión aceptable de los resultados o la desviación estándar establecida, calculada a partir de la siguiente ecuación:



$$Z = \frac{X - X_a}{\sigma_{pt}}$$

Donde:

Xa = resultado de los participantes

X = valor asignado (VA)

σ_{pt} = Desviación estándar establecida para el ensayo de aptitud.

Se pueden presentar diferencias al comparar cálculos de z-score realizados por los participantes con los obtenidos con la planilla de cálculo Excel debido a efectos de redondeo. Los informes de INTECMAR detallan los resultados z-score con 2 decimales.

Se asumirá por tanto una distribución normal, para cualquier grupo de datos con distribución normal los Z-score deberán estar entre el rango de ± 2 a ± 3 .

Los criterios de aceptabilidad están definidos por el valor obtenido por cada laboratorio, que son clasificados de la siguiente manera, en cuanto a su interpretación:

$$\begin{aligned} |Z| < 2 & \text{ satisfactorio} \\ 2 < |Z| < 3 & \text{ cuestionable} \\ |Z| > 3 & \text{ insatisfactorio} \end{aligned}$$

La representación gráfica de los Z-score corresponde a la puntuación Z calculada con la variabilidad aceptable según la desviación típica robusta.

Los valores asignados se determinan mediante valores consensuados entre los laboratorios participantes. Todos deben tener trazabilidad e incertidumbre de medida. Los participantes no conocen los valores asignados durante el desarrollo del ejercicio.

Los valores asignados para los ensayos de aptitud organizados no podrán ser conocidos por los participantes, sino hasta la generación del informe final. Por lo cual, la información que maneja el personal que coordina los ensayos de aptitud y los colaboradores que participen en el proceso de asignación del valor, evaluación estadística o generación del informe, no podrá ser divulgada a terceros o a los participantes.

Las gráficas con los resultados de z-score obtenidos por los laboratorios para cada analito y matriz se detallan en el ANEXO III

10. VARIABLES METODOLÓGICAS Y DESEMPEÑO

El principal objetivo de este ejercicio de intercomparación no radica únicamente en evaluar la competencia técnica individual de los participantes mediante la asignación de un z-score,





sino en identificar aquellas fuentes de variabilidad analítica que condicionan la dispersión global de los resultados. En la determinación de contaminantes orgánicos traza en matrices ambientales complejas, la coelución de interferentes, la carga orgánica latente y los fenómenos físicos de supresión o mejora de la ionización en la fuente de la interfase constituyen un gran desafío analítico.

Con el propósito de establecer un diagnóstico robusto sobre las mejores prácticas disponibles, en este apartado se realiza un análisis de correlación cruzada entre las características operacionales y metodológicas declaradas por cada laboratorio (estrategias de pretratamiento, técnicas de extracción, configuraciones cromatográficas y metodologías de calibración cuantitativa) y la magnitud de las desviaciones obtenidas expresadas en valor absoluto ($|z\text{-score}|$).

A través de esta aproximación estadística, se pretende discernir si los resultados clasificados como cuestionables o insatisfactorios responden a desviaciones aleatorias o si, por el contrario, están intrínsecamente vinculados a limitaciones sistemáticas de configuraciones metodológicas específicas frente a la complejidad de las matrices evaluadas.

Para evaluar como influyen las metodologías analíticas, se han cruzado los resultados del ejercicio ($z\text{-score}$) con las respuestas al cuestionario sobre las características técnicas. En el **ANEXO IV** se presenta un resumen de los datos suministrados por los laboratorios. Dado que el objetivo es medir la magnitud total de la desviación analítica con respecto al valor asignado (independientemente de la dirección del error), se transformaron todos los resultados a su valor absoluto $|z|$. De este modo, una desviación de +1.5 y una de -1.5 se evalúan con la misma importancia (1.5), permitiendo comparar la dispersión real de cada técnica.

10.1. Evaluación de significancia estadística

Con el objetivo de determinar si las diferencias observadas en las magnitudes de los $z\text{-scores}$ son estadísticamente significativas o si corresponden a la variabilidad aleatoria inherente al ejercicio, se aplicó el test no paramétrico de Kruskal-Wallis.

Esta prueba estadística es la alternativa no paramétrica al Análisis de Varianza (ANOVA) de un factor y se seleccionó debido a que los $z\text{-scores}$ del ejercicio no cumplen con los principios de normalidad y homogeneidad de varianzas (homocedasticidad), una característica habitual en los datos de rendimiento de ejercicios intercomparativos donde abundan los valores cercanos a cero y distribuciones asimétricas con presencia de valores atípicos (*outliers*).

El test evalúa la hipótesis nula (H_0) de que las medianas de los $z\text{-scores}$ son iguales entre los diferentes grupos metodológicos comparados (por ejemplo, entre los distintos métodos de cálculo o las técnicas de extracción), frente a la hipótesis alternativa (H_1) de que al menos uno de los métodos presenta una distribución de errores sistemáticamente diferente a los demás. El análisis se ejecutó fijando un nivel de significancia estándar del 5% ($\alpha=0.05$), asumiendo que un valor de p inferior a este umbral ($p < 0.05$) confirma la existencia de un impacto crítico y estadísticamente demostrable de dicha variable analítica en la calidad del resultado reportado por los laboratorios.

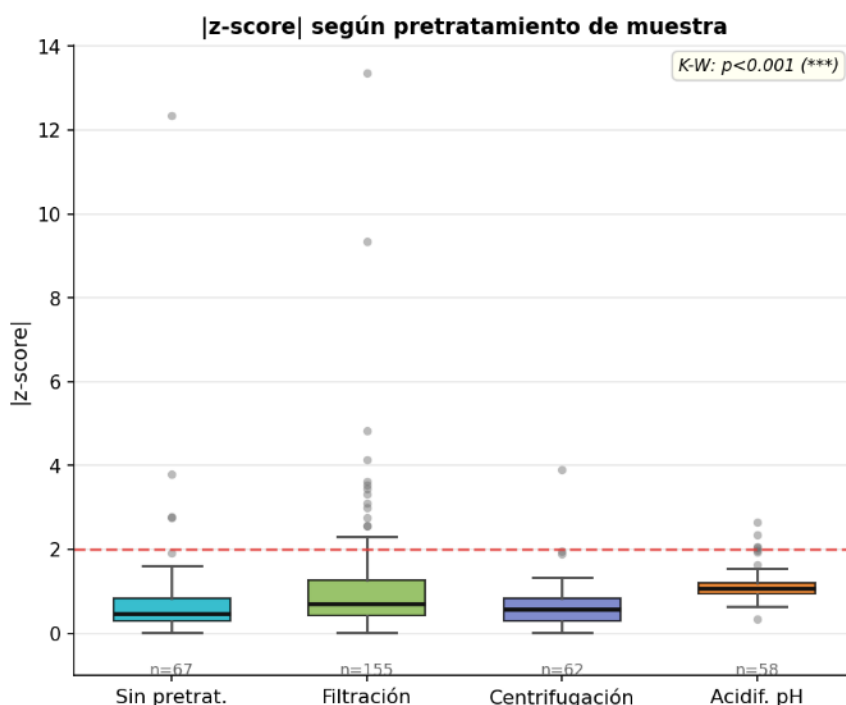


Rendimiento general por laboratorio

Los 12 laboratorios muestran un rango de desempeño amplio ($|z|$ medio de 0,34 a 2,45). La mayoría obtiene resultados satisfactorios ($|z| < 2$), con Lab.01, Lab.09 y Lab.12 alcanzando el 100 % y 83 % de satisfactorios respectivamente, mientras que Lab.07 registra el peor rendimiento ($|z|$ medio = 2,45; 20 % de resultados insatisfactorios). Se identificaron 6 valores de z-score no reportados (Lab.11 y Lab.04 para ciertos compuestos en medio receptor) que fueron excluidos del análisis estadístico.

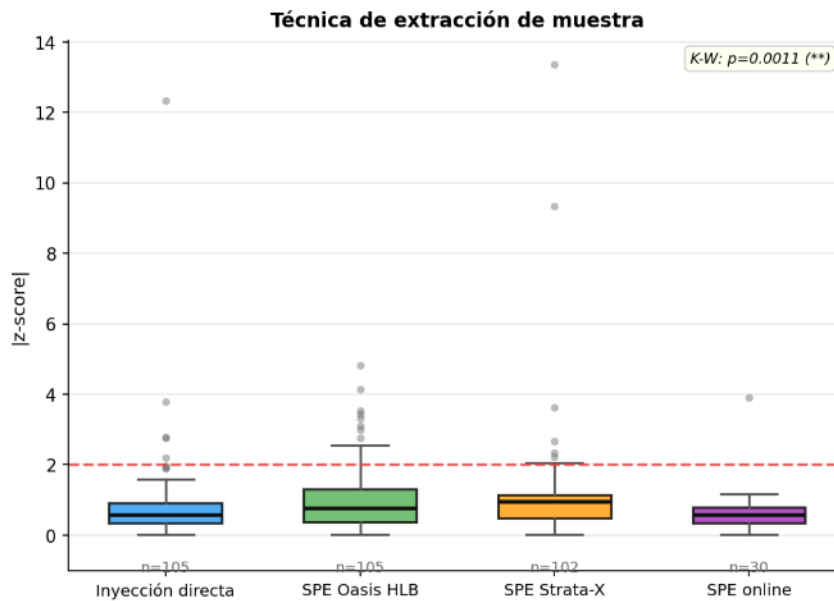
a) Pretratamiento de muestra — diferencia estadísticamente significativa (K-W $p < 0,001$)

Este es el factor con mayor impacto. Los laboratorios que utilizaron centrifugación obtuvieron los $|z|$ más bajos (media = 0,66; 98 % satisfactorios), seguidos de los que no realizaron pretratamiento (media = 0,85). En contraste, los que emplearon filtración mostraron mayor variabilidad y peor rendimiento medio (media = 1,10; 6,5 % insatisfactorios). La acidificación de pH dio resultados intermedios pero con baja dispersión. La diferencia entre centrifugación y filtración es significativa (Mann-Whitney $p = 0,009$).



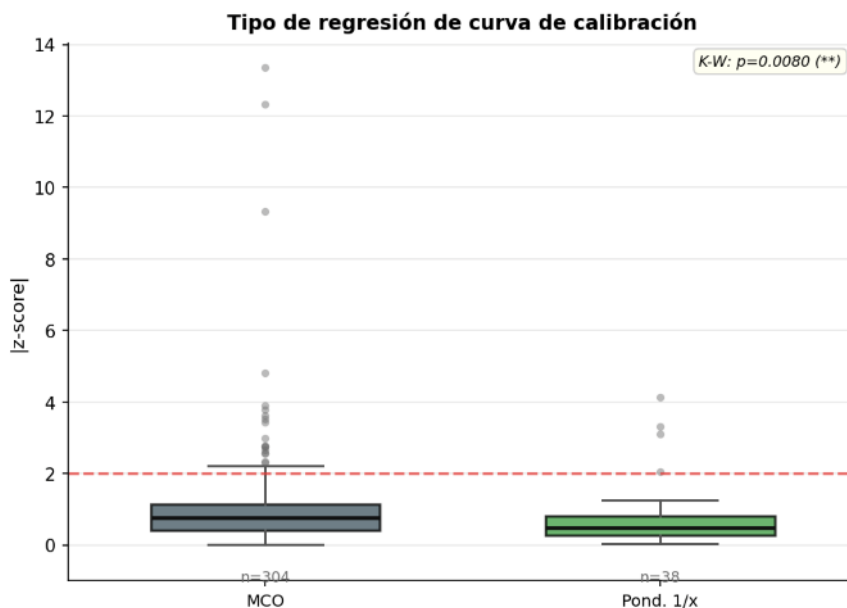
b. Técnica de extracción — diferencia significativa (K-W $p = 0,001$)

La inyección directa y la SPE online presentan los mejores perfiles (medias de 0,86 y 0,66; ~95-97 % satisfactorios). Los cartuchos SPE Oasis HLB y SPE Strata-X muestran mayores $|z|$ (medias 1,03 y 1,15) y mayor dispersión, aunque la diferencia entre inyección directa y SPE Oasis HLB no alcanza significación individual (Mann-Whitney $p = 0,06$). El resultado sugiere que la simplicidad del proceso de extracción favorece la trazabilidad de los resultados.



c. Tipo de regresión — diferencia significativa (K-W $p = 0,008$)

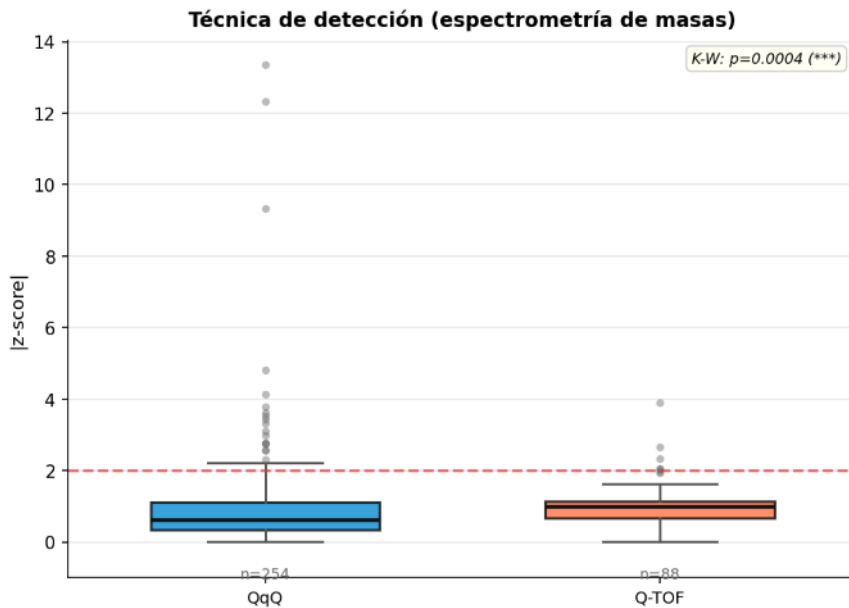
Los laboratorios que aplicaron regresión lineal ponderada con factor $1/x$ obtuvieron $|z|$ notablemente más bajos (media = 0,77; mediana = 0,48) frente a los que usaron mínimos cuadrados ordinarios (media = 1,01; Mann-Whitney $p = 0,008$). Este resultado es consistente con la literatura: la ponderación $1/x$ es más adecuada cuando la relación señal/concentración no es homocedástica, como ocurre habitualmente en LC-MS/MS a bajas concentraciones ambientales.





d. Técnica de detección — diferencia estadística, pero efecto mínimo (K-W $p = 0,0004$)

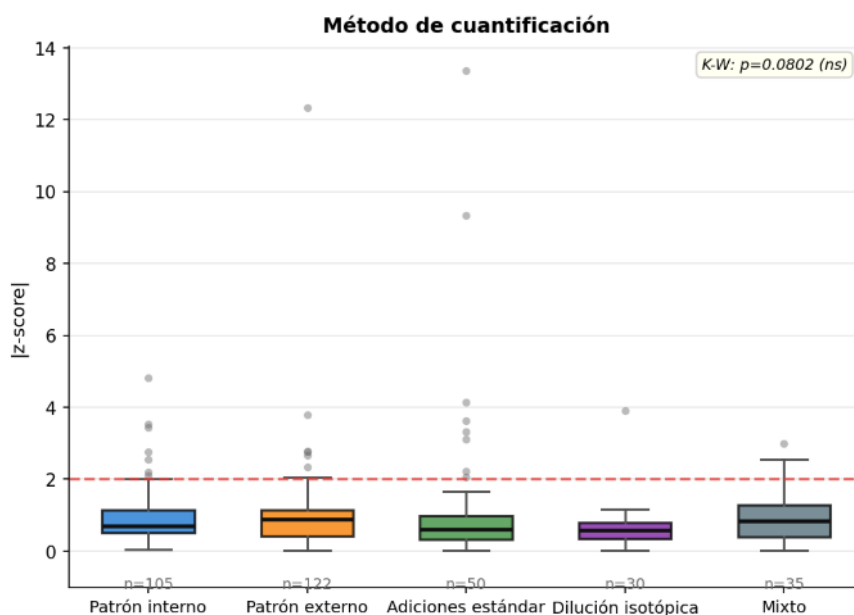
Aunque el test Kruskal-Wallis detecta diferencia entre QqQ y Q-TOF, las medias son prácticamente idénticas (0,98 vs. 0,98). La significación estadística se debe probablemente al desbalance en el número de resultados ($n = 260$ vs. 88) y a la gran varianza intragrupo. En términos prácticos, la técnica MS no es un factor discriminante de rendimiento en este ejercicio.



e. Método de cuantificación — tendencia no significativa (K-W $p = 0,08$)

La dilución isotópica (Lab.08) ofrece el mejor rendimiento (media = 0,66; 97 % satisfactorios), seguida del patrón interno (media = 0,90). Las adiciones estándar presentan la mayor variabilidad (media = 1,30; 12 % insatisfactorios), arrastradas principalmente por Lab.07. La diferencia patrón interno vs. patrón externo no es significativa ($p = 0,33$). La tendencia apunta a que el uso de marcadores isotópicos favorece la exactitud, pero el bajo número de laboratorios por categoría limita la potencia estadística.





Recomendación metodológica: La combinación de pretratamiento por centrifugación, extracción simplificada (inyección directa o SPE online) y regresión ponderada 1/x aparece como la más robusta en este intercomparativo. La adición de marcadores de dilución isotópica, cuando sea posible, ofrece ventajas adicionales de exactitud.

11. DISCUSIÓN DE RESULTADOS

Se han desarrollado y optimizado metodologías analíticas de alto rendimiento orientadas a la determinación de 12 contaminantes emergentes. Asimismo, se ha organizado un ejercicio de comparación interlaboratorios con un doble propósito:

- a) colaborar con otras entidades en la validación de las metodologías desarrolladas, identificando las técnicas instrumentales con mejor exactitud y precisión para cada compuesto.
- b) Apoyar en la evaluación de la eficiencia de las EDAR con tratamientos terciarios/cuaternarios e implementar estudios experimentales orientados a la valoración del potencial de reutilización del agua tratada con fines agrícolas.

Comentarios generales – abordaje estadístico clásico/robusto

1. En la matriz influente, los valores de VA (ppt) son inferiores a los obtenidos en la matriz efluente para 7 analitos. Esta aparente anomalía puede estar relacionada con limitaciones de las técnicas analíticas empleadas, como la falta de resolución instrumental o la proximidad al límite de cuantificación.
2. La presencia de distribuciones bimodales en los datos contribuye, de forma general, a un incremento de la variabilidad observada, reflejado en valores más elevados de RSD %.



3. La variabilidad del ejercicio no depende exclusivamente de la concentración del mensurando, sino que está condicionada por múltiples factores, entre los que destacan el número de participantes, la proporción de valores discrepantes, el nivel de concentración (ppt), las condiciones analíticas empleadas para cada compuesto y la proximidad al LoQ, entre otros.
4. A partir de los resultados obtenidos, se ha establecido de forma orientativa un rango de referencia para el RSD situado entre el 30 % y el 60 %, en función del analito y de la matriz evaluada.

12. CONCLUSIONES

A continuación, se presentan las conclusiones preliminares derivadas del tratamiento estadístico y de la evaluación técnica de los resultados obtenidos en el ejercicio de intercomparación, con el objetivo de aportar evidencia experimental para la validación del método analítico.

1ª.- Reproducibilidad del ejercicio de intercomparación: σ_{objetivo}

Tanto en los ensayos de homogeneidad como en la evaluación del desempeño de los laboratorios participantes, se adoptó como criterio de desviación estándar objetivo (σ_{objetivo}) el 25 % del valor medio obtenido en los estudios de homogeneidad y el 25 % de los respectivos valores asignados en la evaluación mediante z-score.

La adopción de este criterio se fundamenta en la ausencia de datos de ejercicios interlaboratorio previos suficientemente consolidados para la determinación simultánea de un número elevado de contaminantes emergentes en las matrices objeto de estudio.

No obstante, en el ámbito del análisis de residuos de pesticidas a niveles de trazas (ppt), donde existe una mayor disponibilidad de antecedentes bibliográficos y estudios interlaboratorio, es habitual emplear criterios de aceptación equivalentes.

Al comparar los valores experimentales de desviación estándar de reproducibilidad (S_R) obtenidos para cada combinación analito/matriz con los valores estimados de σ_{objetivo} , se observa que, en un número significativo de casos, los valores de S_R resultan superiores al criterio establecido. Este comportamiento pone de manifiesto la elevada complejidad analítica asociada a la determinación de estos compuestos a niveles ultratrazas, así como la influencia de las diferentes estrategias instrumentales y de tratamiento de datos empleadas por los laboratorios participantes.

2ª.- Presencia de bimodalidad de los resultados experimentales

En determinados analitos y matrices se observó una distribución bimodal de los resultados reportados por los laboratorios participantes. Entre las posibles causas asociadas a este comportamiento pueden señalarse las siguientes:

- ✓ Empleo de diferentes técnicas cromatográficas e instrumentales.



- ✓ Aplicación de distintos procedimientos de pretratamiento y extracción de las muestras.
- ✓ Utilización de diferentes estrategias de cuantificación (patrón externo, patrón interno, método de adiciones estándar, entre otras).
- ✓ Aplicación de distintos modelos de regresión en las curvas de calibración:
 - regresión lineal por mínimos cuadrados
 - regresión lineal ponderada y diferente factor de ponderación empleado ($1/x$, $1/x^2$, etc.)

La selección del modelo de regresión adquiere especial relevancia en rangos de concentración extremadamente bajos (ng/L – ppt), donde pequeñas variaciones en el ajuste de calibración pueden traducirse en diferencias significativas en los resultados cuantificados entre laboratorios, originando valores elevados de %RSD.

Desde el punto de vista quimiométrico, en niveles próximos al límite de cuantificación (LQ), el empleo de regresiones ponderadas suele proporcionar un mejor ajuste de los datos experimentales que la regresión lineal clásica por mínimos cuadrados. Esto permite reducir el error relativo (%RE) frente a los patrones de calibración y mejorar los intervalos de confianza asociados a la cuantificación.

Asimismo, si el rango de linealidad del método no se encuentra correctamente establecido, algunos laboratorios podrían estar realizando cuantificaciones en niveles inferiores al LQ efectivo del método, circunstancia que incrementa considerablemente la dispersión de los resultados experimentales y contribuye a la aparición de distribuciones bimodales.

3ª.- Comparativa resultados de estadística clásica frente estadística robusta

El estudio incluye tablas comparativas resumen para todas las combinaciones analito/matriz, recopilando los principales parámetros obtenidos mediante diferentes aproximaciones estadísticas:

- ✓ valor asignado
- ✓ desviación estándar de reproducibilidad (SR)
- ✓ coeficiente de variación relativo (RSD%)
- ✓ distribución de los datos
- ✓ valores de z-score.

La aplicación conjunta de estadística clásica y estadística robusta ha permitido obtener información complementaria de elevado interés técnico para la interpretación global de los resultados y la evaluación del comportamiento analítico de los laboratorios participantes.

De forma general, se observa un incremento de la variabilidad analítica (%RSD) asociado principalmente a los siguientes factores:

- ✓ disminución de la concentración del analito;
- ✓ incremento de la complejidad de la matriz y, por tanto, del efecto matriz asociado (aguas superficiales < aguas residuales efluentes < aguas residuales influentes).



Adicionalmente, la diferente naturaleza fisicoquímica de los compuestos estudiados constituye un factor adicional de variabilidad, debido tanto a las distintas respuestas instrumentales como a la diferente susceptibilidad de cada analito frente a los efectos de matriz presentes en muestras ambientales complejas.

Los valores elevados de S_{RPT} y %RSD observados para determinadas combinaciones analito/matriz podrían considerarse indicativos de que las concentraciones determinadas se encuentran próximas o incluso por debajo del límite de cuantificación efectivo del método. En este contexto, dichos resultados podrían reflejar limitaciones asociadas al límite de cuantificación real del método o al límite de cuantificación instrumental (iLQ) alcanzable para determinados compuestos.

Por otra parte, los resultados de desempeño obtenidos por los laboratorios participantes deberían interpretarse con carácter orientativo, especialmente en aquellos casos en los que se haya observado bimodalidad en la distribución de los datos o una reproducibilidad interlaboratorio limitada.

A la vista de los resultados obtenidos, se recomienda el empleo de metodologías basadas en UHPLC-MS/MS, combinadas con el uso de patrones internos isotópicamente marcados y la aplicación de modelos de regresión ponderada, preferentemente con factor de ponderación $1/x^2$, especialmente en la cuantificación de compuestos presentes a niveles ultratraza.



13. DEFINICIONES

Para la correcta interpretación del presente documento se detallan las siguientes definiciones:

1. **Ensayo de aptitud:** Evaluación del desempeño de los participantes con respecto a criterios previamente establecidos a través de comparaciones interlaboratorios (Ref. ISO 17043)
2. **Ensayo Interlaboratorios:** Tipo de ensayo de aptitud que se realiza con una distribución simultánea a los participantes, se selecciona un material con un valor asignado o certificado contra el cual es comparado el resultado obtenido por el o los participantes. Sirve para indicar el desempeño individual o grupal de los laboratorios.
3. **Ítem de ensayo de aptitud:** Muestra, producto, artefacto, material de referencia, equipo, estándar de medición, conjunto de datos u otra información utilizada para evaluar el desempeño de los participantes en las pruebas de competencia.
4. **Desviación estándar para la evaluación de aptitud:** Medida de la dispersión utilizada en la evaluación de los resultados de la escala de valores, sobre los ensayos de aptitud basados en la información disponible.
5. **Valor asignado:** Valor que se asigna al material o ítem de ensayo utilizado en la ronda de intercomparación con la finalidad de comparar los resultados de los participantes con un valor cercano al “valor verdadero”.
6. **Método estadístico robusto:** método estadístico insensible a pequeñas desviaciones de las hipótesis iniciales del modelo probabilístico implícito.
7. **Técnicas estadísticas robustas:** Técnicas para minimizar la influencia que pueden tener los resultados extremos sobre los estimadores del promedio y de la desviación estándar.
8. **Mediana:** Valor de la variable de posición central en un conjunto de datos ordenados.
9. **Test de Grubbs:** Contraste de datos anómalos que compara la desviación del valor sospechoso y la media muestral con la desviación estándar de la muestra, este contraste supone una distribución de población normal.
10. **Valor anómalo:** Miembro de un conjunto de datos que es inconsistente con los otros miembros de dicho conjunto.
11. **Z-Score:** Índice de desempeño que estima el error que existe entre el resultado informado y el valor asignado del material de ensayo, y la desviación estándar establecida por el organizador del ensayo de aptitud.



14. REFERENCIAS

UNE-EN ISO/IEC 17043:2023. Evaluación de la conformidad – Requisitos General para los ensayos de aptitud.

UNE-EN ISO/IEC 17025:2017. Requisitos generales para la competencia de los laboratorios de ensayo y calibración.

ISO 13528:2022 (E). Statistical methods for use in proficiency testing by interlaboratory comparisons.

The International Harmonized Protocol for Proficiency testing of analytical chemistry laboratories (2006). Pure Appl. Chem. Vol78, pp 145-196.

UNE-EN ISO 9000:2015. Sistemas de gestión de calidad- Fundamentos y vocabulario.

UNE-EN ISO 17000.2004. Evaluación de la conformidad – Vocabulario y principios generales.

Guía EURACHEM “Selection, Use and Interpretation of Proficiency Testing (PT) Schemes – 3ª edición – 2021.

Analytical Methods Committee - Robust Statistic Part I & II. Analyst 114, 1693-1702 (1989).

Iria González-Mariño, Verónica Castro, Rosa Montes, Rosario Rodil, Ana Lores, Rafael Cela y José Benito Quintana: “Multi-residue determination of psychoactive pharmaceuticals, illicit drugs and related metabolites in wastewater by ultra-highperformance liquid chromatography-tandem mass spectrometry” Journal of Chromatography A, Vol.1569 – Septiembre 2018; 91-100.

Rosa Montes, Sandra Méndez, Julio Cobas, Nieves Carro, Teresa Neuparth, Nelson Alves, Miguel M. Santos, José Benito Quintana y Rosario Rodil: “Occurrence of persistent and mobile chemicals and other contaminants of emerging concern in Spanish and Portuguese wastewater treatment plants, transnational river basins and coastal water” Science of the Total Environment 885 (2023) 163737





15. ANEXOS

ANEXO I. DATOS PRELIMINARES. RESUMEN

Resumen preliminar de resultados en matriz agua superficial (medio receptor)

Analito	% RSD	VA (ng/l)	bimodal	nº datos	outliers	Codigo Lab.
Amisulpride	37,64	66,99	no	10	1 (C)	10
Benzotriazole	51,65	104,11	no	10	1 (G)	5
Candesartan (*)	83,94	14,94	sí	6	2 (G), 1 (C)	6 ; 10 / 10
Carbamazepine	28,41	23,55	no	10	3 (C)	5 ; 10 ; 12A
Citalopram	56,60	94,19	no	10	1 (G), 1 (C)	10 / 10
Clarithromycin	46,39	72,10	no	11	1 (G), 2 (C)	7 / 6 ; 10
Diclofenac	19,51	217,76	no	10	2 (G)	3 ; 6
Hydrochlorothiazide	32,72	247,98	sí	8	1 (G)	12A
Irbesartan	26,90	95,81	no	11	-	
Metoprolol	33,62	65,25	no	11	1 (G), 1 (C)	10 / 10
Venlafaxine	44,05	88,52	sí	11	1 (C)	10
4(5)Me-BT	42,97	105,79	sí	9	1 (C)	10

(G) = Discrepancia Grubbs - (C) = Discrepancia Cochran

OBSERVACIONES:

- ✓ Laboratorios discrepantes Test de Grubbs: 9
- ✓ 4 analitos con distribución posiblemente bimodal
- ✓ Promedio %RSD = 42,03



Resumen preliminar de resultados en matriz agua residual influente

Analito	% RSD	VA (ppt)	bimodal	nº datos	outliers	Codigo Lab.
Amisulpride	34,15	124,26	sí	10	1 (G)	8
Benzotriazole	69,99	176,74	sí	10	--	
Candesartan	46,14	181,25	sí	8	1 (C)	2
Carbamazepine	45,04	30,84	no	11	2 (C)	10 ; 11
Citalopram	57,91	78,67	sí	10	1 (C)	5
Clarithromycin	59,02	70,87	sí	10	1 (G), 1 (C)	5 ; 5
Diclofenac	70,04	255,26	sí	12	1 (G)	6
Hydrochlorothiazide	23,48	969,65	sí	8	1 (C)	10
Irbesartan	34,53	435,37	no	11	1 (C)	10
Metoprolol	78,15	12,29	sí	7	--	
Venlafaxine	39,04	330,50	no	11	--	
4(5)Me-BT	51,30	330,60	sí	10	1 (C)	10

(G) = Discrepancia Grubbs - (C) = Discrepancia Cochran

OBSERVACIONES:

- ✓ Laboratorios discrepantes Test de Grubbs: 3
- ✓ 9 analitos con distribución posiblemente bimodal
- ✓ Promedio %RSD = 50,73



Resumen preliminar de resultados en matriz agua residual efluente

Analito	% RSD	VA (ppt)	bimodal	nº datos	outliers	Codigo Lab.
Amisulpride	27,97	187,25	sí	10	1 (C)	10
Benzotriazole	62,01	216,77	sí	10	1 (G), 1 (C)	5 ; 10
Candesartan	43,26	209,02	sí	8	1 (C)	3
Carbamazepine	39,33	29,30	no	11	--	
Citalopram	72,71	59,51	no	10	1 (C)	10
Clarithromycin	34,92	43,84	sí	10	--	
Diclofenac	46,77	553,21	sí	12	1 (C)	10
Hydrochlorothiazide	33,59	1302,96	no	8	1 (G), 2 (C)	3 / 6; 10
Irbesartan	43,21	478,76	no	11	1 (C)	10
Metoprolol	34,06	19,78	no	7	1 (G), 2 (C)	7 / 7;10
Venlafaxine	36,40	265,73	no	11	1 (G), 3 (C)	7 / 5 ; 7;10
4(5)Me-BT	54,28	469,43	sí	10	2 (C)	7 ; 10

(G) = Discrepante Grubbs - (C) = Discrepante Cochran

OBSERVACIONES:

- ✓ Laboratorios discrepantes Test de Grubbs: 4
- ✓ 6 analitos con distribución posiblemente bimodal
- ✓ Promedio %RSD = 44,04

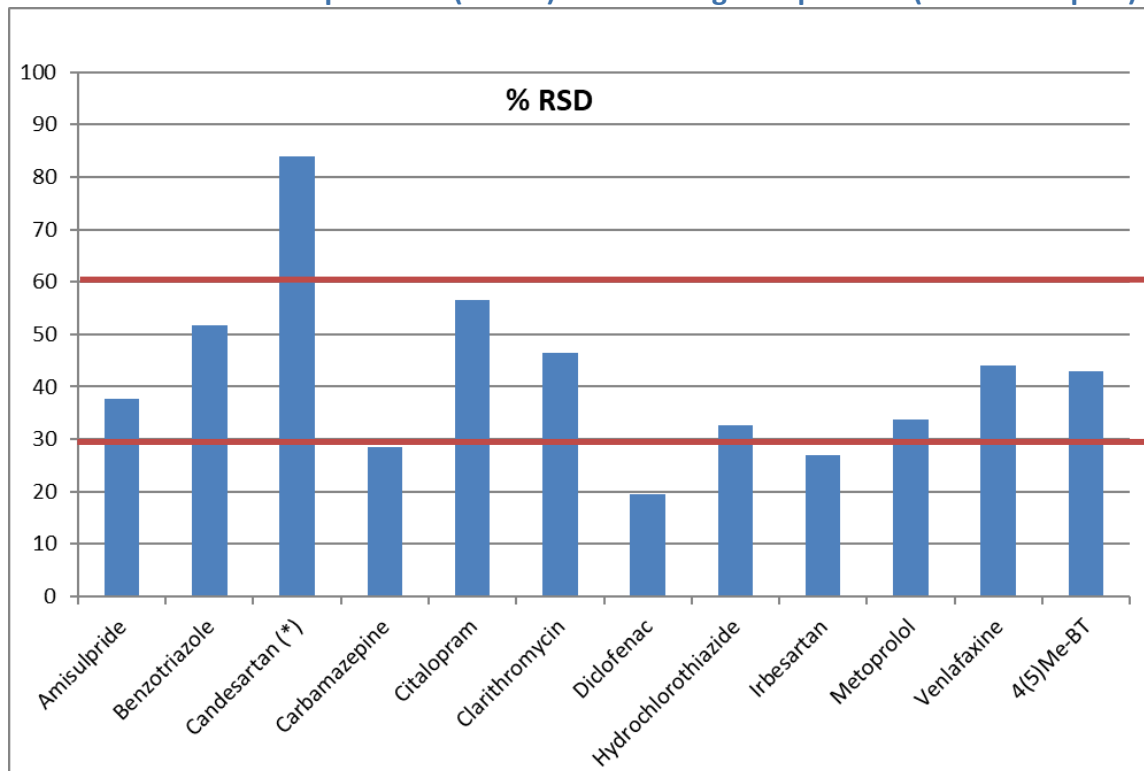




ANEXO II. GRAFICAS DE PRECISI3N (RSD%)

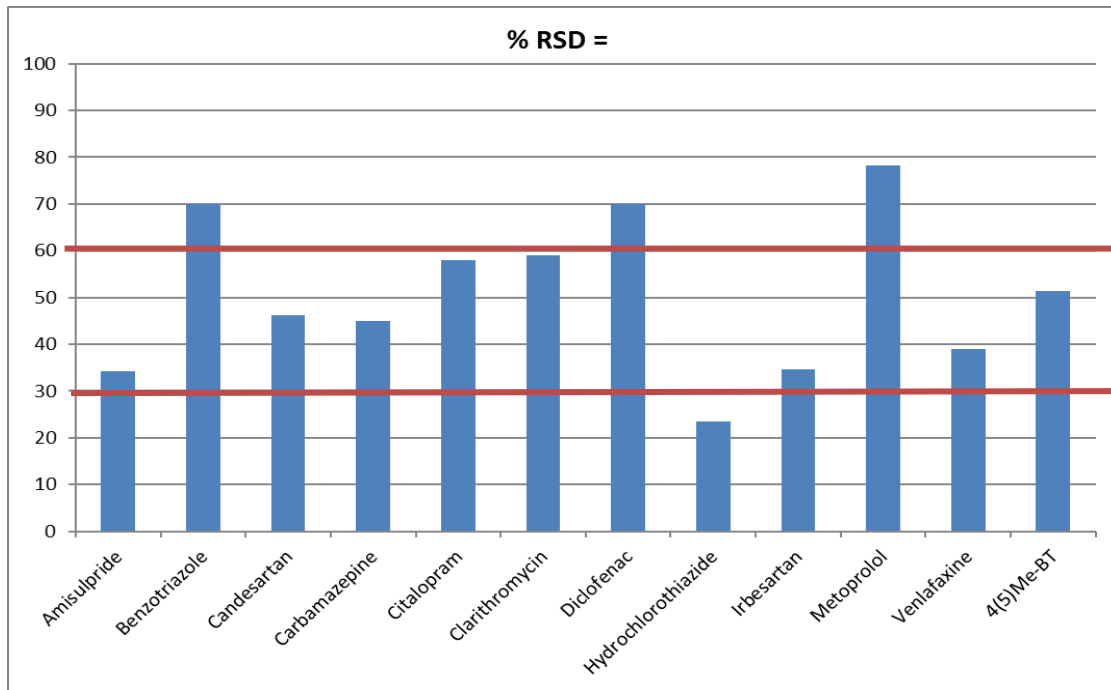
En base a los resultados obtenidos, se ha establecido un valor de referencia arbitrario para el RSD %, situado entre 30% - 60%

Resumen valores de precisi3n (RSD %) en matriz agua superficial (medio receptor)

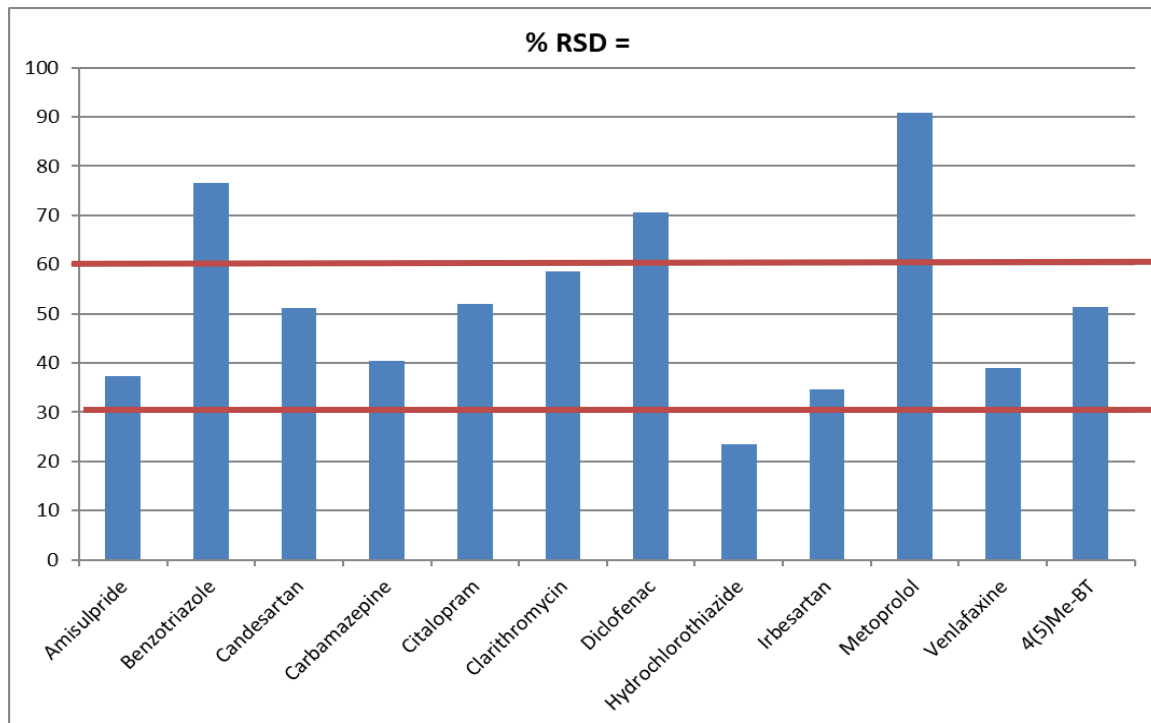




Resumen preliminar de resultados en matriz agua residual influente



Resumen preliminar de resultados en matriz agua residual efluente





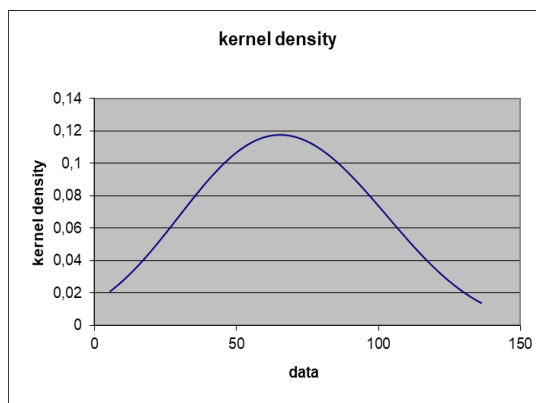
ANEXO III. GRÁFICAS DE RESULTADOS Z-SCORE DE LOS LABORATORIOS

Evaluación del desempeño y representación gráfica. Valores de z-score ordenados de cada participante

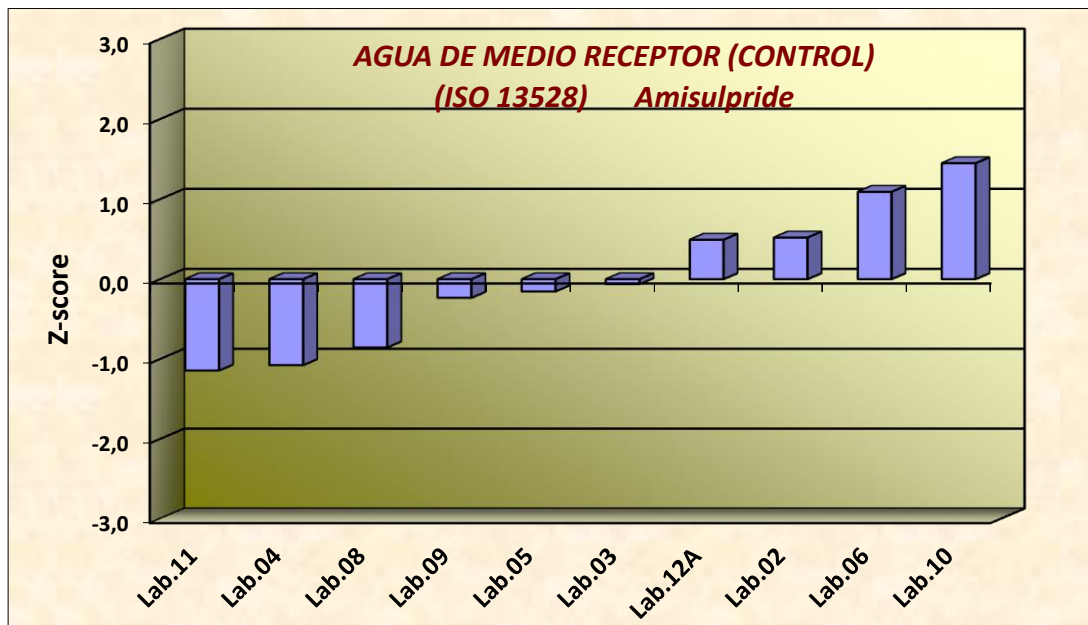
Z norm: valor de z-score normalizado calculado por estadística robusta empleado en la evaluación del desempeño

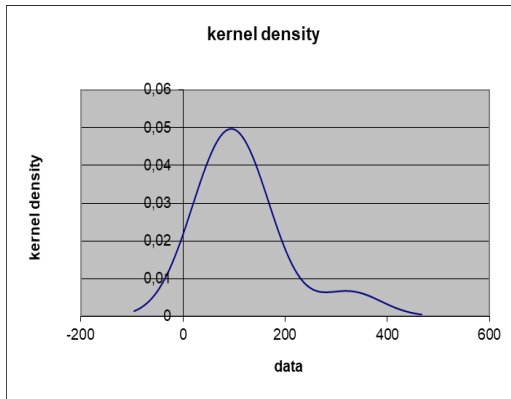
Z (σ_{objetivo}) (25% VA): valor de z-score basado en una desviación estándar objetivo del 25% del valor asignado

A.- MATRIZ AGUA SUPERFICIAL (MEDIO RECEPTOR)

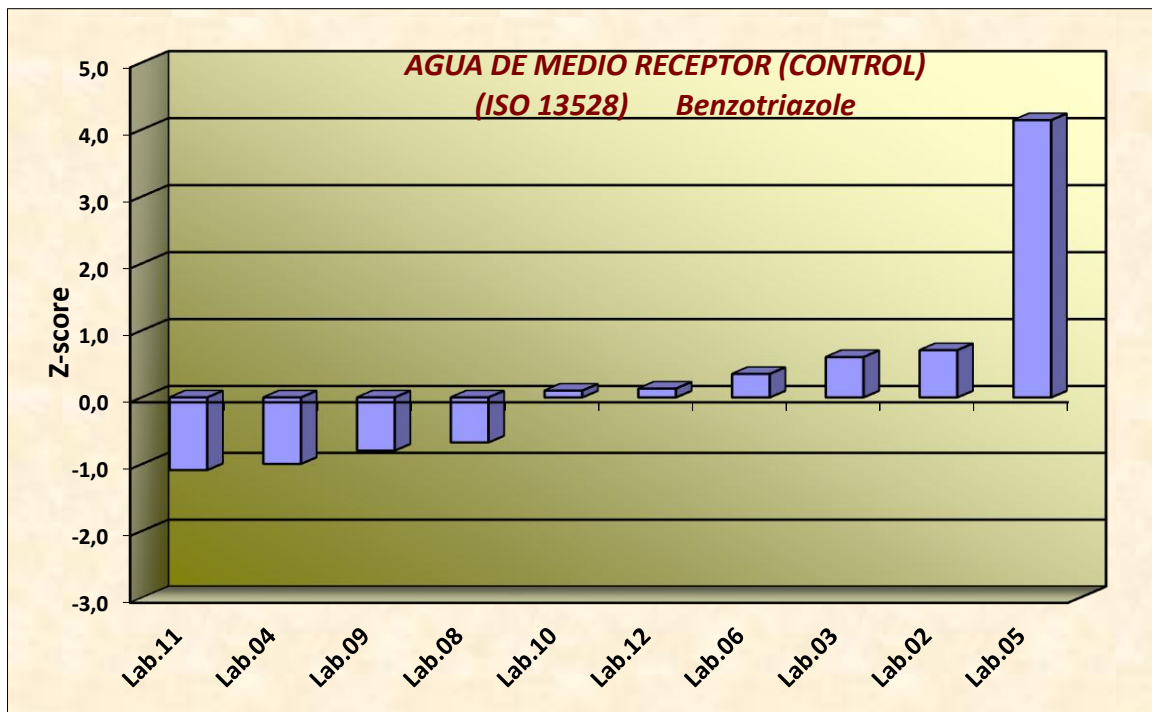


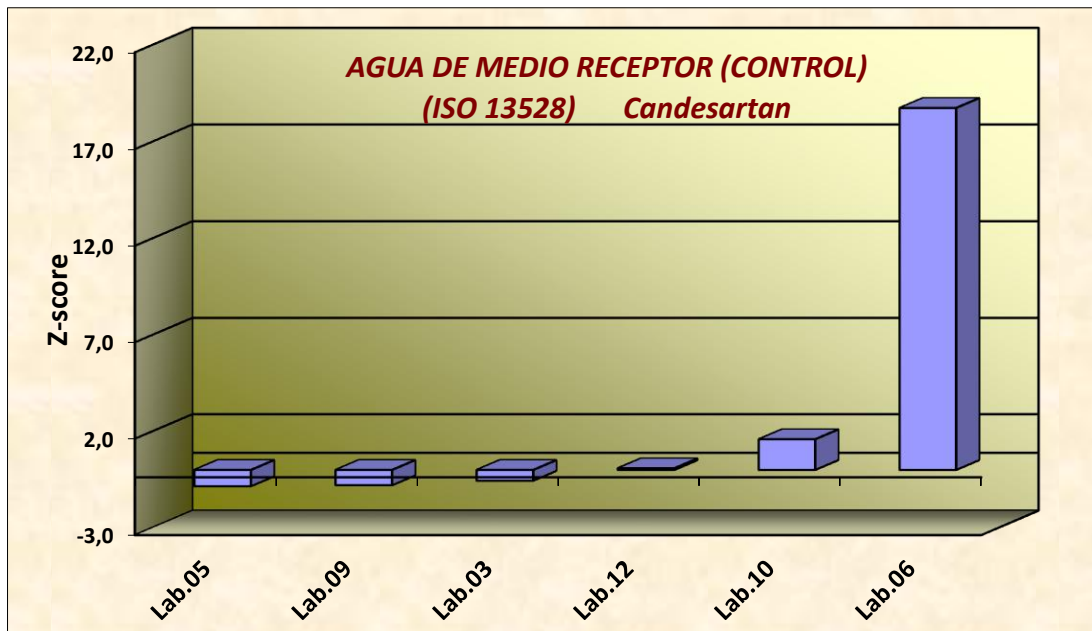
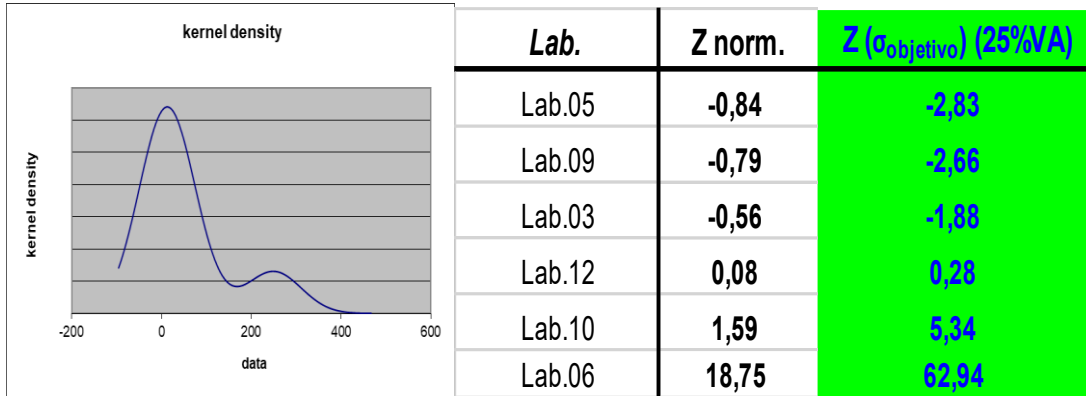
Lab.	Z norm.	Z (σ_{objetivo}) (25%VA)
Lab.11	-1,14	-1,72
Lab.04	-1,08	-1,62
Lab.08	-0,85	-1,29
Lab.09	-0,23	-0,35
Lab.05	-0,15	-0,23
Lab.03	-0,06	-0,09
Lab.12	0,49	0,74
Lab.02	0,52	0,78
Lab.06	1,09	1,64
Lab.10	1,45	2,18

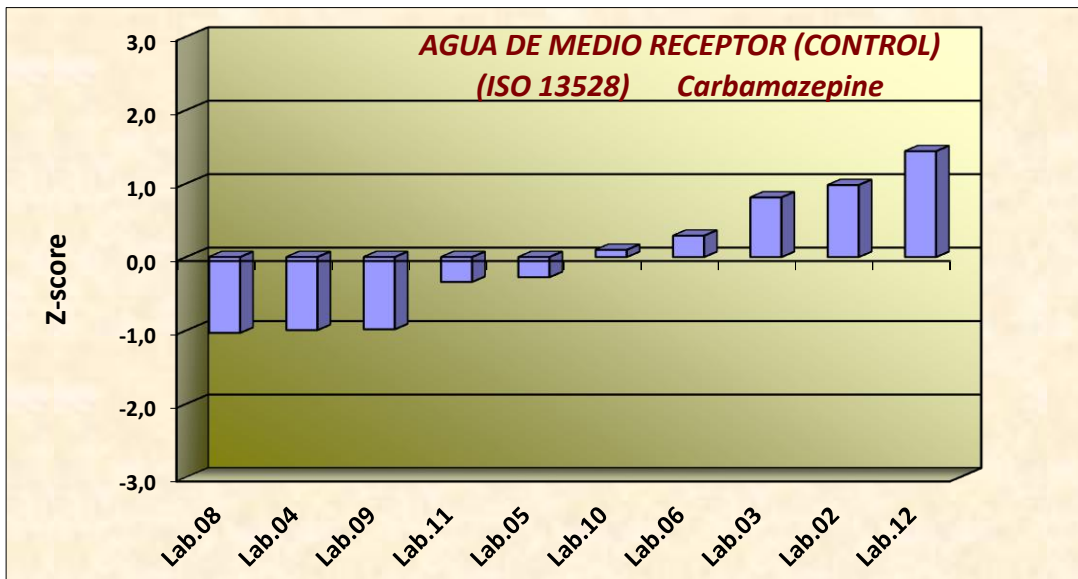
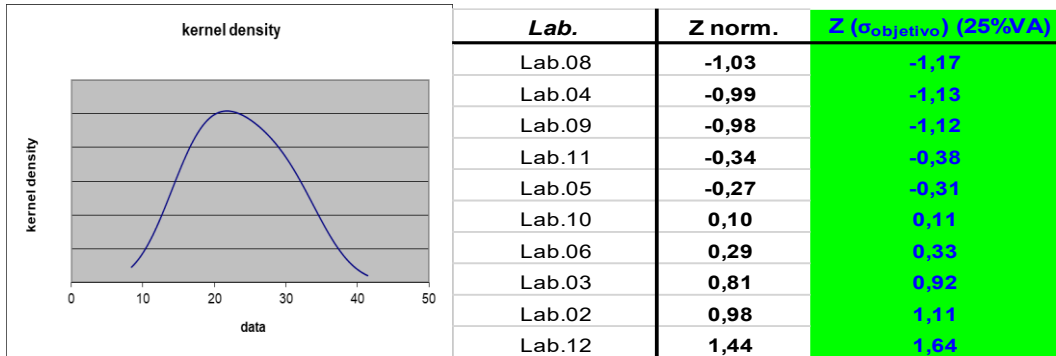


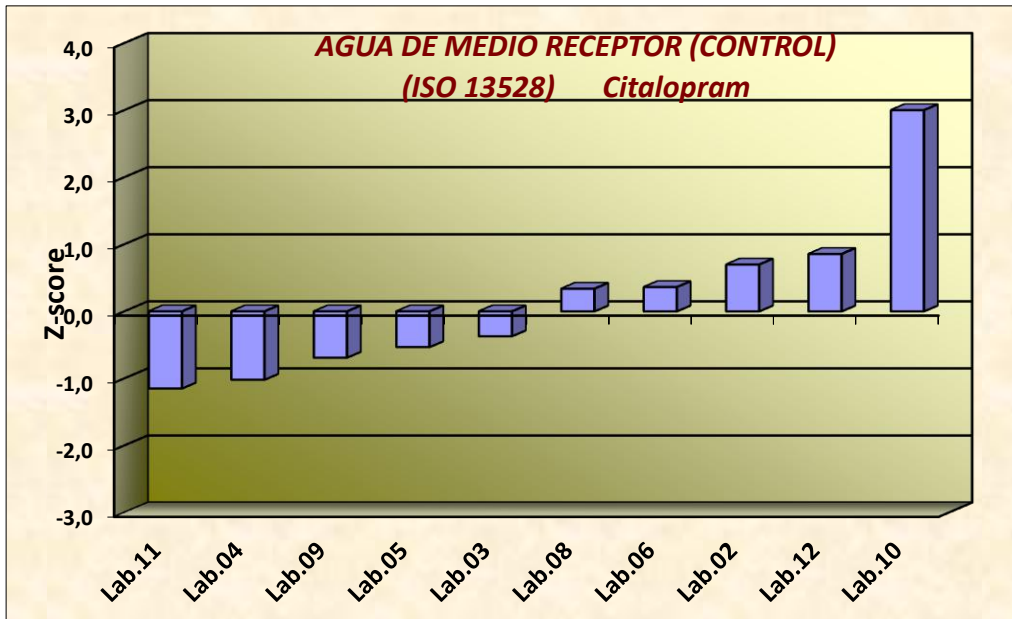
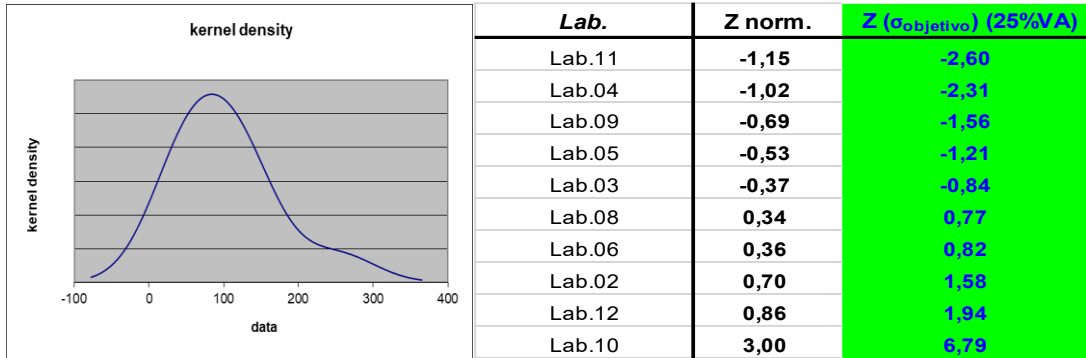


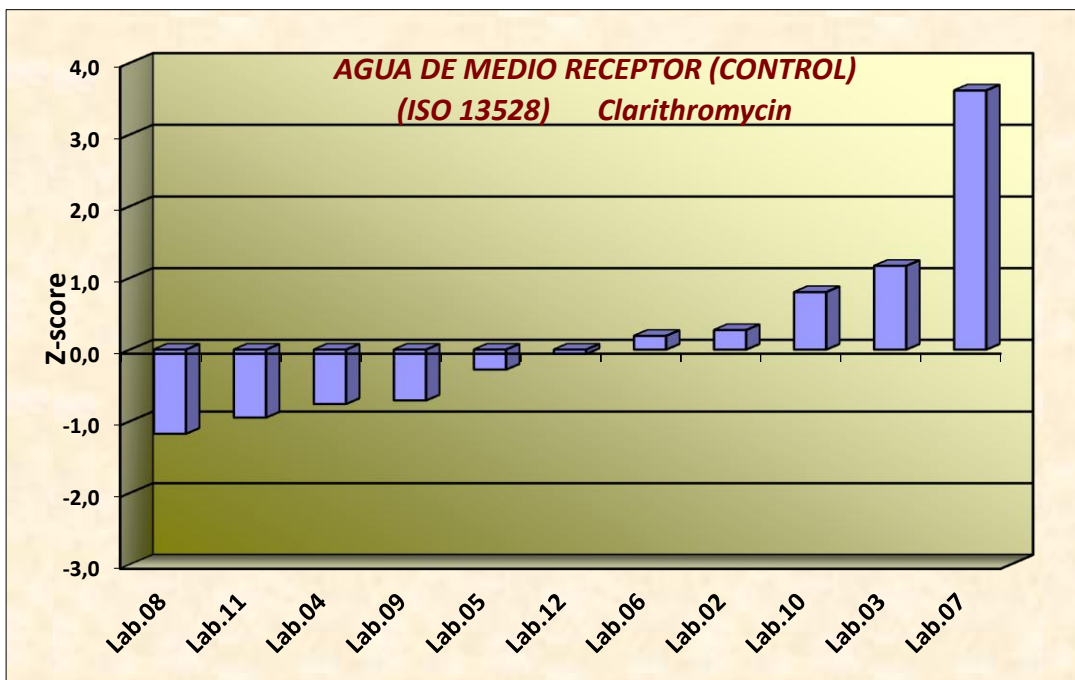
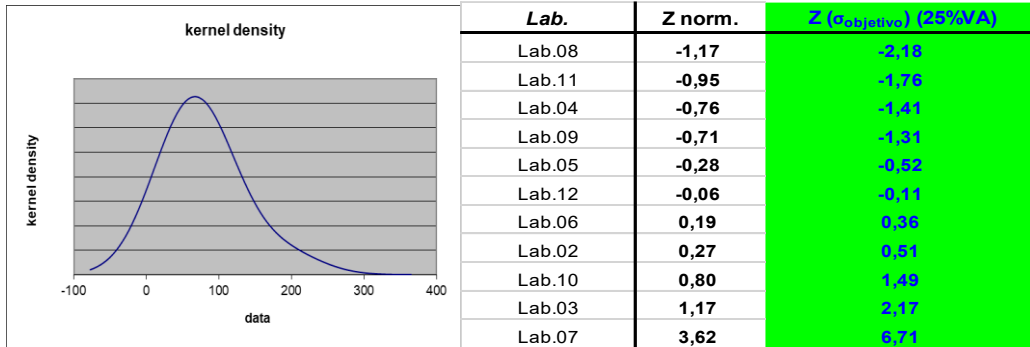
Lab.	Z norm.	Z (Objetivo) (25%VA)
Lab.11	-1,09	-2,24
Lab.04	-0,99	-2,05
Lab.09	-0,79	-1,64
Lab.08	-0,67	-1,39
Lab.10	0,10	0,21
Lab.12	0,13	0,27
Lab.06	0,35	0,73
Lab.03	0,60	1,25
Lab.02	0,71	1,46
Lab.05	4,14	8,55

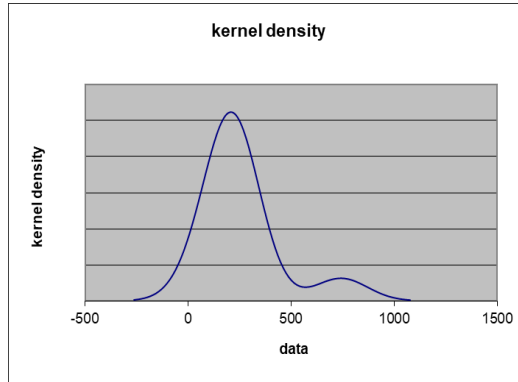




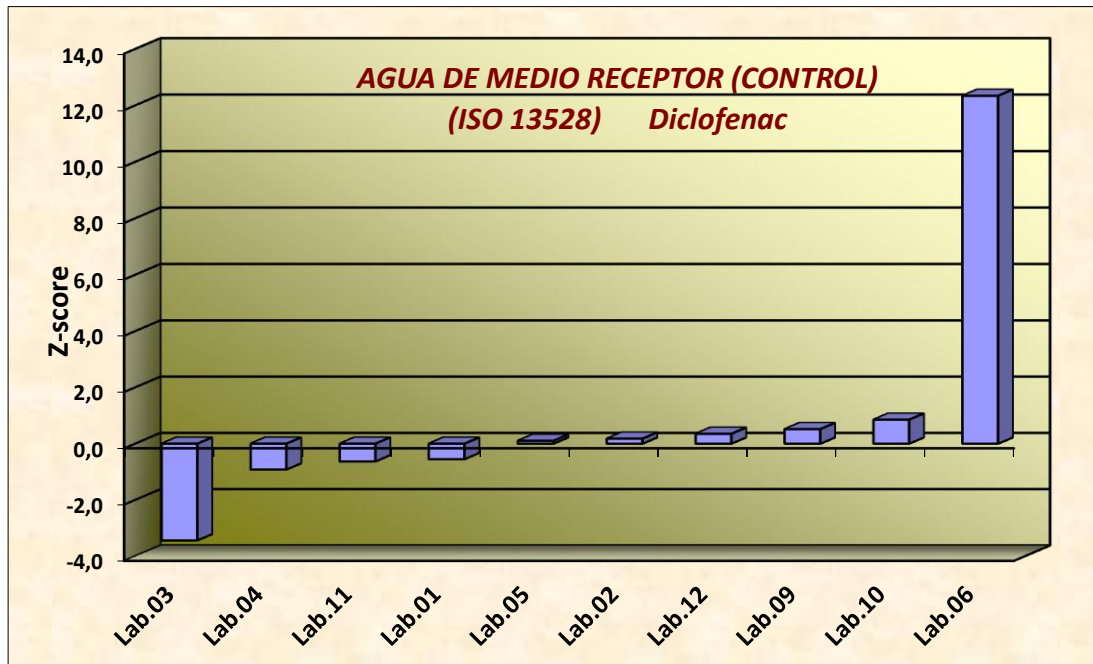


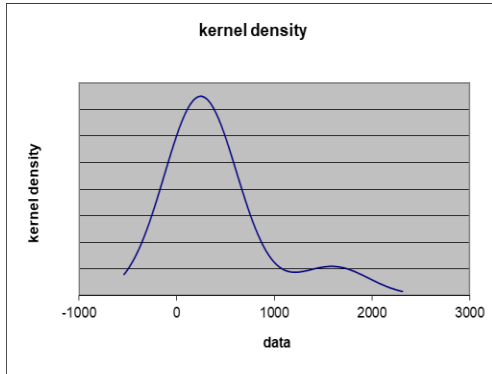




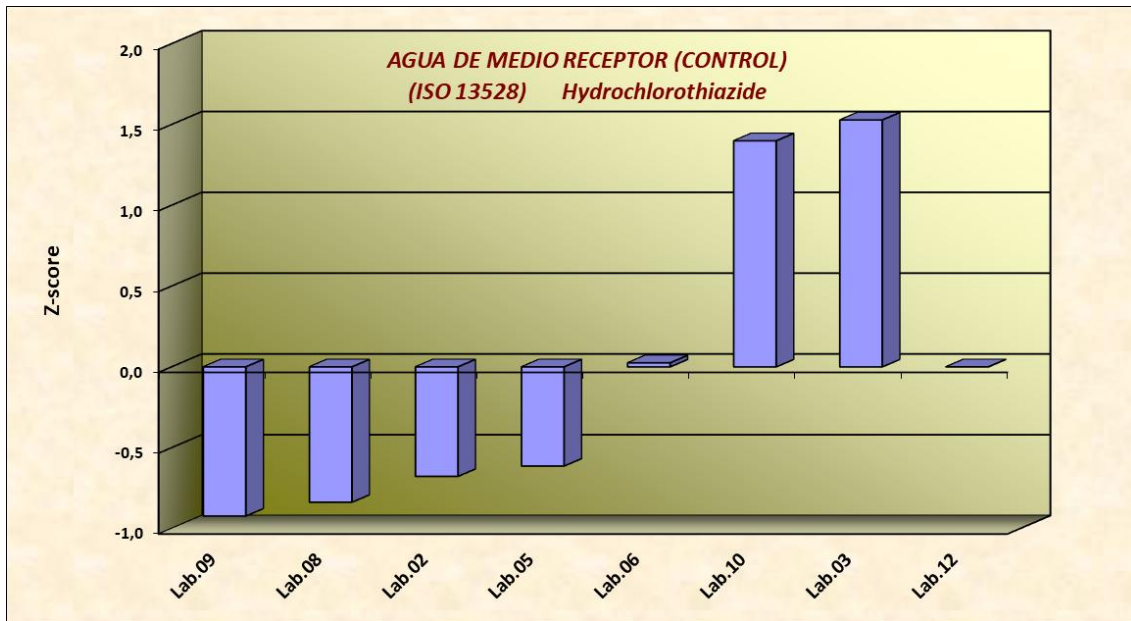


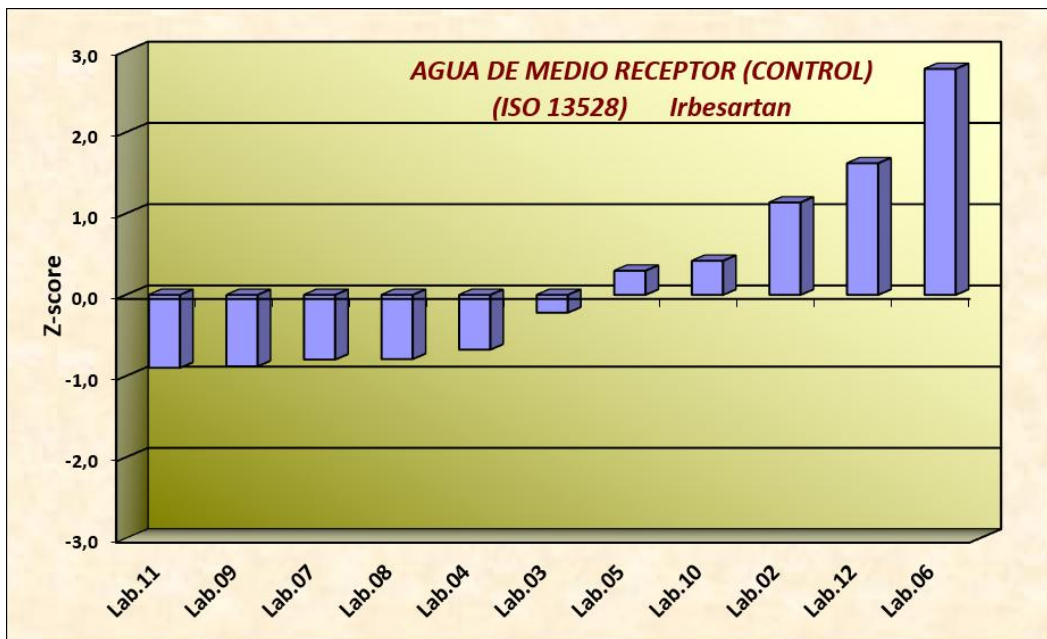
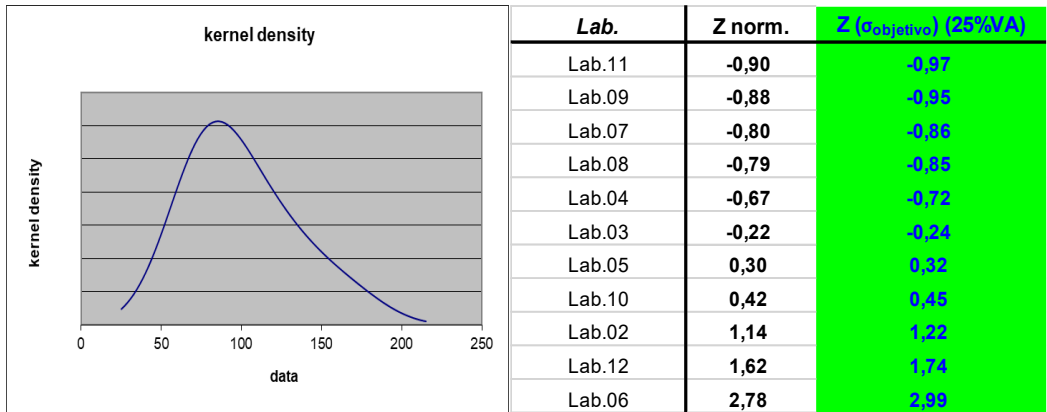
Lab.	Z norm.	Z (Objetivo) (25%VA)
Lab.03	-3,44	-2,68
Lab.04	-0,92	-0,72
Lab.11	-0,63	-0,49
Lab.01	-0,55	-0,43
Lab.05	0,10	0,08
Lab.02	0,18	0,14
Lab.12	0,35	0,28
Lab.09	0,52	0,40
Lab.10	0,85	0,67
Lab.06	12,34	9,63

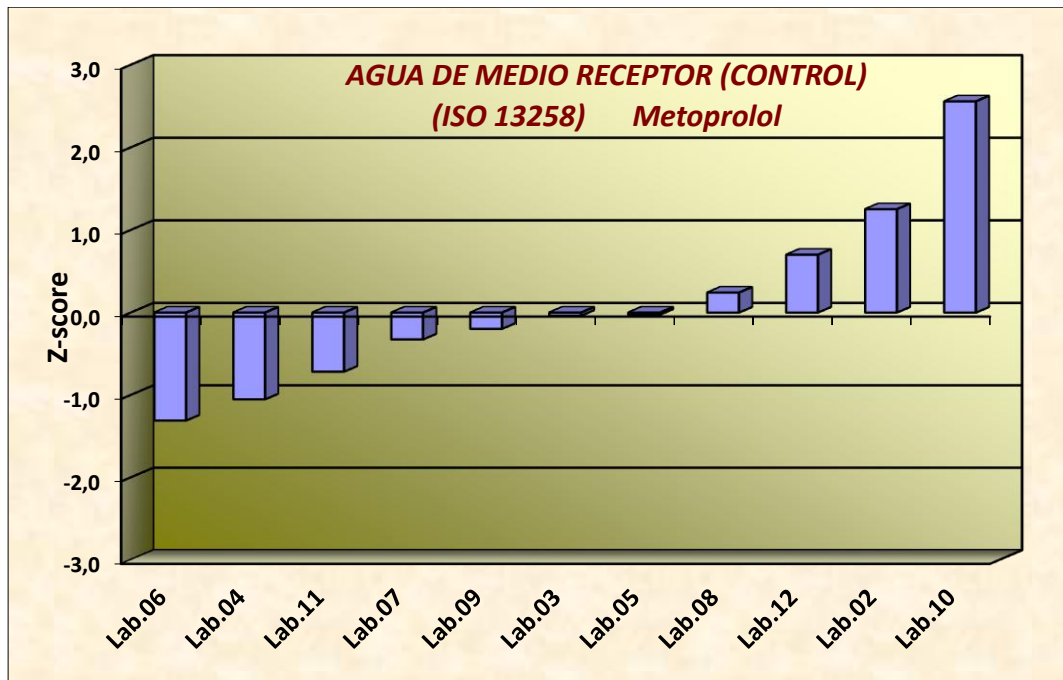
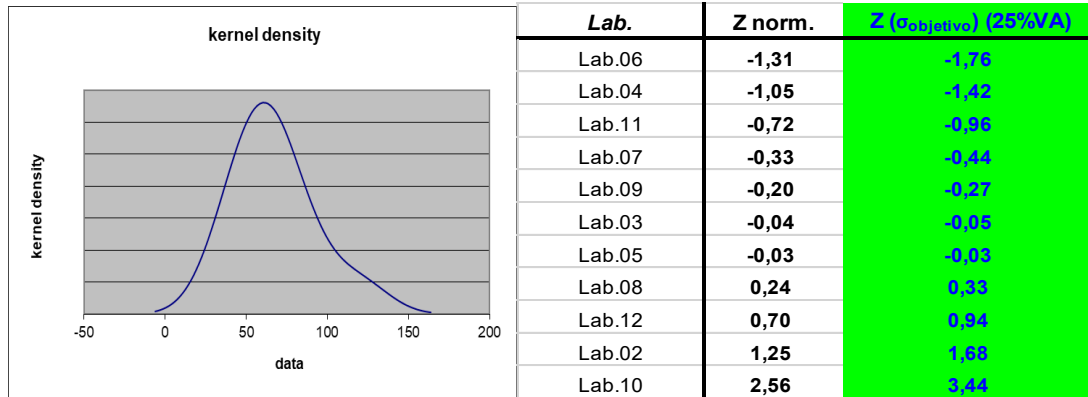


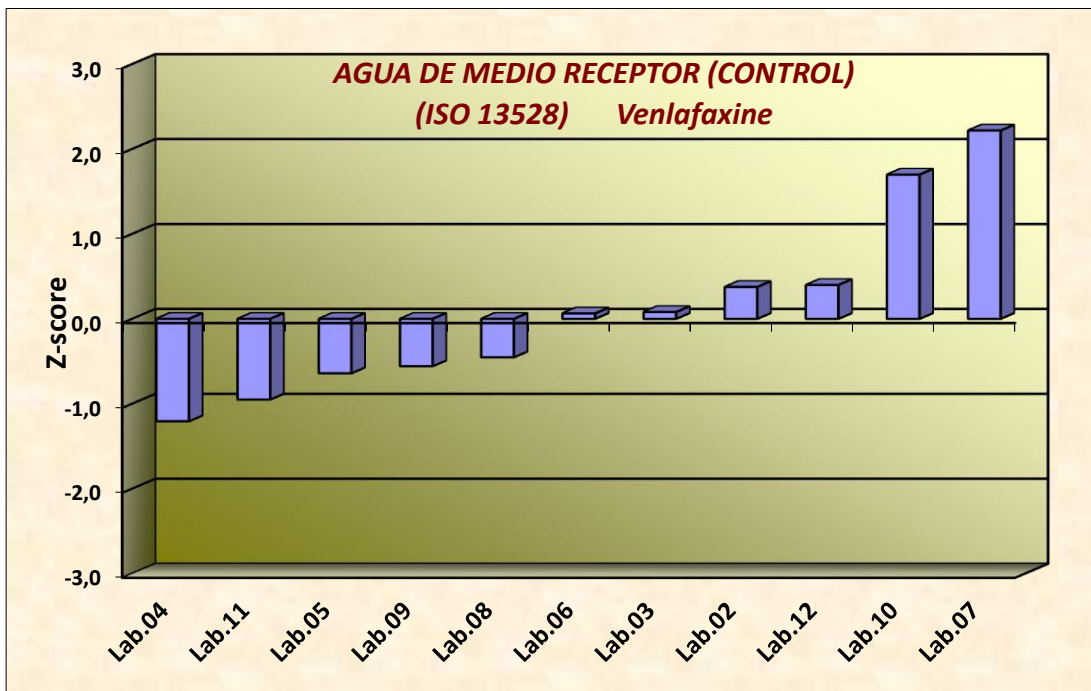
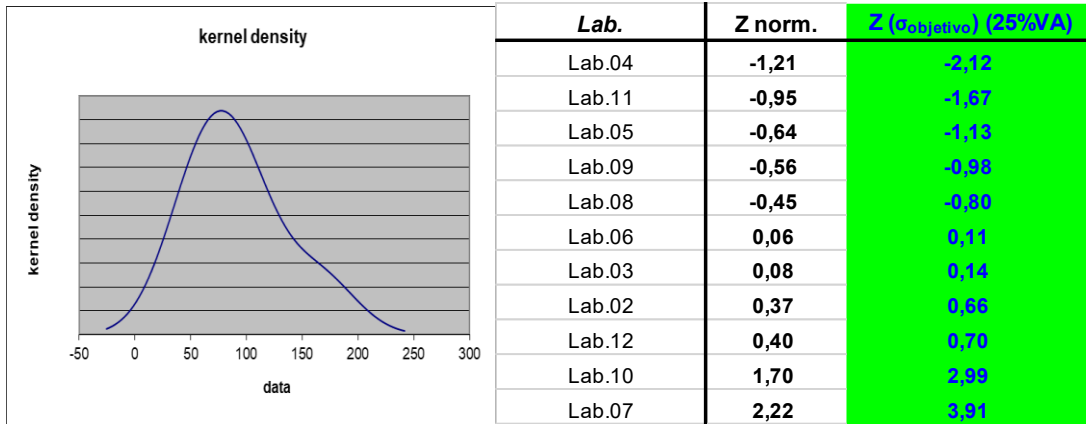


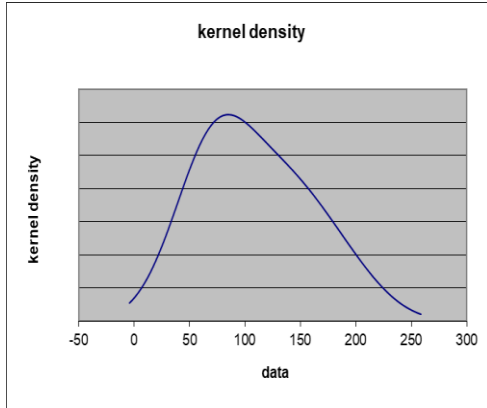
Lab.	Z norm.	Z (σ _{objetivo}) (25%VA)
Lab.09	-0,92	-1,21
Lab.08	-0,84	-1,10
Lab.02	-0,68	-0,89
Lab.05	-0,61	-0,80
Lab.06	0,02	0,03
Lab.10	1,40	1,83
Lab.03	1,53	2,00
Lab.12	-	-



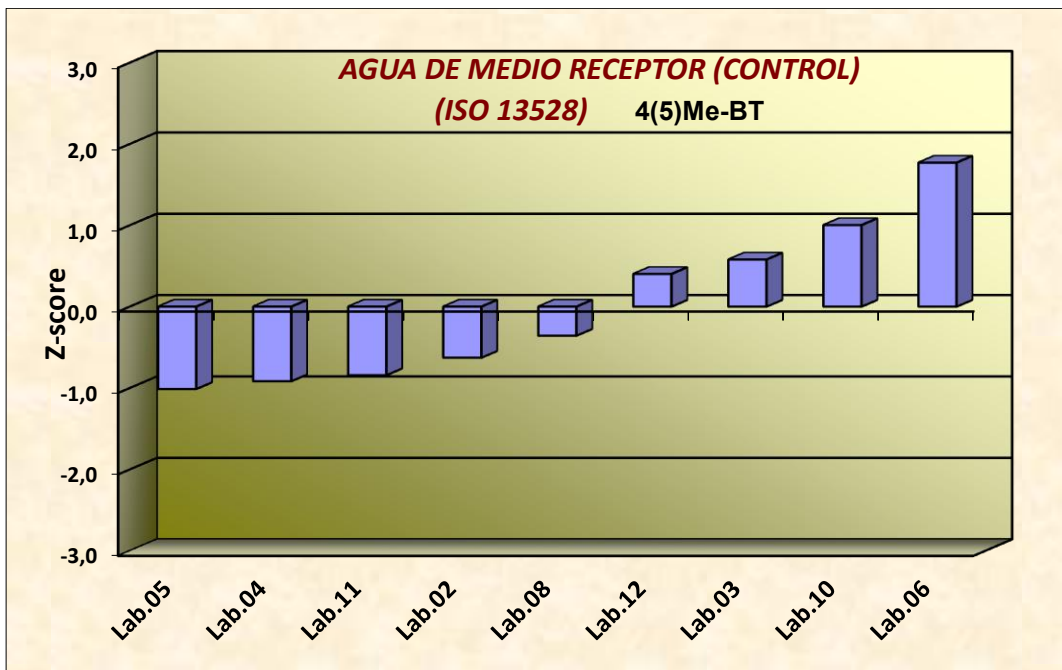






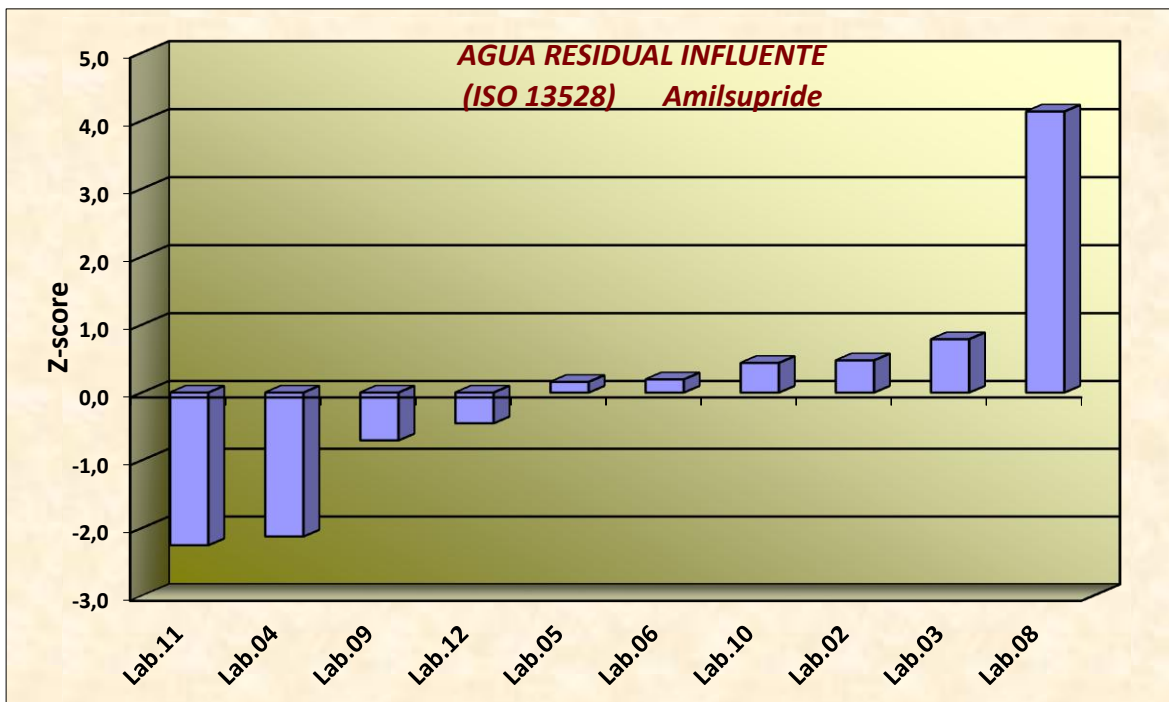
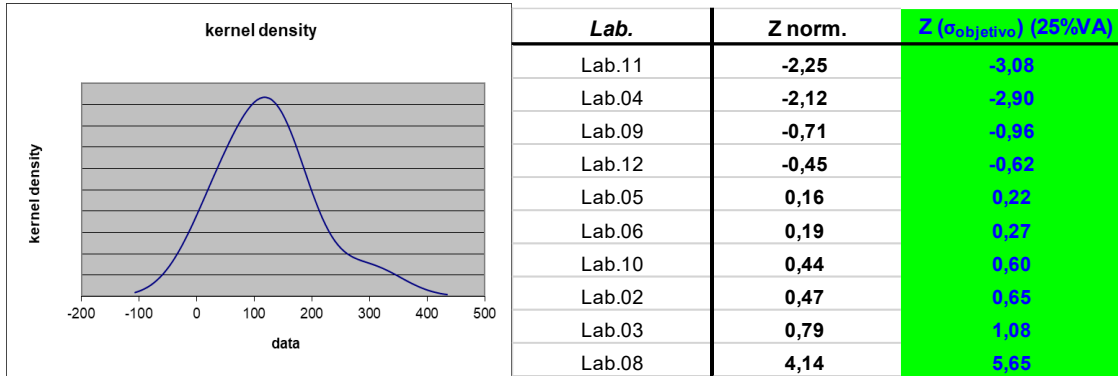


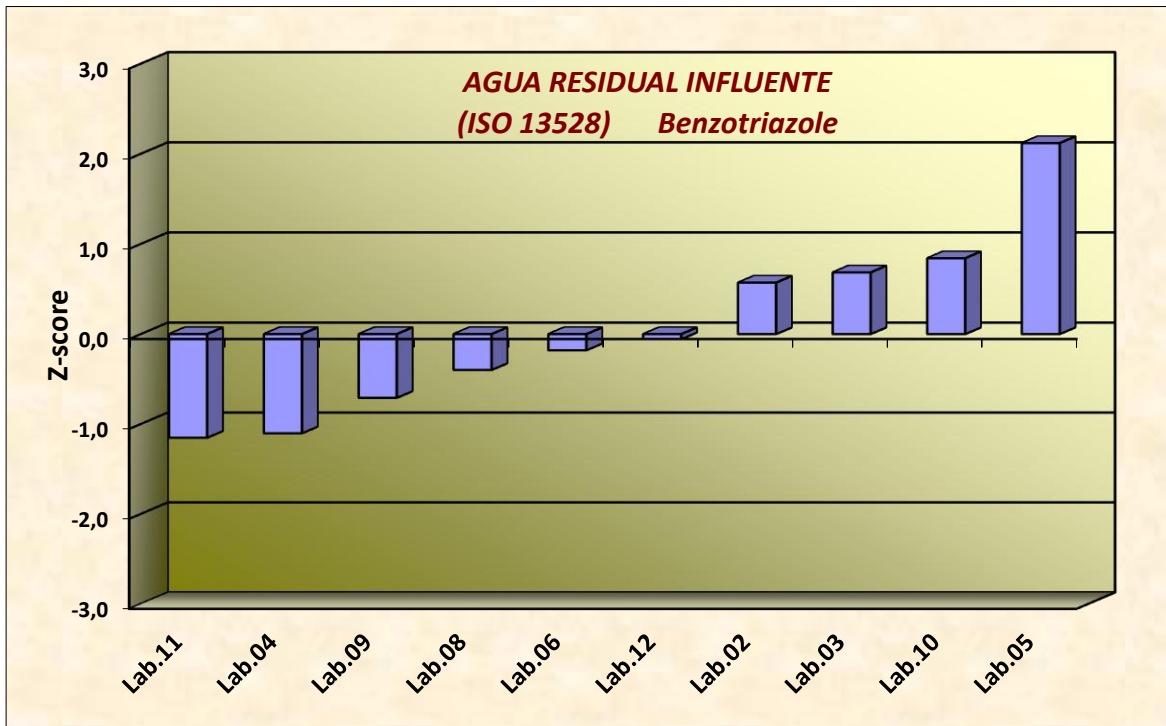
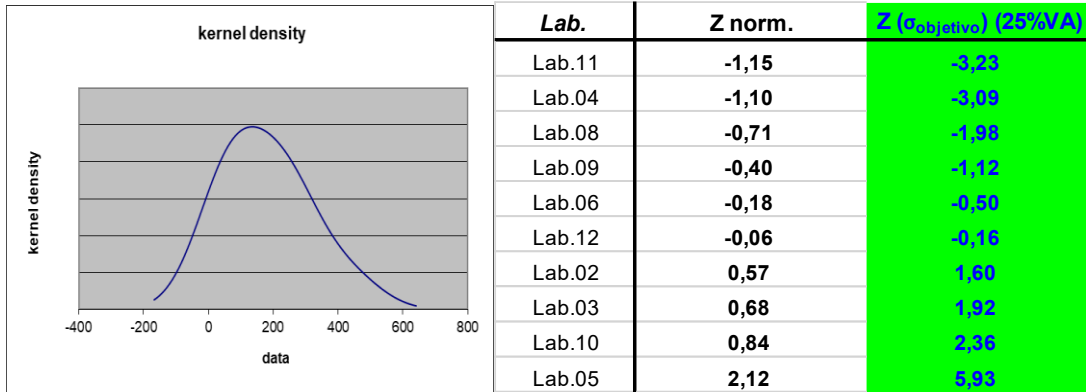
Lab.	Z norm.	Z (σ_{objetivo}) (25%VA)
Lab.05	-0,97	-1,68
Lab.04	-0,88	-1,50
Lab.11	-0,80	-1,37
Lab.02	-0,57	-0,99
Lab.08	-0,29	-0,50
Lab.12	0,49	0,85
Lab.03	0,68	1,17
Lab.10	1,12	1,93
Lab.06	1,92	3,30

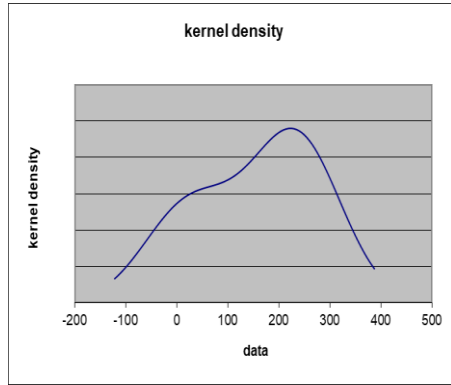




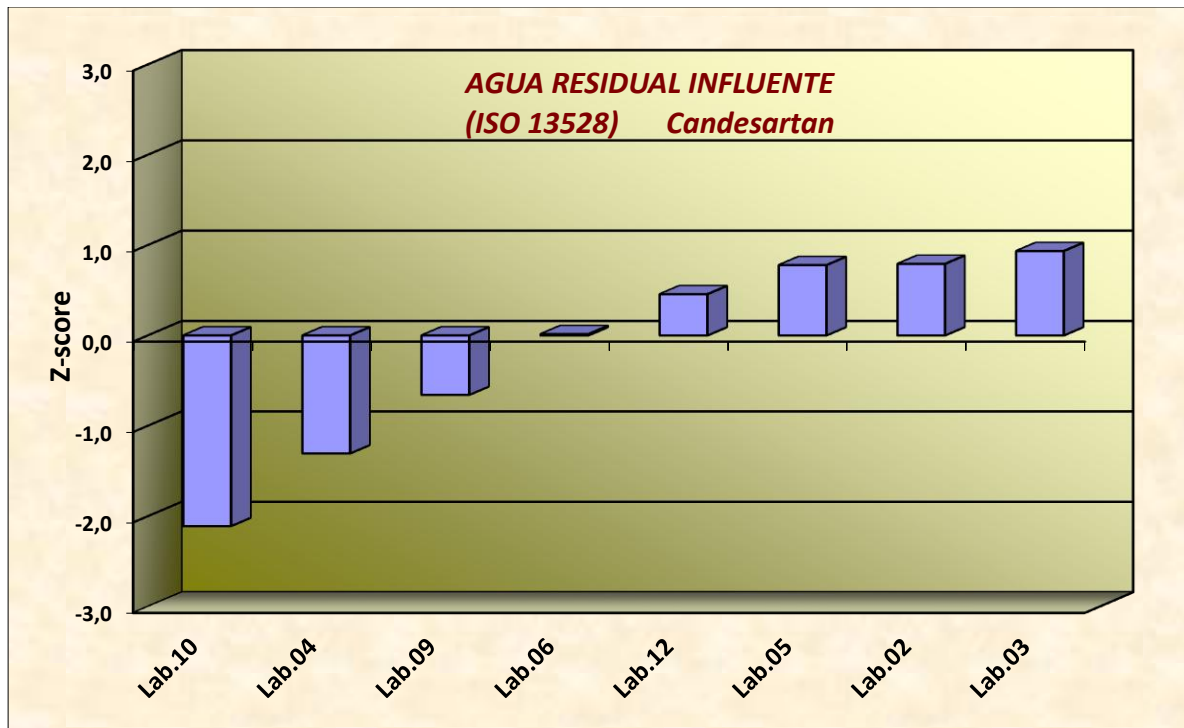
B.- MATRIZ AGUA RESIDUAL INFLUENTE

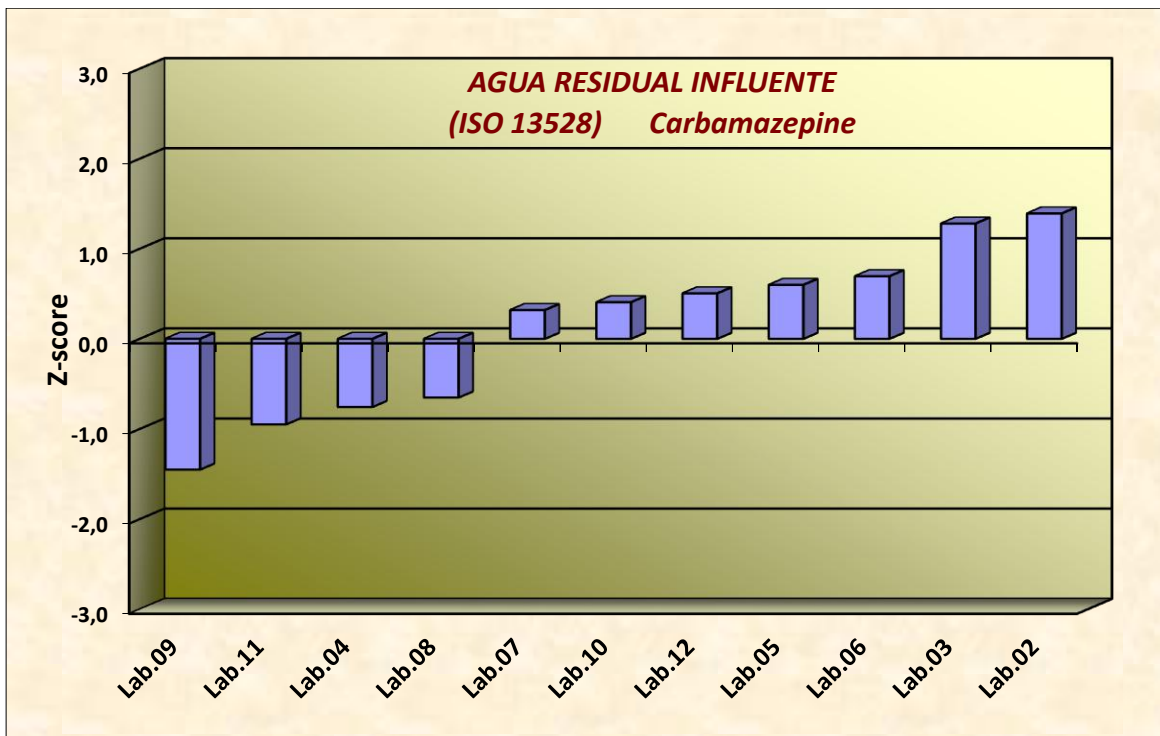
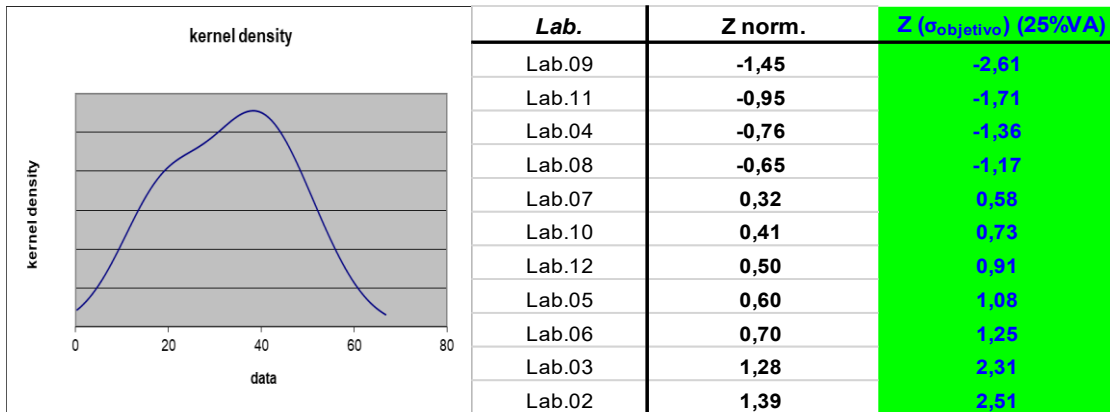


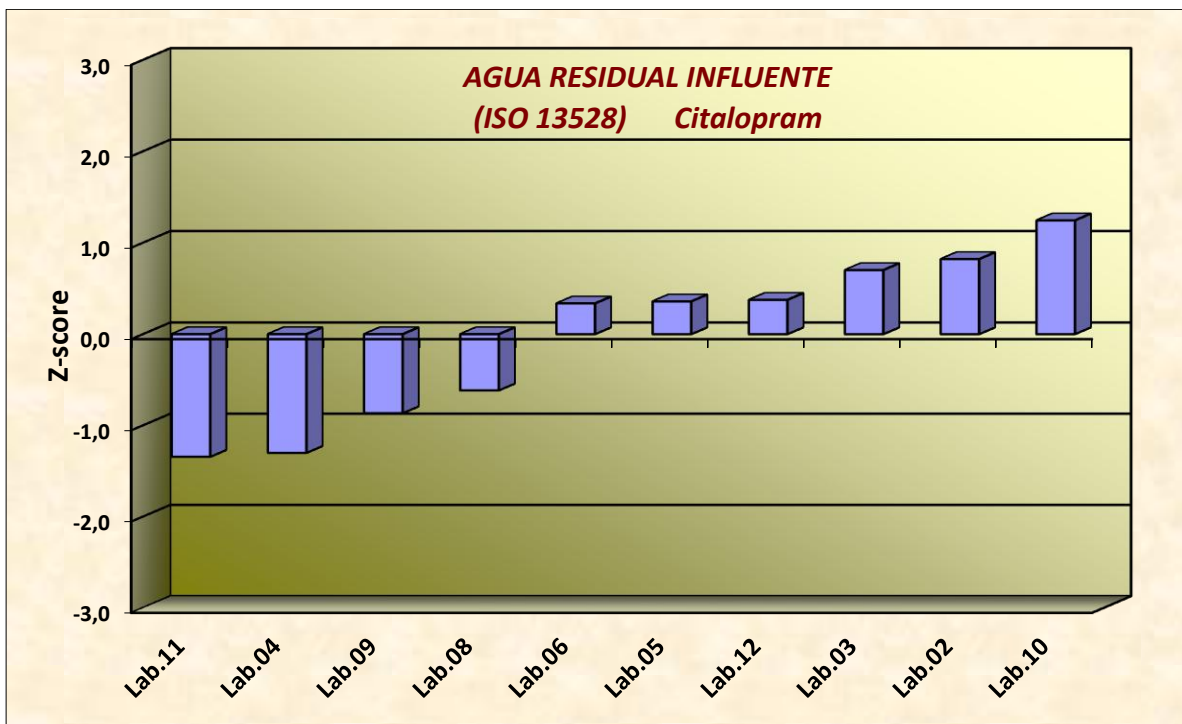
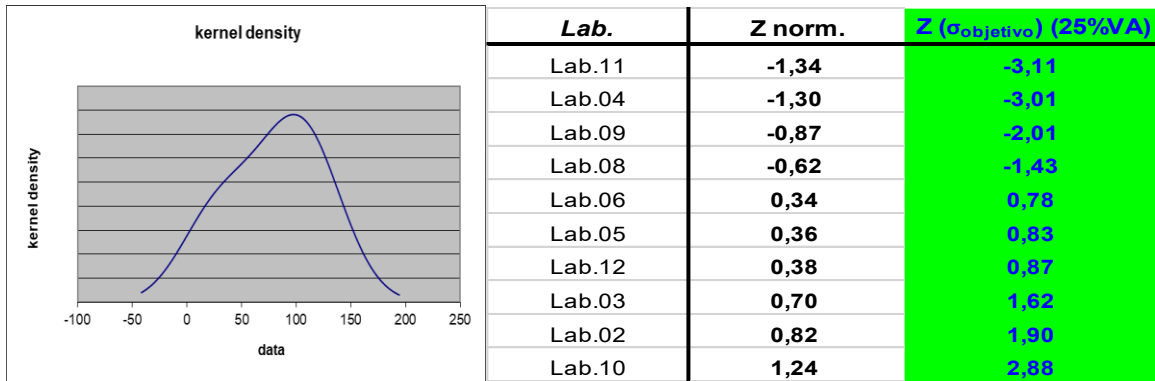


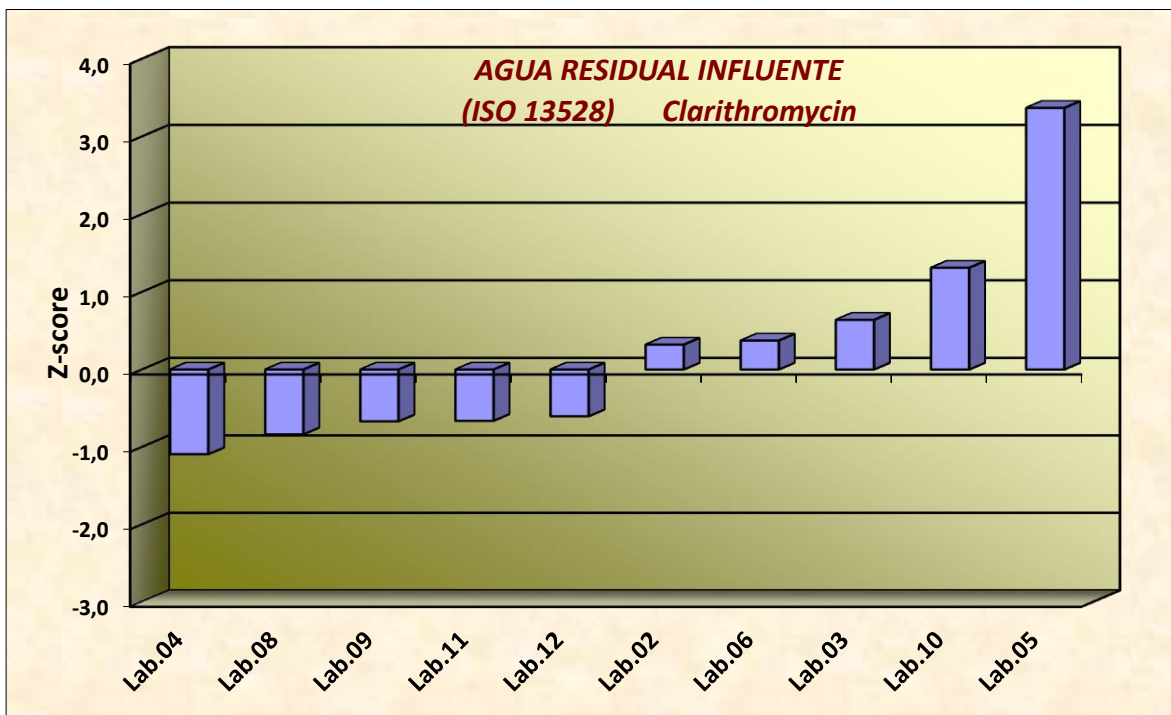
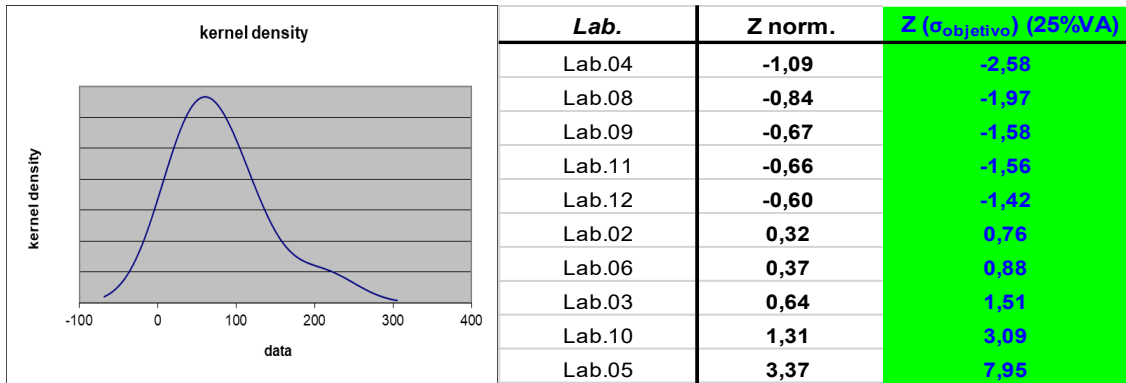


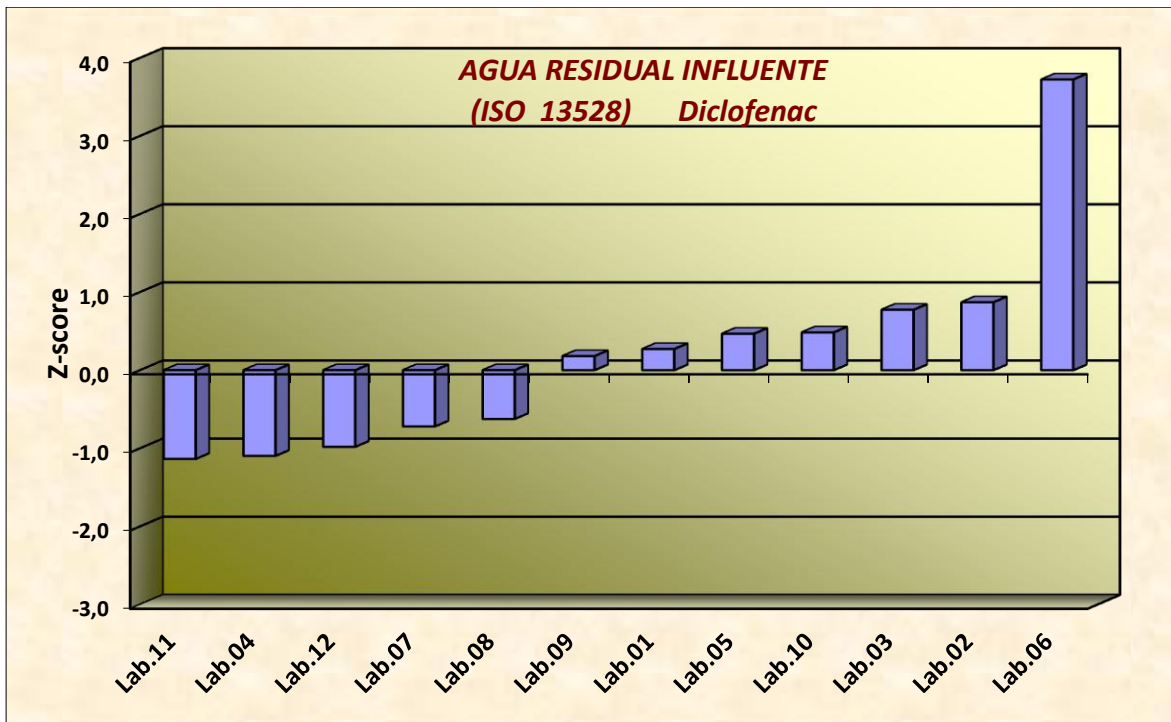
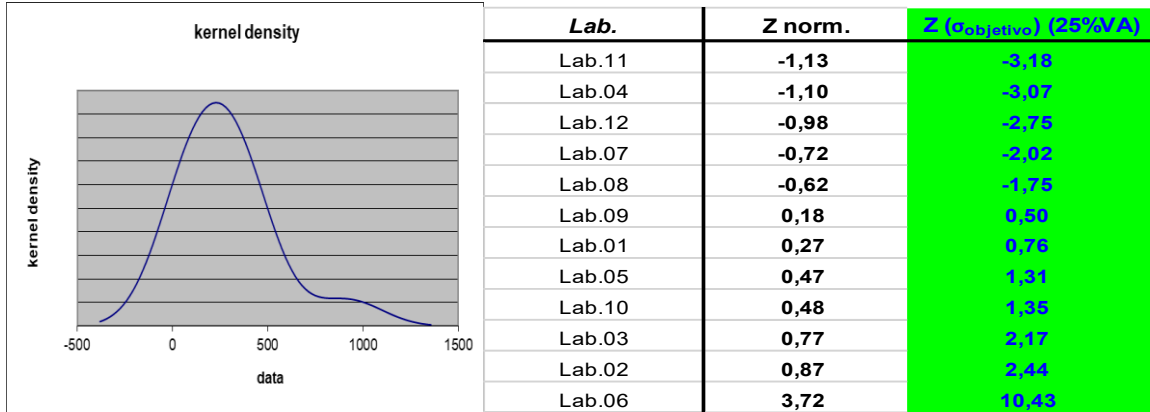
Lab.	Z norm.	Z (σ_{objetivo}) (25%VA)
Lab.10	-2,11	-3,89
Lab.04	-1,30	-2,41
Lab.09	-0,66	-1,21
Lab.06	0,01	0,03
Lab.12	0,46	0,84
Lab.05	0,78	1,44
Lab.02	0,79	1,46
Lab.03	0,93	1,72

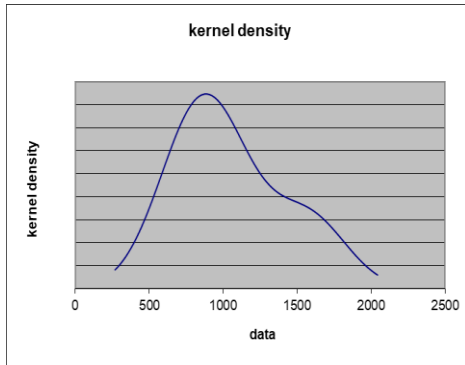




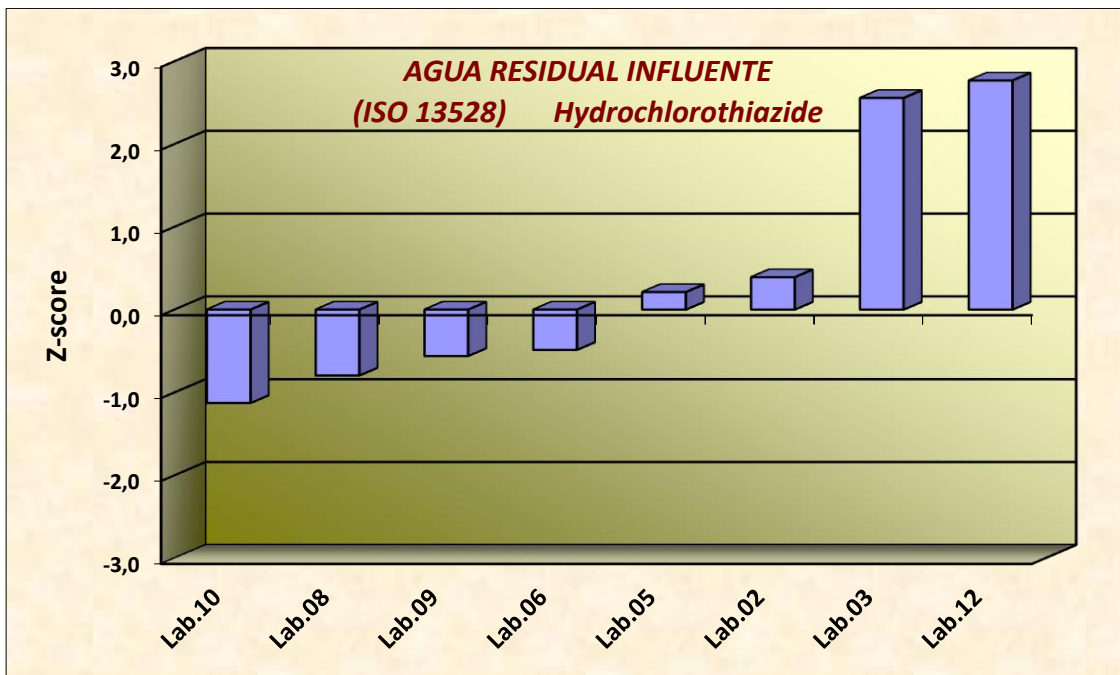


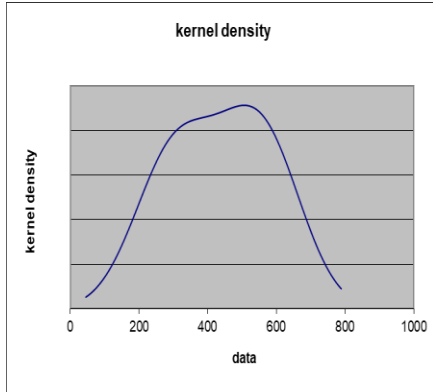




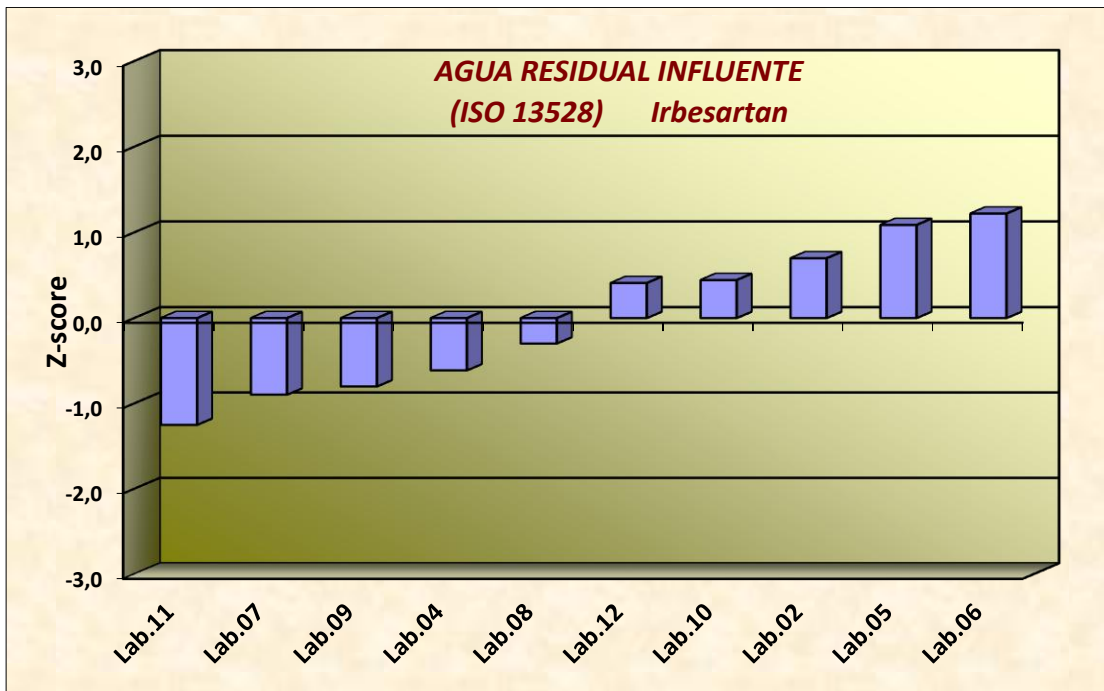


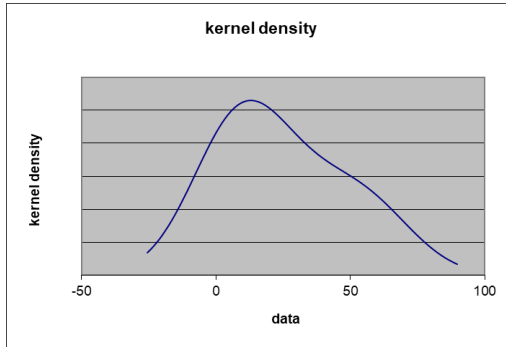
Lab.	Z norm.	Z (σ_{objetivo}) (25%VA)
Lab.10	-1,13	-1,06
Lab.08	-0,80	-0,75
Lab.09	-0,56	-0,53
Lab.06	-0,49	-0,46
Lab.05	0,21	0,20
Lab.02	0,39	0,37
Lab.03	2,56	2,40
Lab.12	2,77	2,60



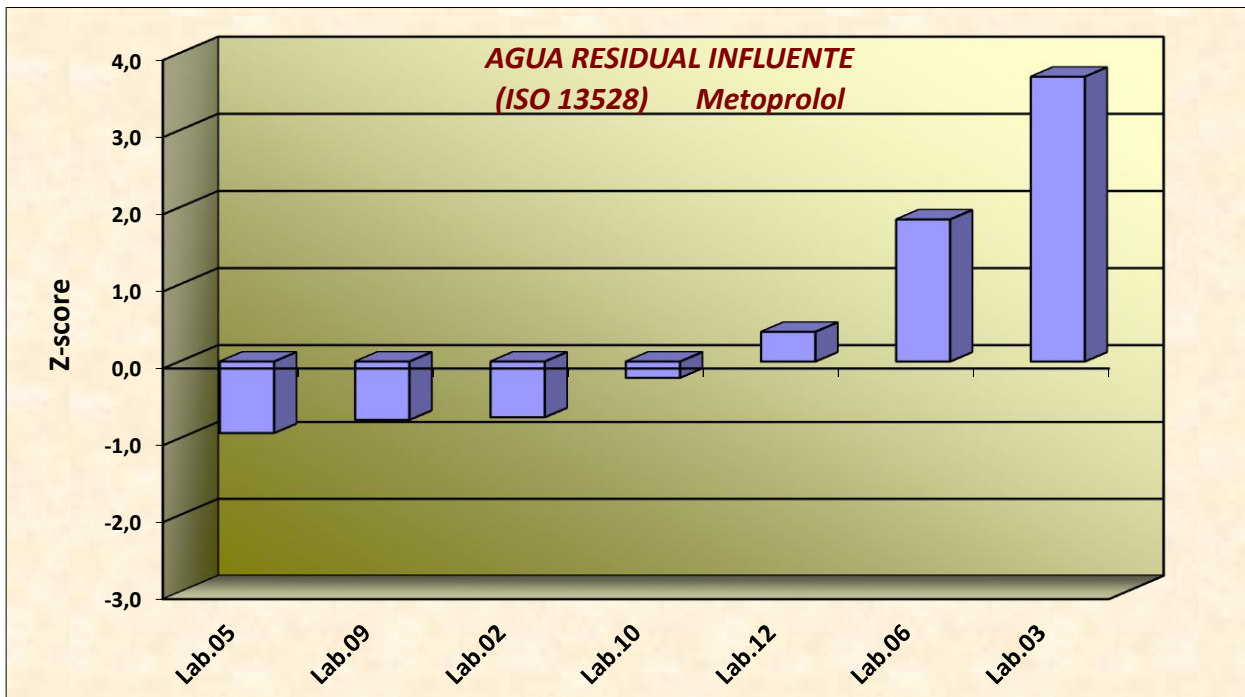


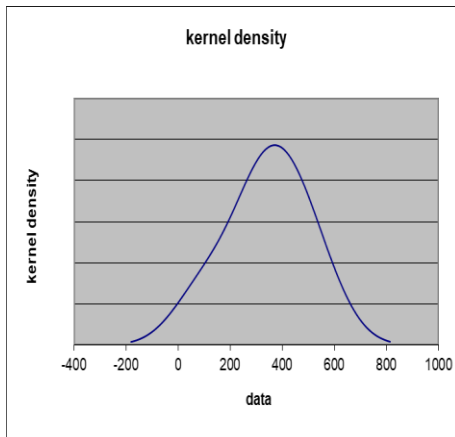
Lab.	Z norm.	Z (σ _{objetivo}) (25%VA)
Lab.11	-1,36	-1,88
Lab.07	-1,01	-1,39
Lab.09	-0,91	-1,26
Lab.04	-0,72	-1,00
Lab.08	-0,41	-0,57
Lab.12	0,30	0,41
Lab.10	0,33	0,46
Lab.02	0,59	0,81
Lab.05	0,98	1,35
Lab.06	1,11	1,53
Lab.03	1,11	1,54



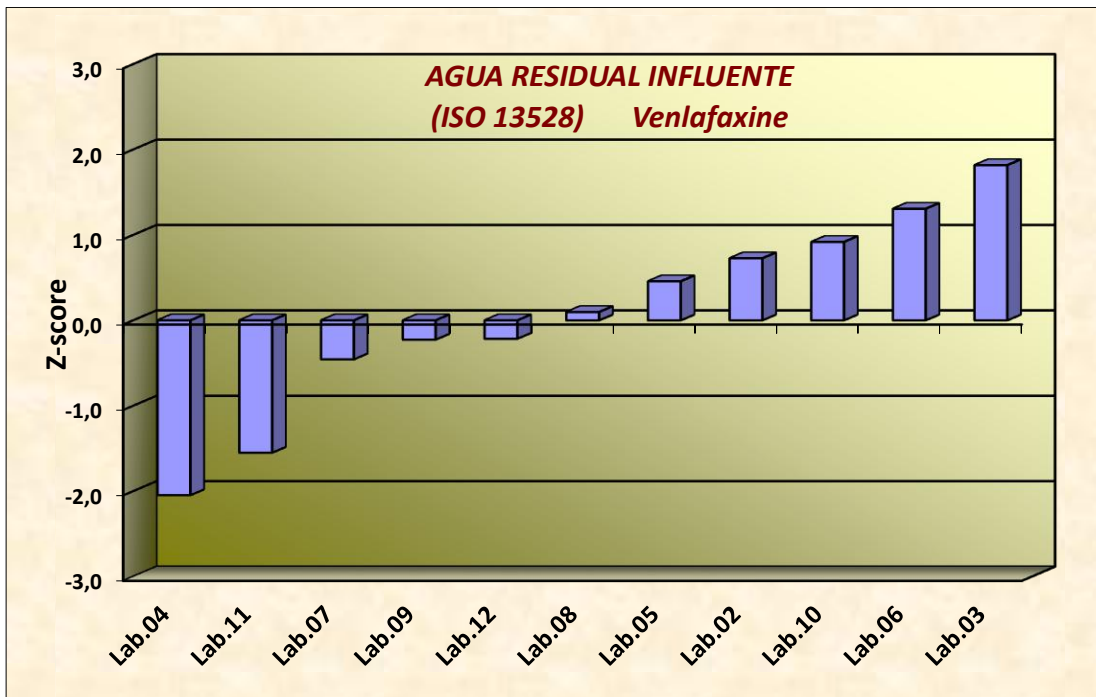


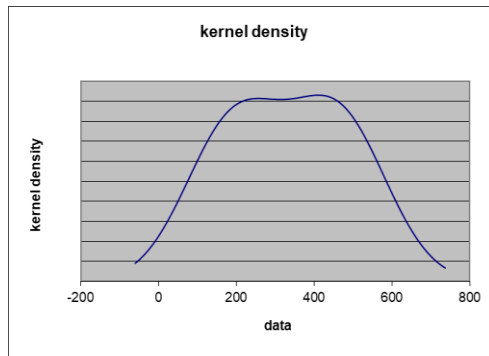
Lab.	Z norm.	Z (σ _{objetivo}) (25%VA)
Lab.05	-0,93	-2,91
Lab.09	-0,76	-2,37
Lab.02	-0,72	-2,27
Lab.10	-0,21	-0,66
Lab.12	0,39	1,21
Lab.06	1,84	5,77
Lab.03	3,69	11,55



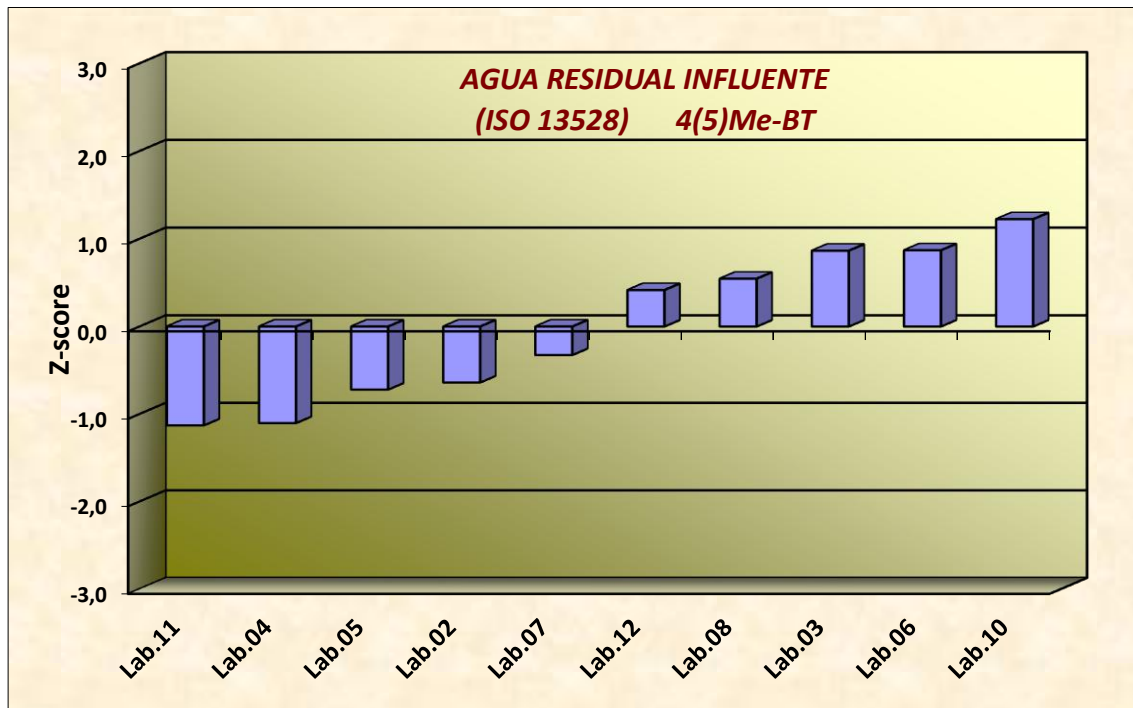


Lab.	Z norm.	Z (σ_{objetivo}) (25%VA)
Lab.04	-2,05	-3,20
Lab.11	-1,55	-2,42
Lab.07	-0,46	-0,71
Lab.09	-0,23	-0,35
Lab.12	-0,22	-0,34
Lab.08	0,10	0,15
Lab.05	0,46	0,71
Lab.02	0,73	1,14
Lab.10	0,92	1,43
Lab.06	1,31	2,04
Lab.03	1,82	2,84



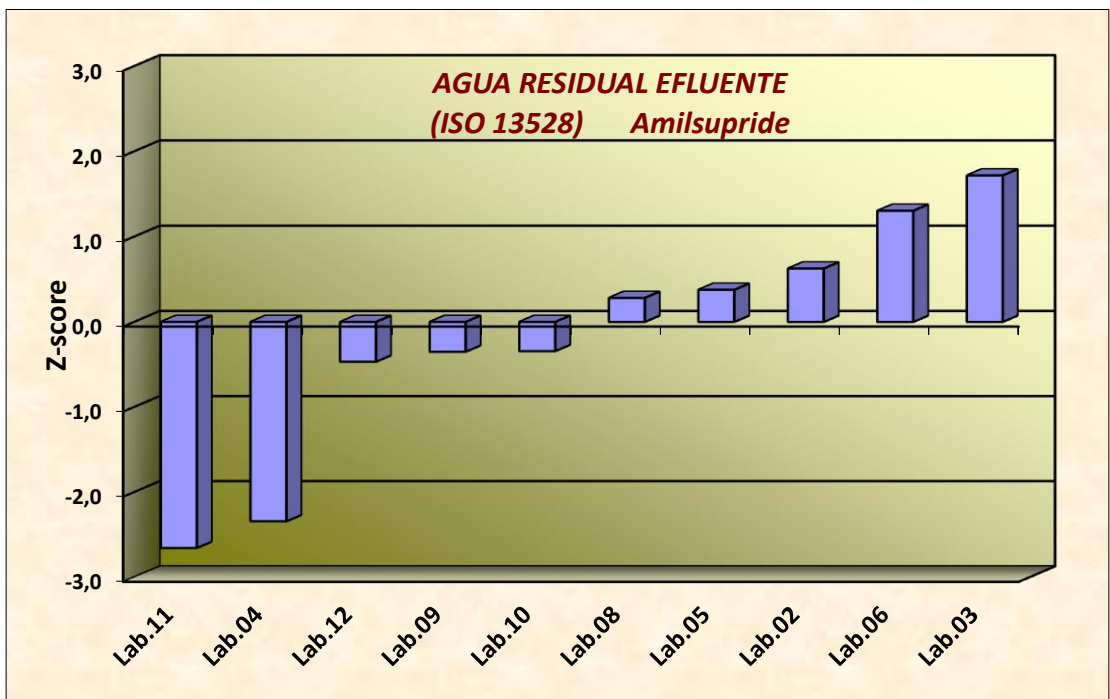
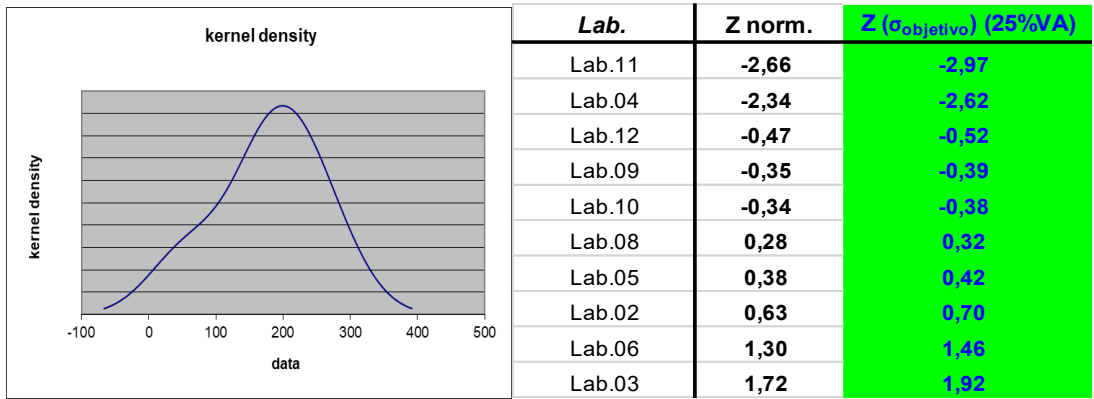


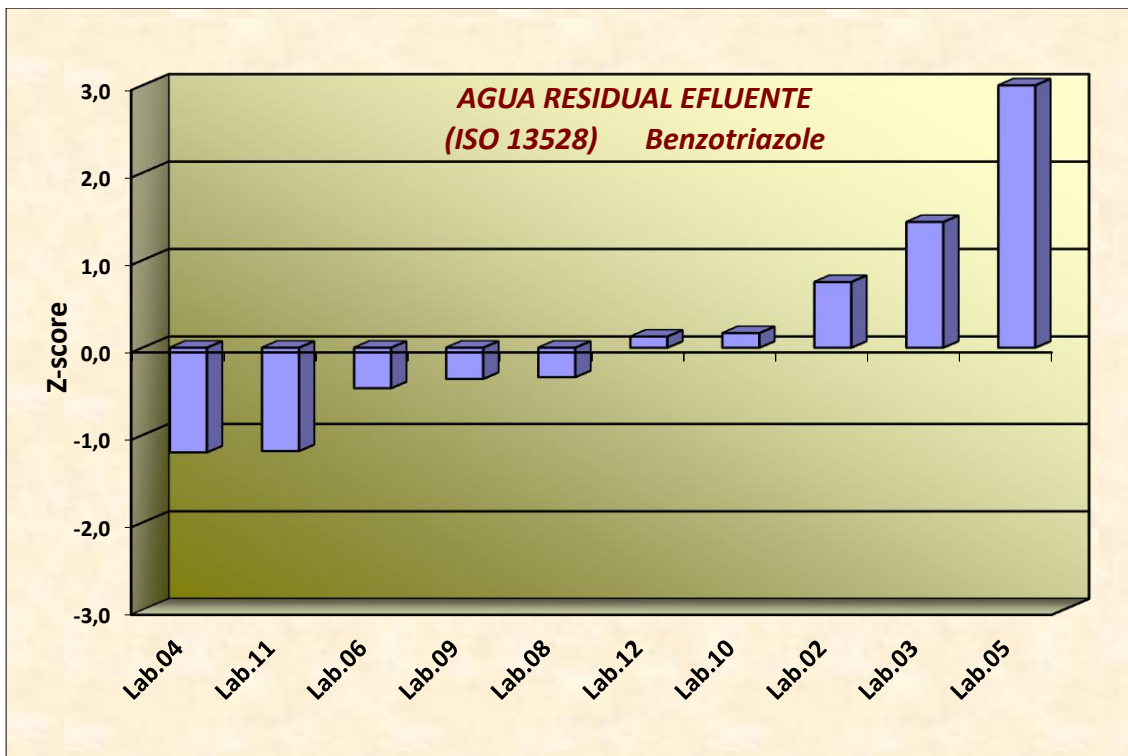
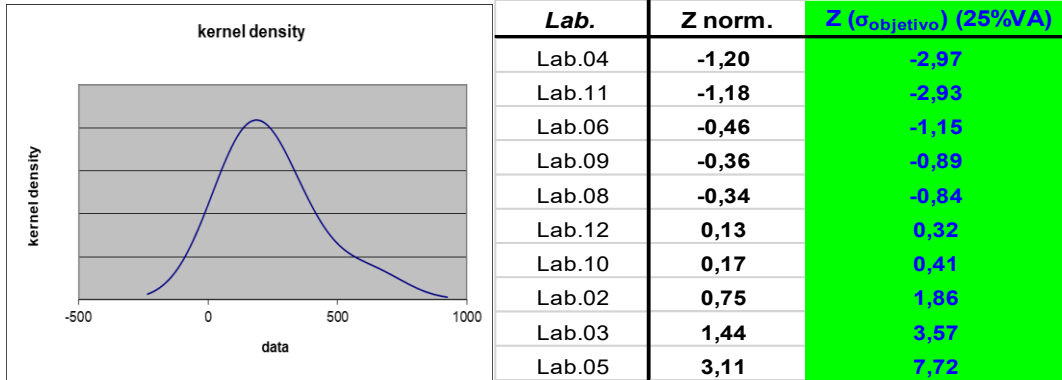
Lab.	Z norm.	Z (objetivo) (25%VA)
Lab.11	-1,13	-2,32
Lab.04	-1,10	-2,26
Lab.05	-0,72	-1,48
Lab.02	-0,64	-1,31
Lab.07	-0,33	-0,67
Lab.12	0,42	0,85
Lab.08	0,55	1,12
Lab.03	0,86	1,77
Lab.06	0,87	1,78
Lab.10	1,22	2,51

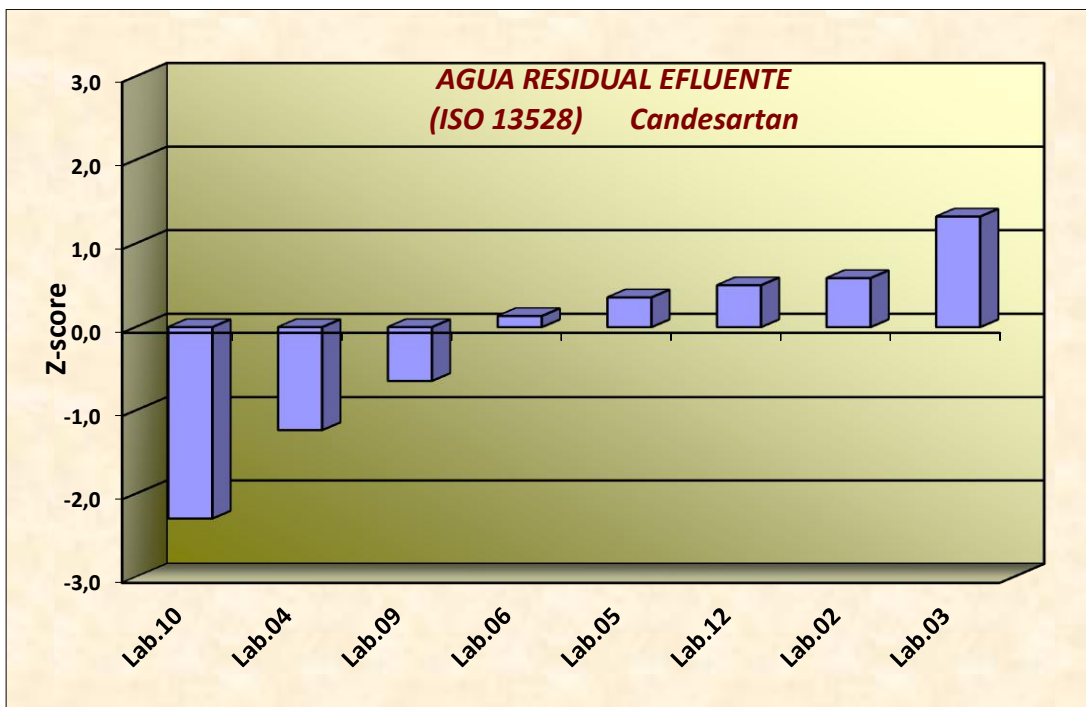
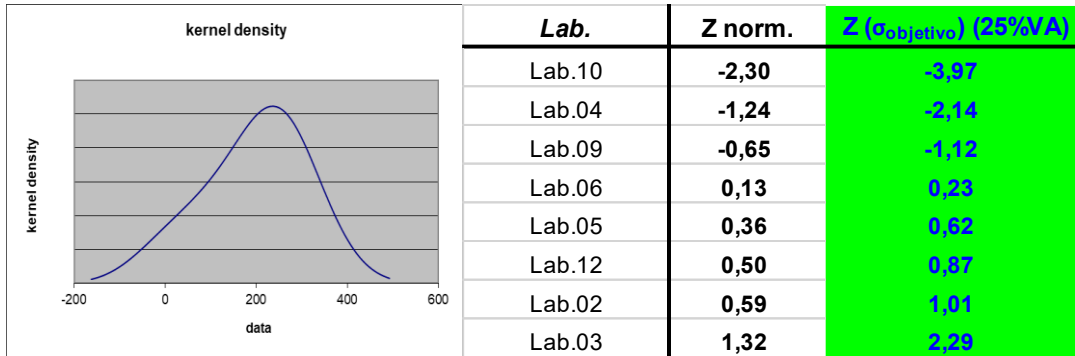


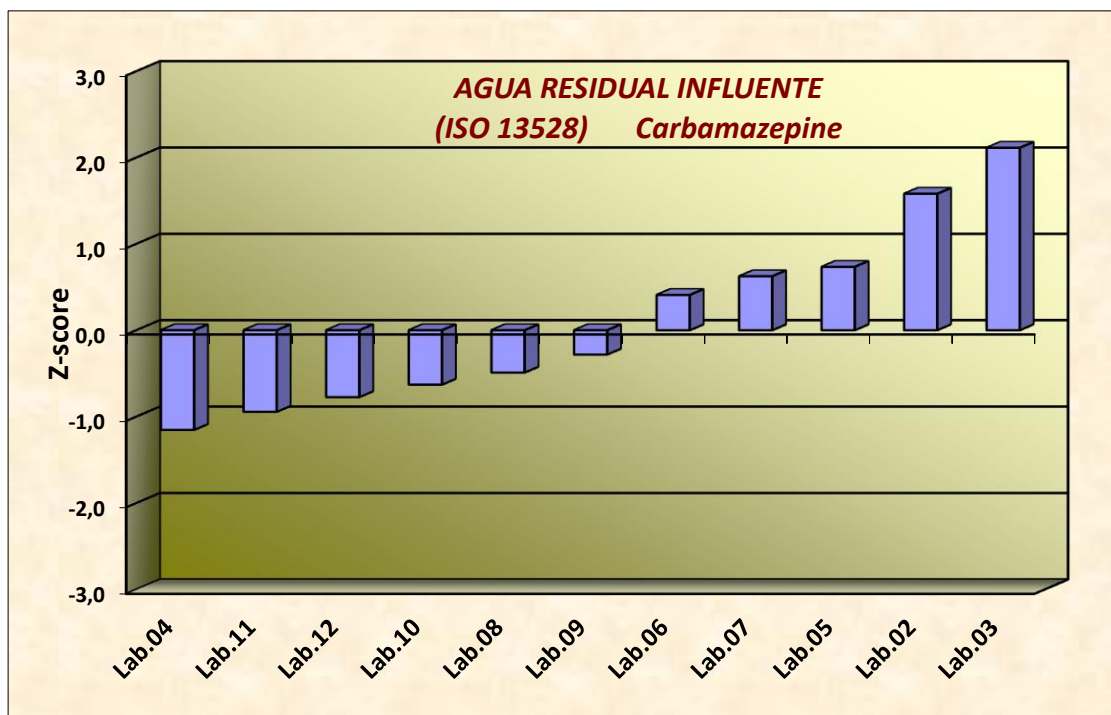
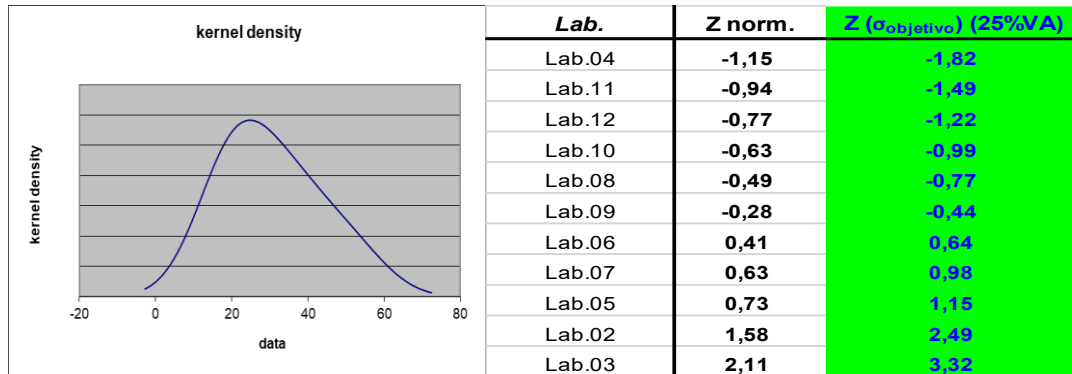


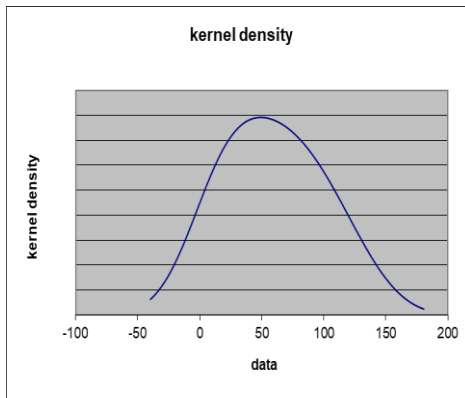
C.- MATRIZ AGUA RESIDUAL EFLUENTE



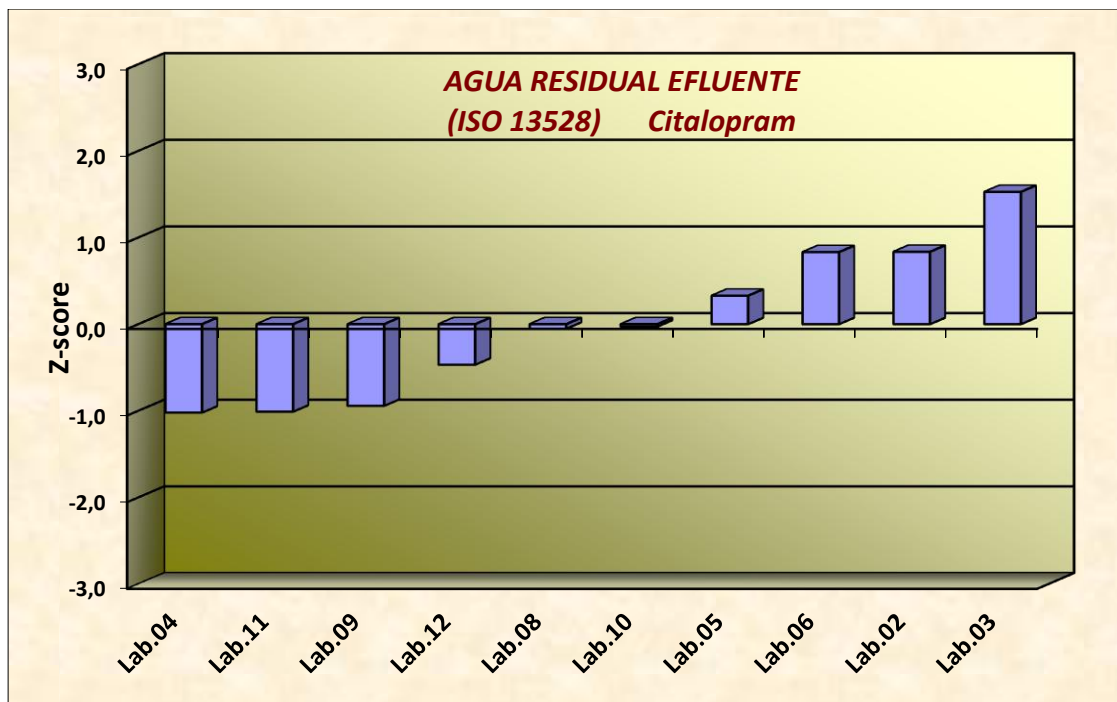


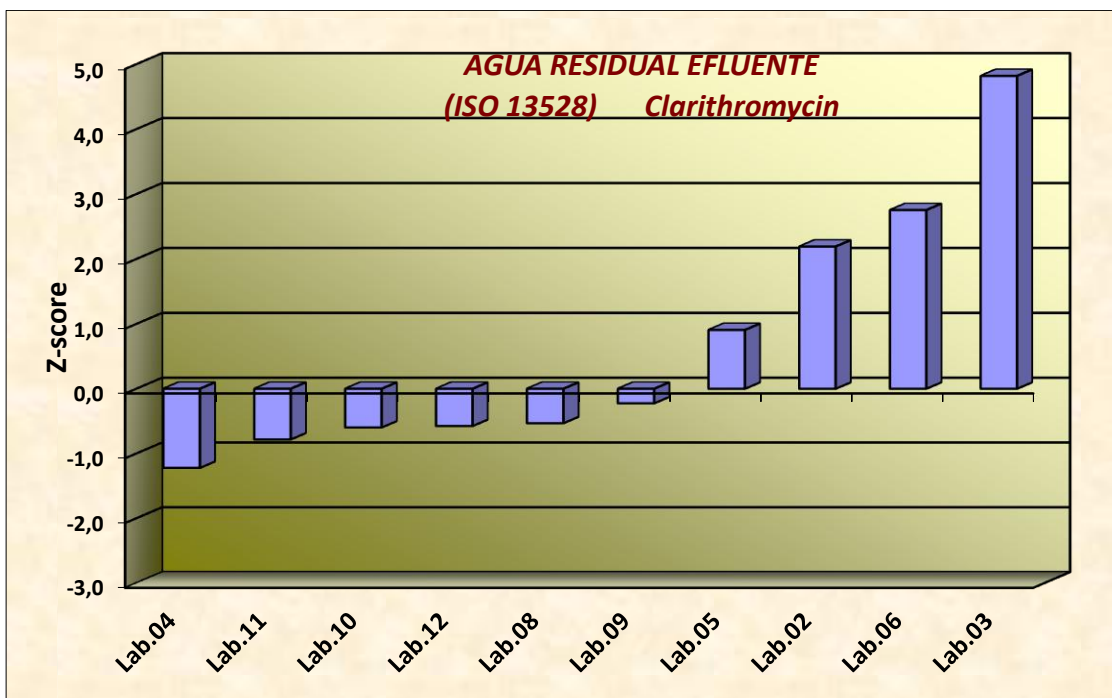
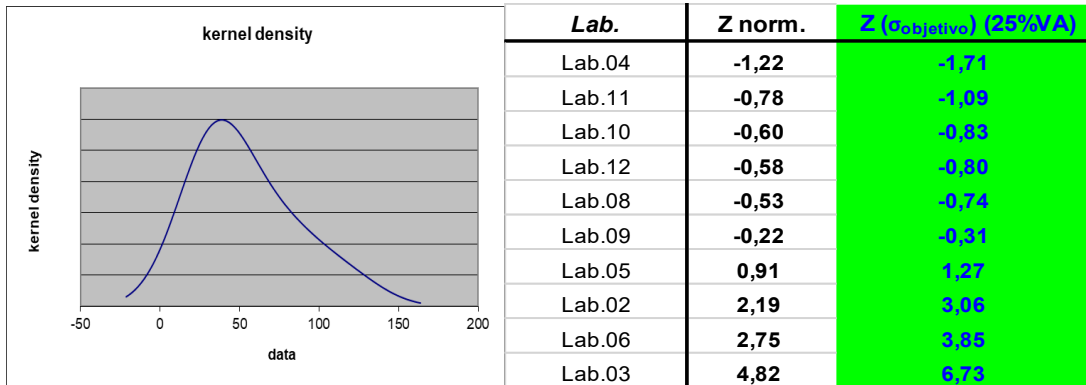


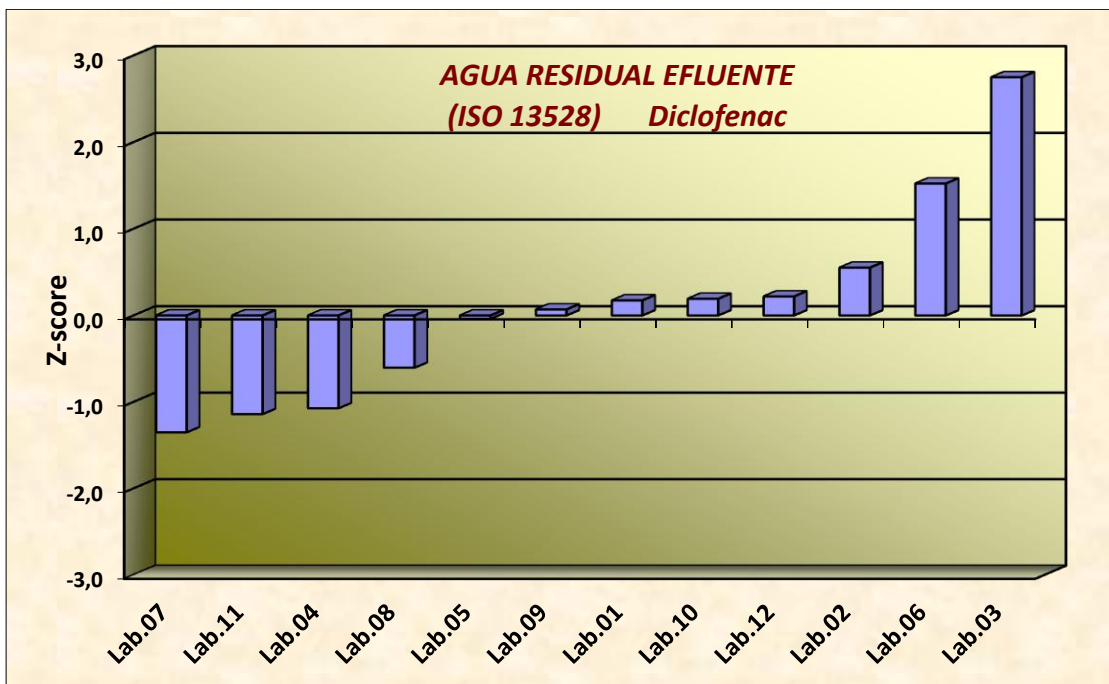
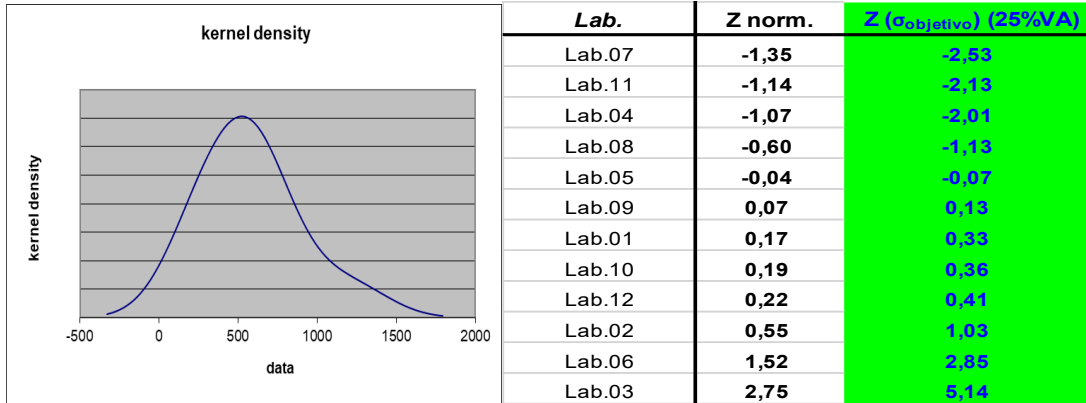


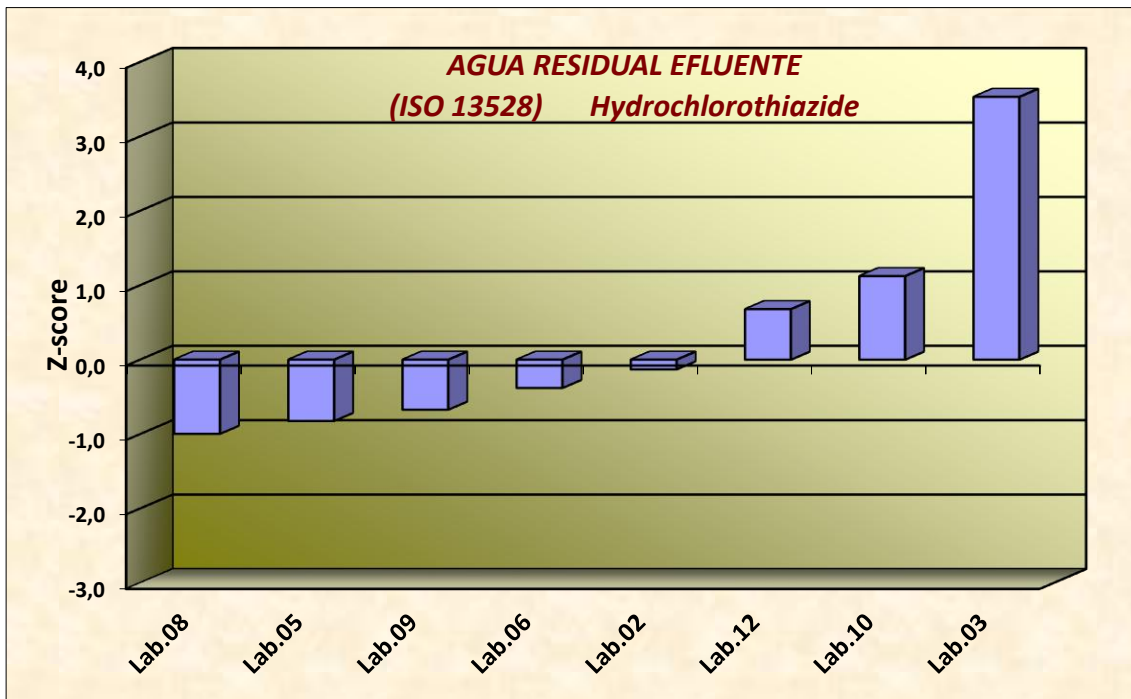
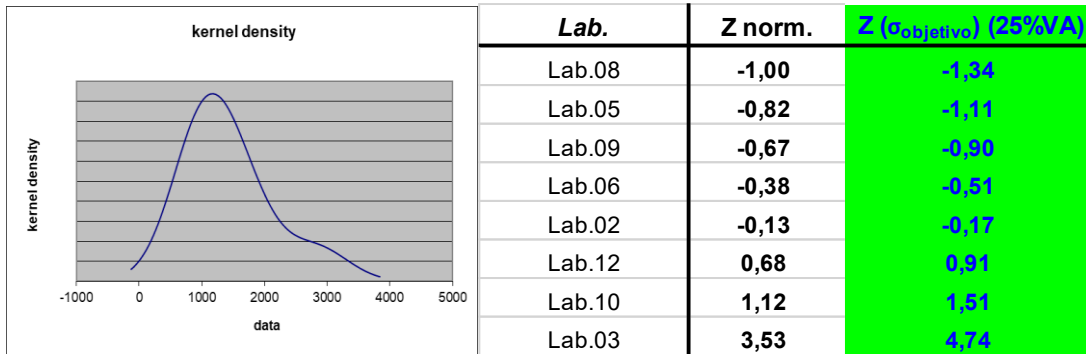


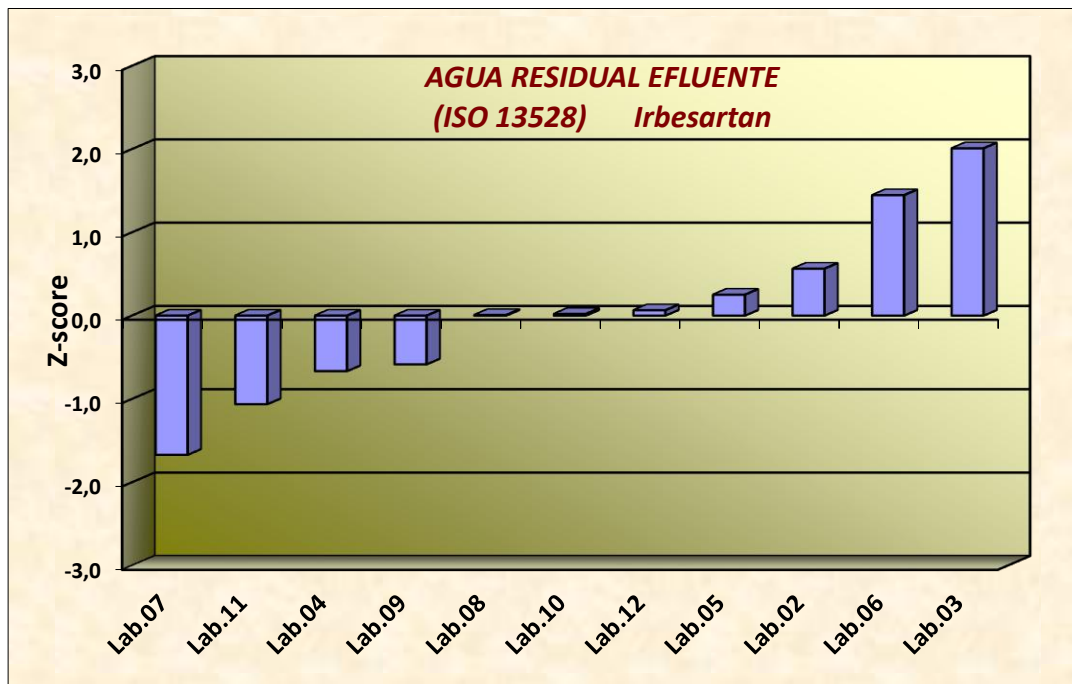
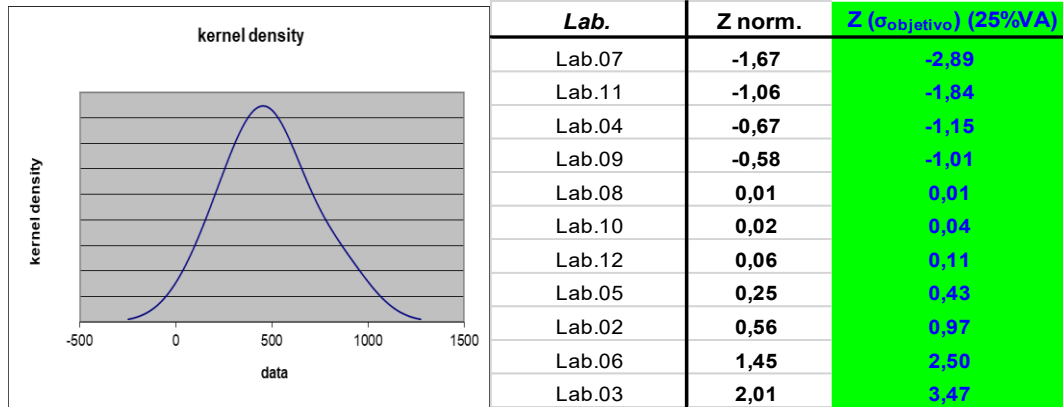
Lab.	Z norm.	Z (σ_{objetivo}) (25%VA)
Lab.04	-1,02	-2,97
Lab.11	-1,01	-2,94
Lab.09	-0,94	-2,74
Lab.12	-0,47	-1,36
Lab.08	-0,05	-0,15
Lab.10	-0,03	-0,08
Lab.05	0,33	0,95
Lab.06	0,83	2,42
Lab.02	0,84	2,43
Lab.03	1,53	4,44

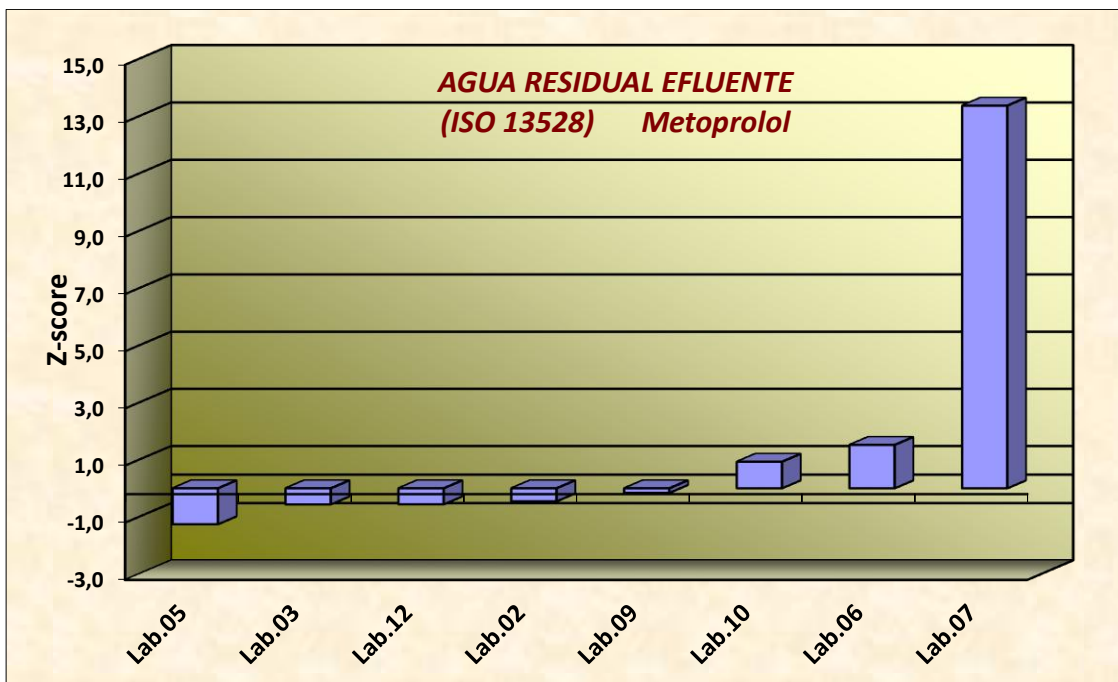
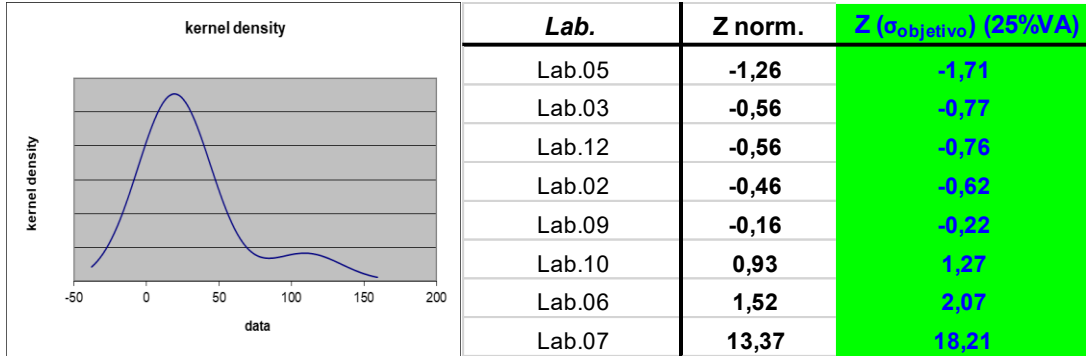


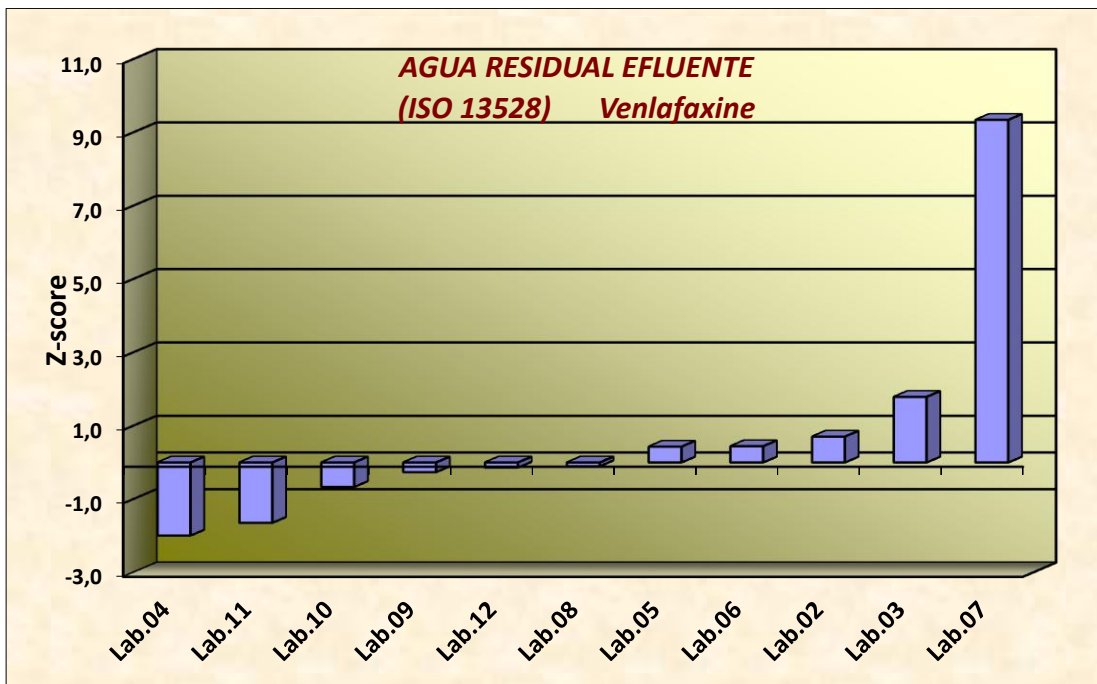
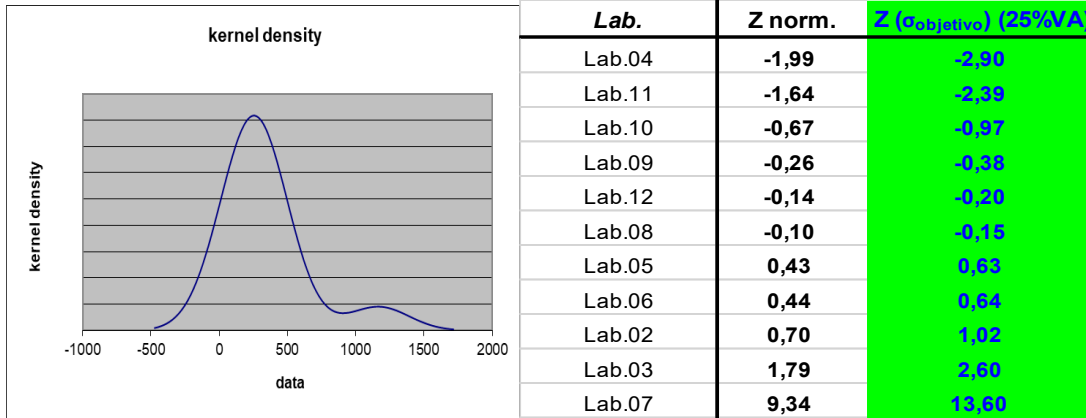


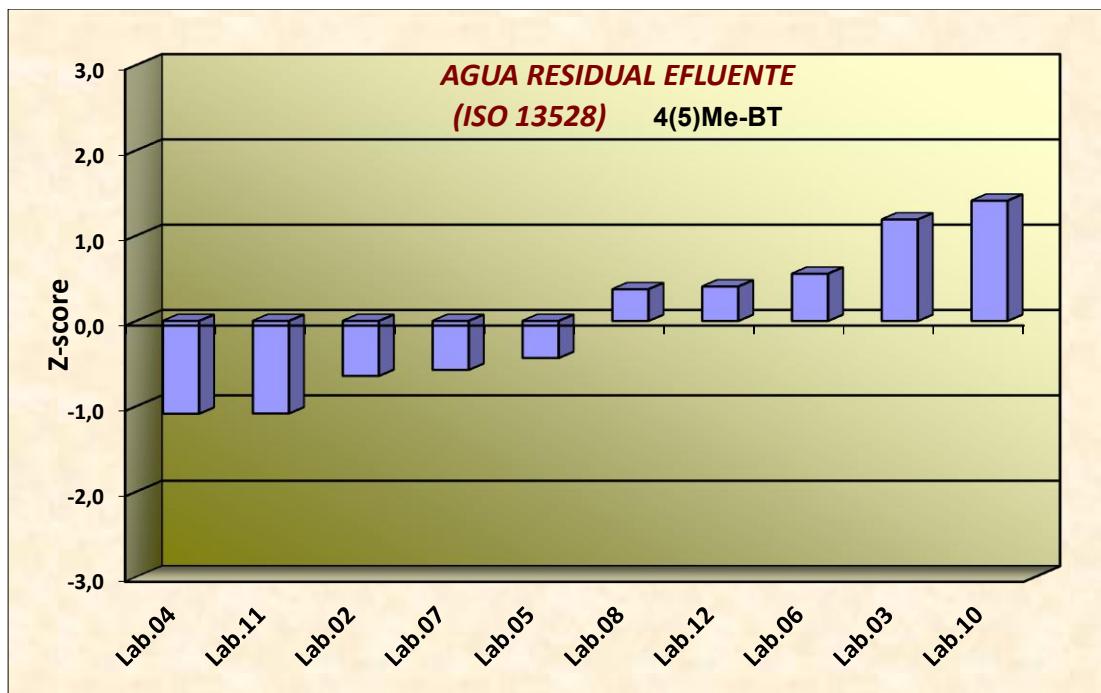
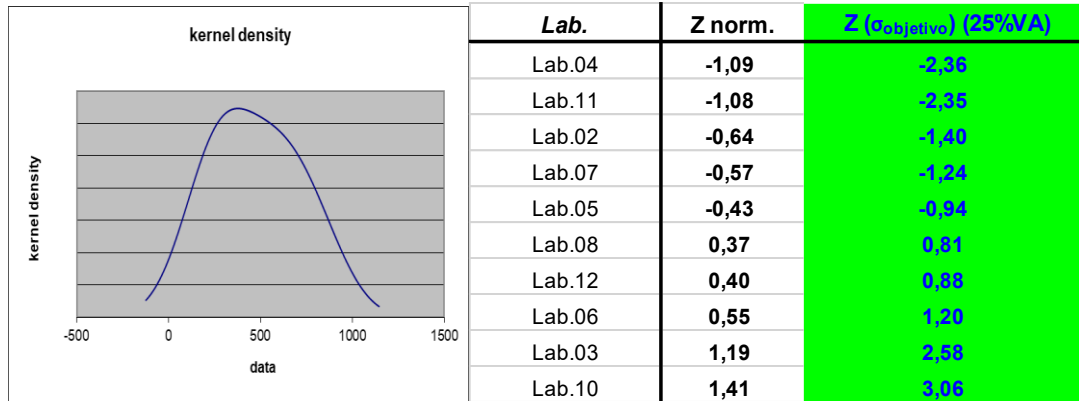














ANEXO IV. RESUMEN DATOS TÉCNICOS DE ENSAYO POR LABORATORIO

Participante	Pretratamiento	Extracción
Lab.01	Sin pretratamiento	Inyección directa
Lab.02	Filtración	Inyección directa
Lab.03	Filtración	Cartuchos SPE (Oasis HLB)
Lab.04	Acidificación pH	Cartuchos SPE (Strata-X)
Lab.05	Filtración	Cartuchos SPE (Oasis HLB)
Lab.06	Sin pretratamiento	Inyección directa
Lab.07	Filtración	Cartuchos SPE (Strata-X)
Lab.08	Centrifugación	SPE online (ENV+)
Lab.09	Centrifugación	Inyección directa
Lab.10	Filtración	Cartuchos SPE (Oasis HLB)
Lab.11	Acidificación pH	Cartuchos SPE (Strata-X)
Lab.12	Sin pretratamiento	Cartuchos SPE (Strata-X)

Participante	Técnica cromatográfica	Columna cromatográfica
Lab.01	LC-ESI-MS/MS (QqQ)	EclipsePlus C18, 50 x 2.1 mm, 1.8 µm
Lab.02	LC-ESI-MS/MS (QqQ)	Zorbax Eclipse XDB-C18, 100 x 2.1 mm, 3.5 µm
Lab.03	LC-ESI-MS/MS (QqQ)	Acquity HSS T3, 50 x 2.1 mm, 1.8 µm (positivo), Acquity BEH C18, 50 x 2.1 mm, 1.7 µm (negativo)
Lab.04	LC-ESI-Q-TOF	Luna Omega PS C18, 50 x 2.1 mm, 1.6 µm
Lab.05	LC-ESI-MS/MS (QqQ)	C18, 150 x 2.1 mm, 1.5 µm
Lab.06	LC-ESI-MS/MS (QqQ)	Waters BEH C18, 75 mm x 2.1, 1.7 µm
Lab.07	LC-ESI-MS/MS (QqQ)	C18, 50 x 2.1 mm, 2.6 µm
Lab.08	LC-ESI-Q-TOF	Bruker Intensity Solo C18-2, 100 x 2,1 mm, 1.8 µm
Lab.09	LC-ESI-MS/MS (QqQ)	Acquity BEH C18, 100 x 2,1 mm, 1.7 µm
Lab.10	LC-ESI-MS/MS (QqQ)	ACE UltraCore 5µm SuperC18 (30 x 21 mm)
Lab.11	LC-ESI-Q-TOF	C18, 100 x 2.1 mm
Lab.12	LC-ESI-MS/MS (QqQ)	CORTECS C18, 100 x 2.1 mm, 1.6 µm



Participante	Método de cálculo	Regresión
Lab.01	Cálculos basados en patrón interno	Regresión lineal ponderada con factor 1/x
Lab.02	Cálculos basados en patrón interno	Regresión lineal por mínimos cuadrados ordinarios
Lab.03	Cálculos basados en patrón interno	Regresión lineal por mínimos cuadrados ordinarios
Lab.04	Cálculos basados sólo en patrón externo	Regresión lineal por mínimos cuadrados ordinarios
Lab.05	Cálculos basados en adiciones estándar	Regresión lineal ponderada con factor 1/x
Lab.06	Cálculos basados sólo en patrón externo	Regresión lineal por mínimos cuadrados ordinarios
Lab.07	Cálculos basados en adiciones estándar	Regresión lineal por mínimos cuadrados ordinarios
Lab.08	Cálculos basados en dilución isotópica	Regresión lineal por mínimos cuadrados ordinarios
Lab.09	Cálculos basados en patrón interno	Regresión lineal por mínimos cuadrados ordinarios
Lab.10	Cálculos basados en patrón externo y adiciones estándar	Regresión lineal por mínimos cuadrados ordinarios
Lab.11	Cálculos basados sólo en patrón externo	Regresión lineal por mínimos cuadrados ordinarios
Lab.12	Cálculos basados sólo en patrón externo	Regresión lineal por mínimos cuadrados ordinarios

