

Proyecto de caracterización de la leche cruda identificando la presencia del Glicomacropéptido de Caseína c-GMP

INFORME ANÁLISIS ESTADÍSTICO

Anexo 3 - 2020



### **Presidente**

Camilo Fernández de Soto

### **Vicepresidente de Agroindustria**

Amira del Pilar Ortiz Olaya

### **Gerente de Agroindustria**

Fabián Camilo Acosta Puentes

### **Profesional Senior**

#### **Sector Agroindustria**

Jenifer Yaneth Guzmán Gualteros

### **Operador Técnico**

Consortio Consultoría PTP (Biotrends Laboratorios S.A.S y RyG Asesorías).

Murcia F., Prieto P., Ávila O., Bernal O.,  
Murcia A., Amaya C., Morales M., Arias W., y Guzman J.

### **Edición**

Jenifer Guzmán  
Profesional Lácteo

### **Diseño**

Leonardo Pérez  
Profesional comunicaciones

Colombia Productiva  
2020

Anexo 3 – Informe Análisis Estadístico  
Caracterización muestras de leche.

Bogotá, Colombia

[www.colombiaproductiva.com](http://www.colombiaproductiva.com)

# Contenido

<b>OBJETIVO GENERAL</b> .....	<b>4</b>
<b>OBJETIVOS ESPECÍFICOS</b> .....	<b>4</b>
<b>RESULTADOS OBTENIDOS</b> .....	<b>5</b>
MUESTREO .....	6
PROCESO ANALÍTICO .....	10
<b>ANÁLISIS ESTADÍSTICO PARA ESTABLECER LAS MEDIDAS DE CENTRALIDAD Y DISPERSIÓN</b> .....	<b>11</b>
REVISIÓN VISUAL DE DATOS .....	11
DETERMINACIÓN DE VALORES ATÍPICOS .....	13
REVISIÓN DE LA DISTRIBUCIÓN DE DATOS .....	14
<b>COMPARACIÓN DE RESULTADOS</b> .....	<b>17</b>
COMPARACIÓN DE LOS RESULTADOS OBTENIDOS ENTRE DEPARTAMENTOS .....	21
<b>CONCLUSIONES</b> .....	<b>23</b>

## Objetivo general

Establecer a través de un estudio estadístico, un valor de referencia de la concentración basal de la molécula glicomacropéptido de caseína (c-GMP), presente en muestras de leche cruda analizadas por

cromatografía líquida de alta eficiencia (HPLC); las cuales fueron obtenidas de 30 unidades productivas rurales, vinculadas a empresas con el CIU 1040, ubicados en 4 departamentos en Colombia.

## Objetivos específicos

Describir la metodología estadística empleada en el análisis de los resultados de ensayo para la determinación de glicomacropéptido de caseína (c-GMP) en las muestras de estudio.

Describir los criterios seleccionados para la determinación del número de muestras a tomar en cada una de las unidades productivas rurales vinculadas al estudio.

Detectar datos anómalos dentro de la muestra estadística a fin de evitar afectaciones en los resultados finales.

Evaluar la normalidad de la muestra estadística de tal forma que la escogencia de los estadísticos sea la más apropiada para el análisis.

Hallar el valor promedio de la concentración de c-GMP en la leche, el cual será determinado como valor de referencia de la concentración basal de la molécula en las muestras de estudio.

Establecer si existen diferencias significativas del estado basal entre los diferentes proveedores de una misma empresa con el fin de obtener información adicional acerca de los valores promedios de la concentración de GMP en las muestras de estudio.

Establecer si existe diferencia significativa del estado basal entre los diferentes departamentos del país, con el fin de obtener información adicional acerca de los valores promedios de la concentración de GMP en las muestras de estudio.

# Resultados obtenidos

## SELECCIÓN DE EMPRESAS Y UNIDADES PRODUCTIVAS RURALES PARTICIPANTES

Las empresas participantes junto con sus respectivas unidades productivas rurales fueron seleccionadas y vinculadas a través de la convocatoria publicada el 19 de mayo de 2020 por Colombia Productiva, cuyo objetivo fundamental fue el establecimiento de criterios para la elección de “10 empresas del sector lácteo con código CIU 1040, con el fin de que participaran en el proyecto de análisis y caracterización la leche, con el propósito de mejorar los estándares de acopio, procesamiento y distribución relevantes para asegurar la calidad e inocuidad del producto final”.

Los criterios que cumplieron las empresas seleccionadas fueron los siguientes:

- Estar debidamente constituida como persona jurídica o cooperativa, perteneciente al sector lácteo con actividad económica del CIU 1040. Para acreditar esta condición se deberá anexar Certificado de Existencia y Representación Legal con fecha de expedición no mayor a treinta (30) días a la fecha de cierre de la convocatoria.
- Sólo se podrá presentar máximo una (1) postulación de unidad productiva de transformación por empresa.
- La unidad productiva de transformación de la Empresa que será sujeta de intervención, deberá estar ubicada geográficamente en alguno de los siguientes departamentos: Antioquia, Boyacá, Caquetá, Cesar, Cundinamarca o Nariño.
- La empresa deberá diligenciar el anexo 2 “Información básica de la empresa y la información de los proveedores”. En este anexo se solicita información de las tres (3) unidades productivas rurales postuladas para la toma de muestra de leche.
- La empresa deberá diligenciar el anexo 1 “Carta de Postulación – Colombia Productiva”.
- No estar en curso o en causal de Disolución o Liquidación.
- En el transcurso del proyecto, deberá diligenciar una evaluación de satisfacción del consultor respecto a las actividades desarrolladas en el marco del proyecto.

## MUESTREO

Para la presente investigación se realizó un muestreo estratificado una vez se conoció la información de acopio diario de las unidades productivas rurales postuladas por las empresas participantes (Tabla No. 1) que fueron aquellas que cumplieron con todos los requisitos de la convocatoria realizada por Colombia Productiva, mencionados en el punto anterior. Una vez obtenida la información de la producción diaria de las unidades productivas rurales

se estableció la estratificación y se asignó la cantidad de muestras a tomar por cada unidad productiva. (Tabla No. 2).

Nota: Con el fin de salvaguardar la confidencialidad de los datos de las empresas participantes en el estudio, se asigna código de identificación a cada empresa y sus respectivas unidades productivas. Dicha codificación se presenta en el Anexo 1 del presente documento y sólo estará disponible para Colombia Productiva.

**Tabla N°.1:** Información del acopio diario de leche cruda por unidad productiva rural

EMPRESA PARTICIPANTE	UNIDAD PRODUCTIVA RURAL	ACOPIO (LITROS / DÍA)
<b>E1.0</b>	UP1.1	3.900
	UP1.2.	5.800
	UP1.3.	7.400
<b>E2.0</b>	UP2.1.	400
	UP2.2.	100
	UP2.3.	100
<b>E3.0</b>	UP3.1.	140
	UP3.2.	140
	UP3.3.	20
<b>E4.0</b>	UP4.1.	800
	UP4.2.	1.400
	UP4.3.	1.200
<b>E5.0</b>	UP5.1.	40.000
	UP5.2.	4.946
	UP5.3.	2.160
<b>E6.0</b>	UP6.1.	280
	UP6.2.	480
	UP6.3.	140
<b>E7.0</b>	UP7.1.	2.000
	UP7.2.	4.800
	UP7.3.	2.500
<b>E8.0</b>	UP8.1.	630

EMPRESA PARTICIPANTE	UNIDAD PRODUCTIVA RURAL	ACOPIO (LITROS / DÍA)
<b>E9.0</b>	UP8.2.	500
	UP8.3.	300
	UP9.1.	5.400
	UP9.2.	4.200
	UP9.3.	4.000
<b>E10.0</b>	UP10.1.	400
	UP10.2.	800
	UP10.3.	800
<b>TOTAL</b>		91036

Basado en la organización de las unidades productivas rurales (No. de unidad productiva rural por estrato), se asignó peso porcentual a cada estrato, estableciendo así el número de muestras a tomar en cada unidad productiva teniendo en cuenta que cada una debería tener al menos 3 muestras (no menos) dentro del análisis total del estudio. (Tabla No. 3)

**Tabla N°.2:** Estratificación de las unidades productivas rurales por acopio diario y cantidad por estrato

ESTRATO	RANGO DE ESTRATIFICACIÓN (L/día)	CANTIDAD DE UNIDADES PRODUCTIVAS POR ESTRATO	PONDERACIÓN POR CANTIDAD DE UNIDADES PRODUCTIVAS (%)	CANTIDAD DE ACOPIO DIARIO POR ESTRATO (L/día)
<b>1</b>	hasta 1.000	16	53,33	6.030
<b>2</b>	1.000 – 10.000	13	43,33	49.706
<b>3</b>	más de 10.000	1	3,33	40.000

**Tabla No.3:** Cantidad de muestras a tomar por unidad productiva

EMPRESA PARTICIPANTE	UNIDAD PRODUCTIVA POSTULADA	ACOPIO (L/día)	ESTRATO	PONDERACIÓN DE ACUERDO AL ACOPIO DIARIO X TOTAL DE LITROS DE CADA ESTRATO (%)	CANTIDAD DE MUESTRAS A TOMAR DE ACUERDO AL MUESTREO ESTRATIFICADO	MÍNIMO DE MUESTRAS A TOMAR ESTABLECIDAS	TOTAL DE MUESTRAS
E1.0	UP1.1	3.900	2	7,80	4	3	7
	UP1.2	5.800	2	11,70	6	3	9
	UP1.3	7.400	2	14,90	7	3	10

EMPRESA PARTICIPANTE	UNIDAD PRODUCTIVA POSTULADA	ACOPIO (L/día)	ESTRATO	PONDERACIÓN DE ACUERDO AL ACOPIO DIARIO X TOTAL DE LITROS DE CADA ESTRATO (%)	CANTIDAD DE MUESTRAS A TOMAR DE ACUERDO AL MUESTREO ESTRATIFICADO	MÍNIMO DE MUESTRAS A TOMAR ESTABLECIDAS	TOTAL DE MUESTRAS
E2.0	UP2.1	400	1	6,60	4	3	7
	UP2.2	100	1	1,70	1	3	4
	UP2.3	100	1	1,70	1	3	4
E3.0	UP3.1	140	1	2,30	1	3	4
	UP3.2	140	1	2,30	1	3	4
	UP3.3	20	1	0,30	0	3	3
E4.0	UP4.1	800	1	13,30	8	3	11
	UP4.2	1.400	2	2,8	1	3	4
	UP4.3	1.200	2	2,4	1	3	4
E5.0	UP5.1	40.000	3	100,0	4	3	7
	UP5.2	4.946	2	10,0	5	3	8
	UP5.3	2.160	2	4,3	2	3	5
E6.0	UP6.1	280	1	4,6	3	3	6
	UP6.2	480	1	8,0	5	3	8
	UP6.3	140	1	2,3	1	3	5
E7.0	UP7.1	2.000	2	4,0	2	3	6
	UP7.2	4.800	2	9,7	5	3	8
	UP7.3	2.500	2	5,0	2	3	5
E8.0	UP8.1	630	1	10,4	6	3	9
	UP8.2	500	1	8,3	5	3	8
	UP8.3	300	1	5,0	3	3	6
E9.0	UP9.1	5.400	2	10,9	5	3	8
	UP9.2	4.200	2	8,4	4	3	7
	UP9.3	4.000	2	8,0	4	3	7
E10.0	UP10.1	400	1	6,6	4	3	7
	UP10.2	800	1	13,3	8	3	11
	UP10.3	800	1	13,3	8	3	11
<b>TOTAL</b>							<b>203</b>

Una vez establecida la cantidad y características para la toma de muestras por unidad productiva, se procedió con el ejercicio en campo el cual fue realizado de acuerdo al procedimiento establecido en el anexo 2. Para el presente estudio, todas las muestras fueron tomadas y analizadas en su totalidad sin tener ninguna desviación a los procedimientos analíticos establecidos.

En la tabla No. 4 se observan las fechas de la ejecución del muestreo para cada una de las unidades productivas postuladas.

**Tabla No. 4:** Fechas de ejecución muestreos realizados por unidad productiva rural.

EMPRESA	UNIDAD PRODUCTIVA RURAL	FECHA EJECUCIÓN
<b>E1.0</b>	UP1.1	2020-07-03
	UP1.2	2020-07-02
	UP1.3	2020-07-02
<b>E2.0</b>	UP2.1	2020-06-12
	UP2.2	2020-06-12
	UP2.3	2020-06-12
<b>E3.0</b>	UP3.1	2020-06-18
	UP3.2	2020-06-17
	UP3.3	2020-06-17
<b>E4.0</b>	UP4.1	2020-06-19
	UP4.2	2020-06-18
	UP4.3	2020-06-17
<b>E5.0</b>	UP5.1	2020-06-10
	UP5.2	2020-06-10
	UP5.3	2020-06-10
<b>E6.0</b>	UP6.1	2020-06-24
	UP6.2	2020-06-23
	UP6.3	2020-06-23
<b>E7.0</b>	UP7.1	2020-06-25
	UP7.2	2020-06-25
	UP7.3	2020-06-26
<b>E8.0</b>	UP8.1	2020-06-23
	UP8.2	2020-06-23
	UP8.3	2020-06-24
<b>E9.0</b>	UP9.1	2020-06-24
	UP9.2	2020-07-01
	UP9.3	2020-07-25
<b>E10.0</b>	UP10.1	2020-06-19
	UP10.2	2020-06-18
	UP10.3	2020-06-18

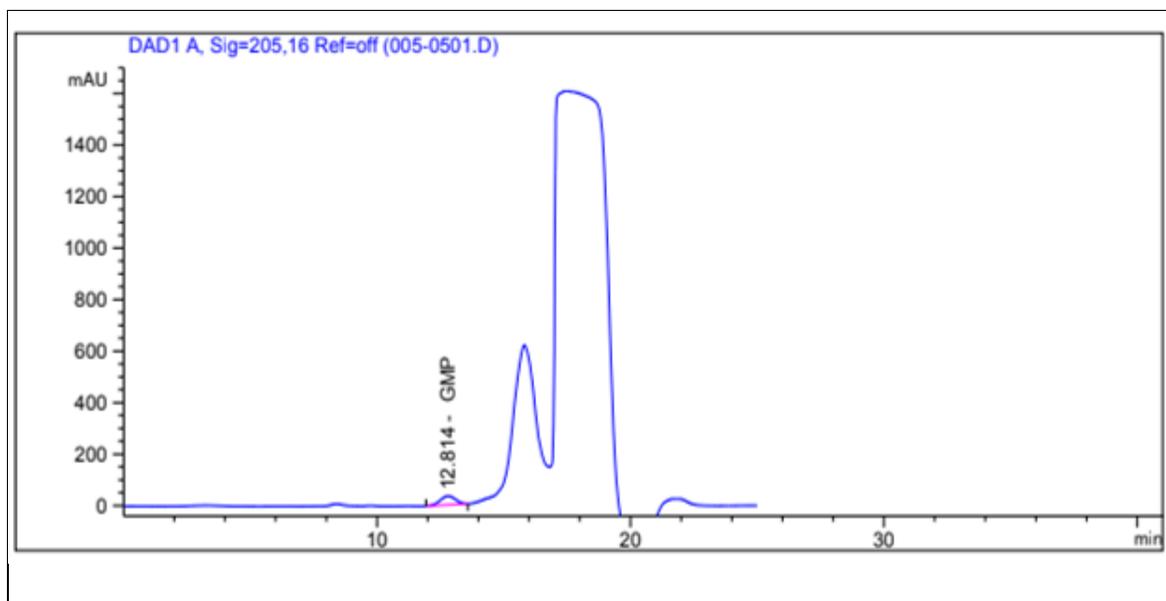
## PROCESO ANALÍTICO

Las muestras fueron allegadas al laboratorio cumpliendo las características determinadas para el muestreo. Posteriormente fueron radicadas en plataforma digital disponible en el laboratorio con el fin de asignar un código de ingreso que garantiza su identificación unívoca, la trazabilidad dentro del proceso y su asignación como ciegos a los analistas con el fin de evitar sesgos.

La cuantificación de c-GMP fue realizada por medio de la técnica analítica cromatografía líquida de alta eficiencia (HPLC), la cual fue validada siguiendo lineamientos internacionales (EURA-

CHEM, AOAC) plasmados en procedimientos propios del Laboratorio y cuyos resultados se presentan en el documento F-Cso-002 Informe de Validación Y/O Confirmación N°. FQ-00001-0 Determinación del Glicomacropéptido de Caseína (C-Gmp) en Leche por Cromatografía Líquida de Alta Eficiencia (Hplc)", en la gráfica N°. 1 se observa el perfil cromatográfico típico, obtenido del análisis de c-GMP por cromatografía líquida

Los ensayos fueron realizados de acuerdo con el documento P-CSO-001 "Método de análisis para la determinación del Glicomacropéptido de Caseína (c-gmp) en Leche por Cromatografía Líquida de Alta Eficiencia (hplc)".



**Gráfica No. 1:** Perfil cromatográfico obtenido de una muestra de leche cruda para la cuantificación de c-GMP.

# Análisis estadístico para establecer las medidas de centralidad y dispersión

## REMOCIÓN DE DATOS EQUIVOCADOS

La NTC-ISO/IEC 17043, literal B.2.5 y el Protocolo Armonizado IUPAC recomiendan: "Eliminar equivocaciones obvias de un conjunto de datos en una etapa temprana de un análisis, antes de emplear cualquier procedimiento robusto o cualquier prueba a fin de identificar valores atípicos estadísticos"<sup>1</sup>. De acuerdo con la NTC-ISO 13528:2015 se indica que estos datos deberían ser tratados por separado y retenerse del conjunto para evaluar su pertinencia de inclusión dentro del análisis estadístico.

Los resultados obtenidos de las muestras analizadas de la unidad productiva rural UP 2.2, mostraron una evidente distancia del conjunto de muestras analizado inicialmente; razón por la cual y ante duda razonable de dichos resultados, se procedió a tomar nuevas muestras para confirmar los hallazgos, encontrando en el segundo conjunto de muestras correlación con el conjunto de datos de resultados de análisis de la población muestreada. Es decir, existió desviación aparente en el primer conjunto de muestras.

Cabe aclarar que se realizó verificación de los procedimientos de muestreo aplicados y no se encontró ninguna desviación.

## REVISIÓN VISUAL DE DATOS

Un primer acercamiento al comportamiento de los resultados obtenidos en la caracterización de c-GMP en leche cruda es por medio de la estadística descriptiva (tabla No.5); aunque no es concluyente, si puede dar indicios del comportamiento de la muestra estadística tomada en el estudio y evidenciar posibles datos anómalos que puedan afectar la estimación estadística.

**Tabla N°. 5:** Estadística descriptiva

<b>Contenido de c-GMP (mg/L)</b>	
<b>Media</b>	44,57
<b>Error típico</b>	1,91
<b>Mediana</b>	40,5
<b>Moda</b>	36,6
<b>Desviación estándar</b>	27,25
<b>Varianza de la muestra</b>	742,68
<b>Curtosis</b>	36,18
<b>Coficiente de asimetría</b>	5,04
<b>Rango</b>	283,5
<b>Mínimo</b>	0
<b>Máximo</b>	283,5
<b>Suma</b>	9047,89

<sup>1</sup> NTC-ISO 17043:2010: Evaluación de la conformidad. Requisitos generales para los ensayos de aptitud.

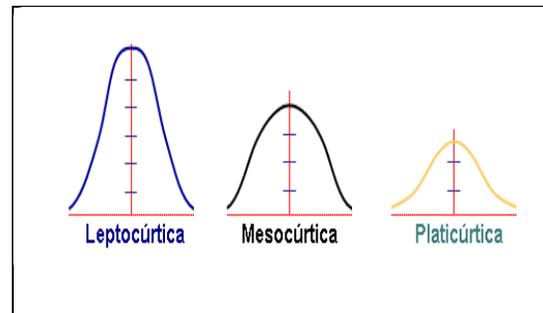
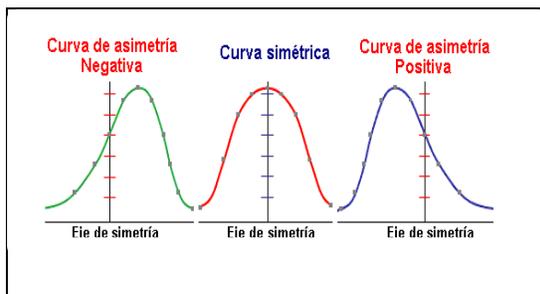
<b>Contenido de c-GMP (mg/L)</b>	
<b>Cuenta</b>	203
<b>Nivel de confianza (95,0%)</b>	3,771

De acuerdo con los resultados obtenidos, se observa que la muestra estadística presenta una desviación estándar alta, lo que puede indicar gran dispersión en los resultados o la posible presencia de datos atípicos.

El coeficiente de asimetría indica si los datos se distribuyen de forma uniforme alrededor del punto central (ver figura No. 1). Para el caso, se presenta un coeficiente de asimetría positivo mayor a cero, indicando que los valores encontrados se tienden a reunir más en la parte izquierda que en la derecha de la media.

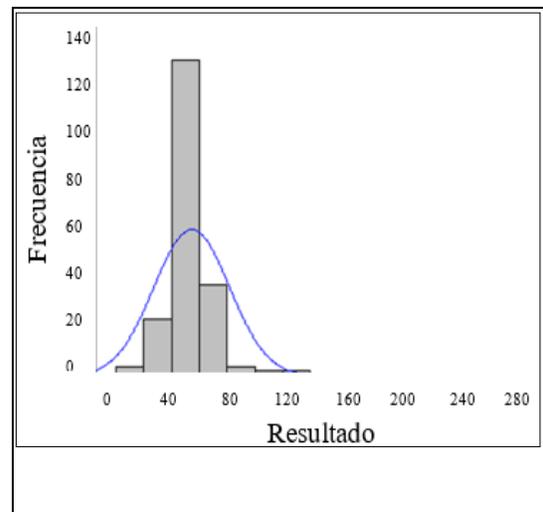
El coeficiente de curtosis determina el grado de concentración que presentan los valores en la región central de la distribución (ver figura N°. 1).

El coeficiente (mayor a 3) indica que la muestra es leptocurtica (gran concentración de valores en la región central).



**Figura N°.1.** Clases de curvas estadísticas en el análisis de datos.

Posteriormente se construyó el histograma de la muestra estadística obtenida del estudio, los resultados se observan a continuación (Gráfica N°.1)

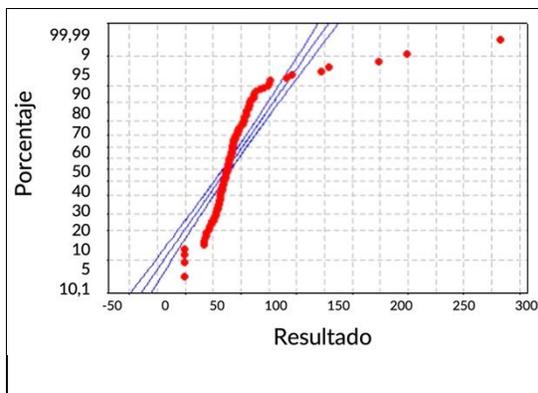


**Gráfica N°. 2:** Histograma obtenido del análisis de datos (previo al análisis de anomalos)

El histograma muestra una curva cercana a la gaussiana (en forma de campana), se evidencian posibles datos anómalos dentro de la muestra estadística

analizada y se confirma el comportamiento de los datos analizado previamente, con los resultados de la estadística descriptiva: **“curva asimétrica positiva leptocurtica.”**

Para finalizar el primer acercamiento a los datos obtenidos, se realizó una gráfica de probabilidad a fin de detectar posibles datos atípicos (Gráfica N°. 3).



**Gráfica N°. 3:** Gráfica de probabilidad obtenido del análisis de datos (previo al análisis de anómalos)

De acuerdo con la gráfica, se confirma la existencia de valores atípicos, los cuales se alejan de la población en general, adicionalmente la muestra al parecer no presenta un comportamiento normal; sin embargo para estimarlo se debe realizar un tratamiento para la remoción de posibles datos anómalos y emplear estadística inferencial para concluir.

## DETERMINACIÓN DE VALORES ATÍPICOS

Entiéndase como valor atípico aquel que dentro de un conjunto de datos es incoherente con los otros miembros del mismo conjunto<sup>2</sup>; por lo que se hace necesario y de acuerdo a la descriptiva analizada previamente, estimar aquellos datos atípicos de la muestra obtenida.

En el presente estudio se utilizó el estadístico de Grubbs para la determinación de los valores aberrantes; la estadística de Grubbs,  $G_p$ , es:

$$\text{Para el valor más alto: } G_p = \frac{x_p - \bar{x}}{s}$$

$$\text{Para el valor más bajo: } G_1 = \frac{\bar{x} - x_1}{s}$$

Las hipótesis definidas para el estadístico son:

$H_0$  = No hay valores atípicos en el conjunto de datos.

$H_1$  = Hay exactamente un valor en el conjunto de datos.

Determinación del nivel de significancia ( $\alpha$ ):

En general es utilizado un  $\alpha = 0.05$  para proyectos de calidad

Una vez realizado el tratamiento de datos atípicos, se obtuvieron 7 datos atípicos de 203 datos analizados (aproximadamente el 3% de los datos recolectados); en la siguiente tabla se discriminan los valores.

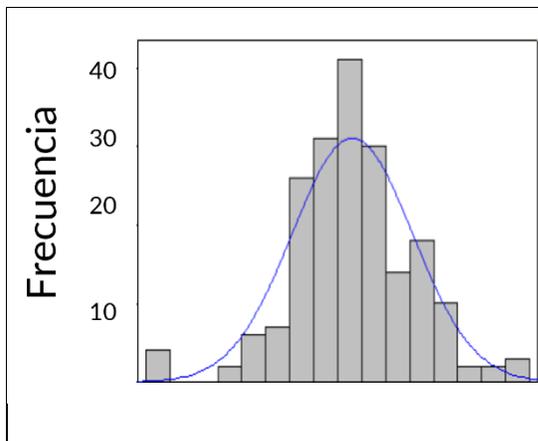
<sup>2</sup> ISO 3529-1: Accuracy (trueness and precision) of measurement methods and results. Part 1: General principles and definitions

EMPRESA	UNIDAD PRODUCTIVA RURAL	CONTENIDO DE c-GMP (mg/L)
E5.0	UP5.2	199,6
E3.0	UP3.2	283,5
E3.0	UP3.3	122,6
E3.0	UP3.3	129,4
E2.0	UP2.2	91,3
E8.0	UP8.2	95,8
E6.0	UP6.1	173,8

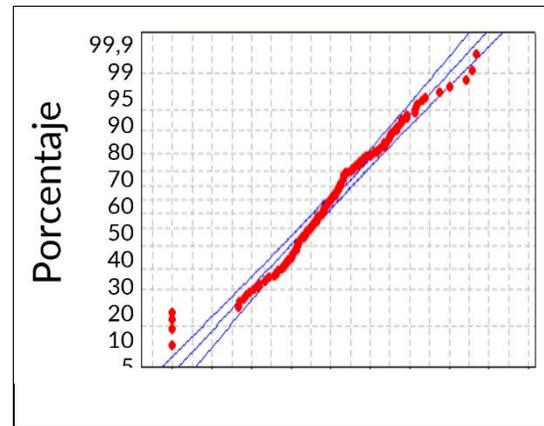
**Tabla N°.6:** Datos atípicos excluidos del análisis estadístico.

## REVISIÓN DE LA DISTRIBUCIÓN DE DATOS

Una vez excluidos los datos anómalos se observa un ajuste en la distribución de los mismos, tal y como se evidencia en el histograma sin los datos atípicos (Gráfica N°. 4) y la gráfica de probabilidad (Gráfica N°. 5).



**Gráfica N°. 4:** Histograma sin datos atípicos.



**Gráfica N°. 5:** Gráfica de probabilidad sin datos atípicos.

Posteriormente se procede a revisar la normalidad de la muestra con el fin evaluar la distribución de los datos y así establecer los estadísticos más apropiados a utilizar para el conjunto de datos. Es de anotar, que en ausencia de normalidad se deben utilizar estadísticos de centralidad y de distancia, no paramétricos y robustos que no se vean afectados por esta característica.

Para el estudio de la normalidad se utilizaron tres pruebas estadísticas, a saber:

- Prueba de Anderson-Darling
- Prueba de normalidad de Ryan-Joiner
- Prueba de normalidad de Kolmogorov – Smirnov

El desarrollo de dichas pruebas está dado como se presenta a continuación:

1. Planteamiento de hipótesis  
 $H_0$ =La distribución de datos es normal.

$H_a$ =La distribución de los datos es no normal.

2. Determinación del nivel de significancia ( $\alpha$ )

En general  $\alpha=0.05$  para proyectos de calidad.

3. Regla de decisión:

Si el p-valor calculado es menor que el nivel de significancia (0.05), se rechaza  $H_0$  es decir, la distribución de los datos no es normal.

La estimación de la normalidad, con los diferentes estadísticos arrojó los siguientes resultados:

ESTADÍSTICO DE PRUEBA	$\alpha$ (nivel de significancia)	p-VALOR CALCULADO	CONCLUSIÓN
Anderson-Darling	0,05	<0,005	La muestra estadística no presenta normalidad
Ryan-Joiner	0,05	<0,010	
Kolgomorv-Smirnof	0,05	<0,010	

**Tabla N°. 7:** Estimación de la normalidad de los resultados obtenidos.

Con los resultados obtenidos se concluye que las muestras analizadas no presentan una distribución normal, por lo tanto, se deben utilizar estadísticos no paramétricos para establecer la tendencia central y distribución del conjunto de datos de los resultados obtenidos en el presente estudio.

### DETERMINACIÓN DE LA MEDIDA CENTRAL Y DISPERSIÓN DE LOS RESULTADOS

De acuerdo con el comportamiento evidenciado del conjunto de datos, se establece que uno de los estadísticos más apropiados para hallar el promedio y dispersión de los resultados es el algoritmo

$A^3$ , el cual transforma los datos originales mediante un proceso denominado winsorización. Los estimados robustos  $x^*$  (media robusta) y  $s^*$  (desviación estándar robusta) se derivan mediante un cálculo iterativo, hasta que el proceso converja. Se asume convergencia cuando no existe cambio de dos iteraciones.

El algoritmo A se calcula como sigue:

<sup>3</sup> ISO 5725-5: 2005: Accuracy (trueness and precision) of measurement methods and results. Part 5: Alternative methods for the determination for the precision of a standard measurement method.

- Se designan los p datos, colocados en orden creciente, por:

$x_1, x_2, \dots, x_i, \dots, x_p$  Se designan por  $x^*$

y  $s^*$  la media y la desviación estándar de robustez de estos datos.

- Se calculan los valores iniciales de  $x^*$  y  $s^*$ , mediante:

$$x^* = \text{mediana de } x_i \quad (i = 1, 2, \dots, p)$$

$$s^* = 1,483 \times \text{mediana de } |x_i - x^*| \quad (i = 1, 2, \dots, p)$$

- Se actualizan los valores de  $x^*$  y  $s^*$  de la siguiente forma

Para cada valor  $x_i$  ( $i = 1, 2, 3, \dots, p$ ), se calcula

$$x_i^* = \begin{cases} x^* - \varphi & \text{si } x_i < x^* - \varphi \\ x^* + \varphi & \text{si } x_i > x^* + \varphi \\ x_i & \text{en cualquier otro caso} \end{cases}$$

Se calculan los nuevos valores de  $x^*$  y  $s^*$  mediante:

$$x^* = \sum_{i=1}^p x_i^* / p$$

$$s^* = 1,134 \sqrt{\sum_{i=1}^p (x_i^* - x^*)^2 / (p-1)}$$

En donde la suma se extiende a i.

Una vez realizada la estimación mediante el procedimiento descrito anteriormente, se obtuvieron los siguientes resultados:

PROMEDIO ROBUSTO ( $x^*$ ) (mg/L)	DESVIACIÓN ESTÁNDAR ( $s^*$ ) (mg/L)	INCERTIDUMBRE* (mg/L)
40,5	10,5	4,2

\*Incertidumbre expandida U del valor estimado, con un factor de cobertura de  $k=2$ , para un intervalo de confianza de aproximadamente el 95%.

Una vez estimada la medida central (promedio robusto) se procede a calcular los límites permitidos teniendo en cuenta la incertidumbre y la desviación estándar establecida para el análisis; cabe aclarar, que dado que el objetivo del presente estudio, busca establecer el valor máximo permitido, el cual corresponde a valores de contenido basal de cGMP en leche, sólo se establece el límite superior calculado.

A continuación, se presenta la fórmula para su cálculo:

$$\text{Límite superior} = X^* + 2 (\sqrt{(s^*)^2 + U^2})$$

En donde:

$x^*$  = Promedio robusto estimado

$s$  = Desviación estándar robusta estimada (igual a  $s^*$ )

U = Incertidumbre expandida estimada

Finalmente, el límite superior permitido de acuerdo a la ecuación es de:

Límite Superior (mg/L)
63,1

# Comparación de resultados

## COMPARACIÓN DE LOS RESULTADOS OBTENIDOS ENTRE LOS PROVEEDORES DE UNA MISMA EMPRESA

Se realizó una comparación de los resultados obtenidos entre las unidades productivas en cada una de las empresas participantes para conocer diferencias estadísticamente significativas del estado basal del c-GMP. Para esta comparación se utilizó el estadístico no paramétrico mediana de Mood, en el cual se contempla lo siguiente:

### 1. PLANTEAMIENTO DE HIPÓTESIS

$H_0$  = Las medianas de las unidades productivas no son significativamente diferentes.

$H_a$  = Las medianas de las unidades productivas son significativamente diferentes.

### 2. DETERMINACIÓN DEL NIVEL DE SIGNIFICANCIA (A)

En general  $\alpha=0.05$  para proyectos de calidad

### 3. REGLA DE DECISIÓN

Si el p-valor calculado es menor que el nivel de significancia (0.05), se rechaza  $H_0$  es decir, dos o más mediana son diferentes.

Se obtuvieron los siguientes resultados:

**Tabla N°.8:** Comparación de unidades productivas

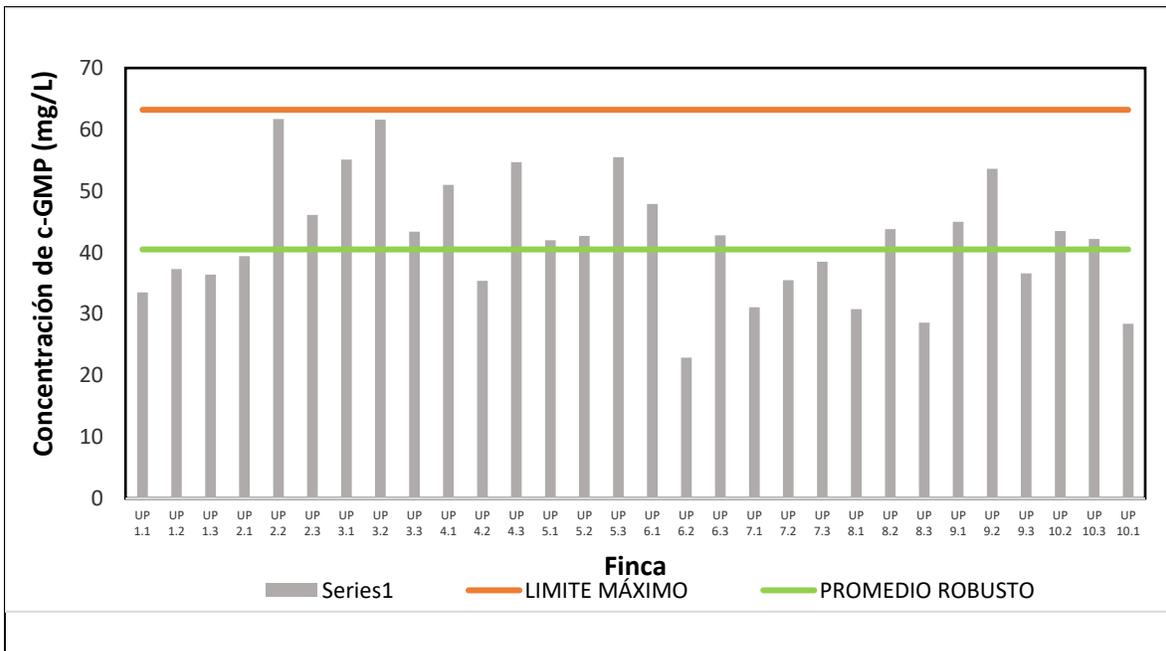
EMPRESA	UNIDAD PRODUCTIVA	MEDIANA POR CADA UNIDAD PRODUCTIVA	MEDIANA GENERAL DE EMPRESA	p-valor	CONCLUSIÓN
E1.0	UP1.1	33,5	36,4	0,319	No hay diferencia significativa entre las medianas obtenidas de las tres unidades productivas
	UP1.2	37,3			
	UP1.3	36,4			

EMPRESA	UNIDAD PRODUCTIVA	MEDIANA POR CADA UNIDAD PRODUCTIVA	MEDIANA GENERAL DE EMPRESA	p-valor	CONCLUSIÓN
E2.0	UP2.1	39,4	43,6	0,023	Hay diferencias significativas entre las medianas obtenidas de las tres unidades productivas. La unidad productiva rural que es diferente a las demás es <b>UP2.2</b>
	UP2.2	61,7			
	UP2.3	46,1			
E3.0	UP3.1	55,1	55,1	0,659	No hay diferencia significativa entre las medianas obtenidas de las tres unidades productivas.
	UP3.2	61,6			
	UP3.3	43,4			
E4.0	UP4.1	51,0	49,1	0,080	No hay diferencia significativa entre las medianas obtenidas de las tres unidades productivas.
	UP4.2	35,4			
	UP4.3	54,7			
E5.0	UP5.1	42,0	43,1	0,157	No hay diferencia significativa entre las medianas obtenidas de las tres unidades productivas
	UP5.2	42,7			
	UP5.3	55,5			
E6.0	UP6.1	47,9	38,9	0,008	Hay diferencias significativas entre las medianas obtenidas de las tres unidades productivas. La unidad productiva rural que es diferente a las demás es <b>UP6.2</b>
	UP6.2	22,9			
	UP6.3	42,8			
E7.0	UP7.1	31,1	31,9	0,298	No hay diferencia significativa entre las medianas obtenidas
	UP7.2	35,5			
	UP7.3	38,5			

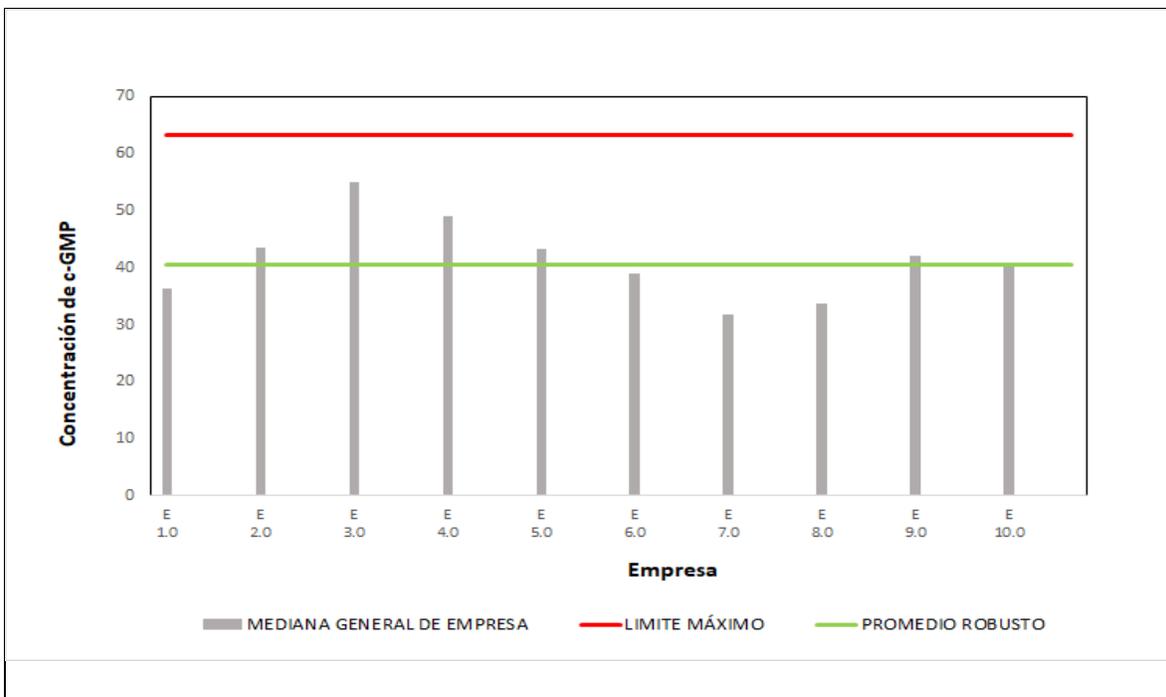
EMPRESA	UNIDAD PRODUCTIVA	MEDIANA POR CADA UNIDAD PRODUCTIVA	MEDIANA GENERAL DE EMPRESA	p-valor	CONCLUSIÓN
					de las tres unidades productivas
<b>E8.0</b>	<b>UP8.1</b>	30,8	33,6	0,163	No hay diferencia significativa entre las medianas obtenidas de las tres unidades productivas
	<b>UP8.2</b>	43,8			
	<b>UP8.3</b>	28,6			
<b>E9.0</b>	<b>UP9.1</b>	45,0	42,2	0,013	Hay diferencias significativas entre las medianas obtenidas de las tres unidades productivas. La unidad productiva rural que es diferente a las demás es <b>UP9.2</b>
	<b>UP9.2</b>	53,6			
	<b>UP9.3</b>	36,6			
<b>E10.0</b>	<b>UP10.1</b>	28,4	40,5	0,021	Hay diferencias significativas entre las medianas obtenidas de las tres unidades productivas. La unidad productiva rural que es diferente a las demás es <b>UP10.1</b>
	<b>UP10.2</b>	43,5			
	<b>UP10.3</b>	42,2			

En las siguientes gráficas se encuentra el comportamiento de la concentración de c-GMP por unidad productiva rural (Gráfica N°. 5) y empresas (Gráfica N°. 6) comparado con el promedio y el máximo obtenidos en el presente estudio.

**Gráfica N°. 6:** Comportamiento de la concentración de c-GMP por unidades productivas.



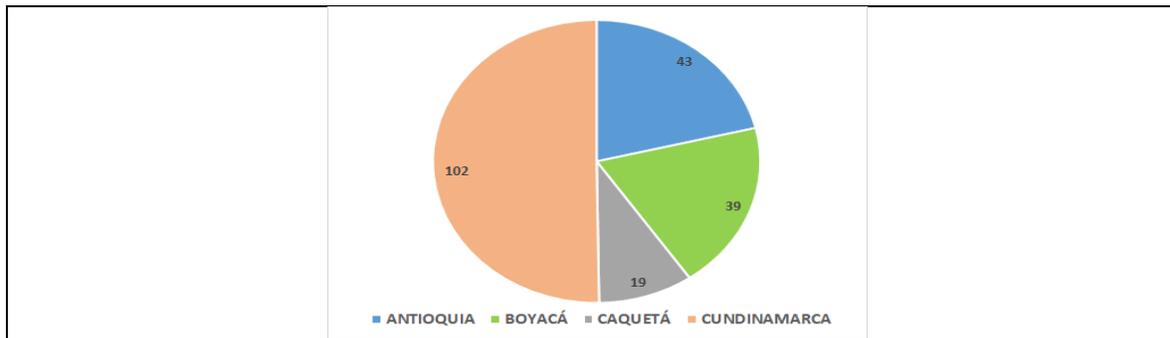
**Gráfica N° 7:** Comportamiento de la concentración de c-GMP por Empresa.



## COMPARACIÓN DE LOS RESULTADOS OBTENIDOS ENTRE DEPARTAMENTOS

Hicieron parte del presente estudio 30 unidades productivas rurales, ubicadas geográficamente en 4 departamentos de Colombia, Antioquia, Cundinamarca, Caquetá y Boyacá, la cantidad de muestras tomadas en cada uno de los sitios se puede evidenciar en la siguiente gráfica:

**Gráfica N° 8:** Distribución del número de muestras analizadas por departamento.



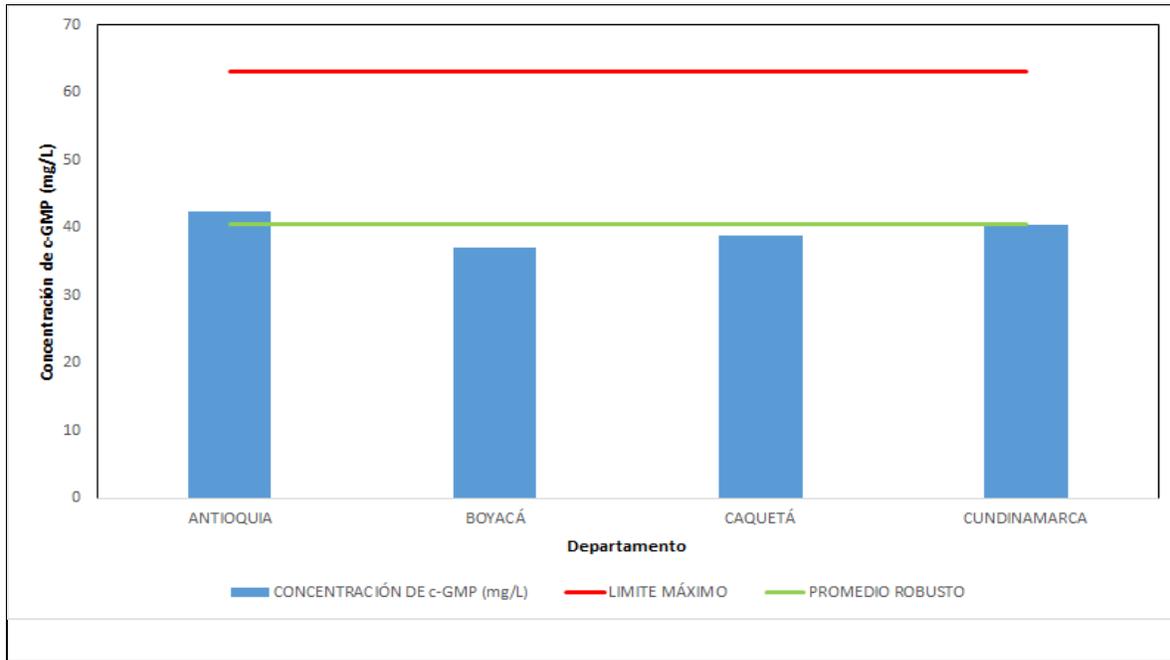
Se realizó una comparación de los resultados obtenidos entre los departamentos que hicieron parte del presente estudio, con el fin de conocer diferencias estadísticamente significativas de la concentración basal del c-GMP. Para esta comparación se utilizó el estadístico no paramétrico mediana de Mood, explicado anteriormente. Se obtuvieron los siguientes resultados:

**Tabla N°.9:** Comparación de los resultados obtenidos entre los departamentos.

DEPARTAMENTO	MEDIANA GENERAL DEPARTAMENTO	MEDIANA GENERAL GLOBAL	p-valor	CONCLUSIÓN
ANTIOQUIA	42,5	40,5	0,114	No hay diferencia significativa entre las medianas obtenidas de los cuatro departamentos que participaron en el presente estudio
BOYACÁ	37,0			
CAQUETÁ	38,9			
CUNDINAMARCA	40,5			

En la siguiente gráfica se encuentra el comportamiento de la concentración de c-GMP por departamento (Gráfica N°. 9) comparado con el promedio y el máximo obtenidos en el presente estudio.

**Gráfica N° 9:** Comportamiento de la concentración de c-GMP por departamento.



## Conclusiones

- La metodología empleada en el estudio permitió realizar el análisis estadístico robusto de los resultados de ensayo para la determinación de glicomacropéptido de caseína (c-GMP) en las muestras de estudio.
- Los criterios para la determinación del número de muestras a tomar por empresa/unidad productiva rural, permitieron dar representatividad al estudio dada la organización por estratos fundamentado en volumen de producción de cada predio.
- La aplicación del test de Grubbs en la muestra estadística, permitió la identificación de datos anómalos del estudio, garantizando la validez de los datos finales.
- La evaluación de la normalidad en la muestra evidenció ausencia de la misma, lo que orientó la elección del uso de estadística no paramétrica robusta para el análisis de los datos.
- Se encontraron diferencias significativas entre los diferentes proveedores de una misma empresa, lo que significa que existen variaciones entre los proveedores que pueden orientar evaluaciones posteriores acerca de sus procesos y productos.
- No se encontraron diferencias significativas en el estado basal entre los diferentes departamentos del país, lo que indica que la ubicación geográfica de las unidades productivas no influencia los resultados. Así mismo y entendiendo que los departamentos de Antioquia, Cundinamarca, Boyacá y Caquetá representan cerca del 80%<sup>4</sup> de acopio de la leche del país y al verificar que no hay influencia de la región geográfica, los resultados de este estudio son un punto de partida para ser tomados como nacionales. Es necesario ampliar la cobertura del proyecto y ampliar el número de muestras para robustecer el estudio.
- El valor promedio robusto obtenido fue de 40,5 mg/L, el cual representa el punto de partida para la determinación del valor basal de concentración de c-GMP. El cálculo posterior teniendo en cuenta su incertidumbre y la desviación estándar encontrada en este estudio, arrojó un valor de 63,1 mg/L; el cual resulta más apropiado para la determinación de un valor de referencia, y será incluido como valor máximo permitido en la propuesta normativa.

---

<sup>4</sup> Valor estimado para el mes de mayo de 2020 según USP (Unidad de Seguimiento de Precios) Volumen de leche captada por la industria, total departamental. Mayo 2019 - Mayo 2020. Recuperado desde <http://uspleche.minagricultura.gov.co/volumen-de-acopio-total-departamental.html>



2021