

METODOLOGÍA PARA EL CONTROL DE LA CALIDAD DEL CAP AUTOMÁTICO ADOPTADA POR LA COMUNIDAD AUTÓNOMA DE ARAGÓN

El Departamento de Desarrollo Rural y Sostenibilidad del Gobierno de Aragón acorde con el plan de mejora de las informaciones del SIGPAC, adoptó a lo largo del año 2014 los procesos de cálculo masivos del Coeficiente de Admisibilidad en Pastos que tienen en cuenta las limitaciones de acceso del ganado a determinadas zonas de las parcelas mediante la definición de un nuevo factor de vegetación obtenido a partir de datos LIDAR. (CAP-2015).

El Ministerio de Agricultura, Alimentación y Medio Ambiente, una vez obtenidos los resultados del CAP sobre la mayor parte del territorio nacional, se ha elaborado un procedimiento que determina cómo deben hacerse las comprobaciones de los resultados obtenidos en un determinado ámbito geográfico. Este procedimiento se aplica en zonas geográficas en las que se considera que las características de los pastos son homogéneas. En las zonas en las que los resultados no se consideren válidos y en las que no pueda establecerse un factor de corrección de los mismos será necesario determinar la admisibilidad de los terrenos de pastos con otros procedimientos técnicos complementarios.

El procedimiento general se plantea como un proceso que consta de dos etapas. En la primera etapa se diseña un plan de muestreo en tres fases para estimar la superficie admisible de pasto total en la zona y construir un intervalo de confianza para la verdadera superficie admisible de pasto (verdad-terreno). En la segunda etapa se valida el procedimiento de cálculo automático del CAP a nivel de zona: el procedimiento automático se considera que es válido si la superficie admisible de pasto en la zona que resulta de su aplicación (CAP_ 2015 automático), se encuentra dentro del intervalo de confianza construido en la primera etapa. La población de la zona será estratificada a fin de poder estudiar por separado los grupos de recintos con características diferenciadas. Las tres fases en las que se fracciona el plan de muestreo son las siguientes: la primera consiste en el proceso de obtención del dato de Coeficiente de Admisibilidad en Pastos automático por recinto, esta fase se ha realizado a lo largo del año 2014, la segunda fase consiste en la selección de una

muestra aleatoria de recintos que serán estudiados posteriormente por fotointerpretación y **la tercera** fase consiste en la selección de una muestra aleatoria de recintos entre los fotointerpretados para su comprobación final sobre el terreno.

Este proceso se ha realizado en el territorio aragonés de la siguiente forma:

1º) Extracción de todos los recintos de la CCAA de Aragón en que su uso del suelo estaba calificado como PA, PS y PR.

Los atributos obtenidos fueron:

- Identificativos del recinto: Provincia, municipio catastral, agregado, zona, polígono, parcela y recinto.
- Rendimiento para el seco. Su dominio está definido desde 1,5 a 2,7 tm/ha
- Superficie catastral del recinto.
- Uso del suelo (PA, PS, PR).
- Atributos de parcela en relación a su CAP tales como:
 - o Pendiente. Accesibilidad pastable relacionada con la pendiente de la parcela.
 - o Suelo. Cobertura del mismo obtenido mediante el NDVI (índice que relaciona las diferencias del Infrarrojo cercano con el rojo con la suma del Infrarrojo cercano con el rojo). Da idea de la productividad pastable.
 - o Altura de la vegetación obtenido mediante metodología Radar. Define la altura de la vegetación.
- CAP automático. Capacidad de admisibilidad de pastos calculado con los tres atributos anteriores (pendiente, factor suelo y LIDAR).

El total de recintos es de 681.952, con una superficie total de 950.117 ha, este será el universo a validar del Coeficiente de Admisibilidad de Pastos (CAP) de cada uno de los recintos.

Por tanto, Aragón pretende validar los Coeficientes de Admisibilidad de Pastos considerando la totalidad del territorio con uso de suelo susceptible a ser pastado.

2º) Proceso de extracción de la muestra para la fotointerpretación:

La CCAA de Aragón optó establecer un muestreo estratificado aleatorio, considerando once estratos según tres tipos de variables combinadas de aquellos recintos con CAP, siendo:

1. Rendimientos del seco (rango de 1,5 a 2,7 tm/ha).
2. Usos del suelo (PA, PS, PR).

3. Dimensión del CAP automático imputado.

De esta forma obtuvo tres zonas de comportamiento productivo homogéneo. A su vez cada zona isoproductiva fue clasificada en estratos que adsorbieran la máxima variabilidad entre ellos, siendo las variables de diferenciales el uso y el CAP. La primera y segunda zonas están segregadas en cuatro estratos cada una y la tercera en tres.

Las razones de este tipo de muestreo estratificado aleatorio son:

1. Los estratos están constituidos por CAP lo más homogéneo posible. Imputando la heterogeneidad en las diferencias de estratos.
2. Esta estratificación proporciona una ganancia en la precisión estadística neta muy considerable.
3. Hay una optimización en los costes económicos ya que los estratos van ligados a zonas o regiones agroclimáticas homogéneas (de comportamiento isoproductivo) y por tanto hay disminución de costes en los desplazamientos.
4. Con este tipo de muestreo se puede llegar a estudiar la totalidad de las parcelas aragonesas que disponen de la variable CAP.

La afijación de la muestra es el reparto o asignación o distribución del tamaño muestral en los diferentes estratos. La afijación es proporcional al tamaño total del estrato (número de parcelas) y a la varianza del estrato.

Para este tipo de muestreo se ha adoptado un Error de Muestreo del 3,1% de la media del CAP y un Coeficiente de confianza del 95% para la afijación.

De los 681.952 recintos se extrajeron los recintos mayores de 0,5 ha y menores de 500 ha que hubieran estado declarados en los últimos 5 años. Como resultado el universo a muestrear se redujo a **94.322** recintos y **766.299** ha y 11 estratos.

ESTRATIFICACIÓN DE LA POBLACIÓN		POBLACIÓN OBJETO DE ESTUDIO		DIMENSIÓN DE LA MUESTRA		
		Nº Recintos	Superficie	Nº Recintos	Superficie	%Error
1	PR_1,5-2_25_55	14.856	131.745	253	2.541	3,13
2	PR_1,5-2_>55-100	30.991	254.655	97	730	3,11
3	PA_1,5-2_25_100	7.748	107.753	651	10.588	3,24
4	PS_1,5-2_25_100	2.228	5.023	198	389	3,25
5	PR_2,2-2,5_25-55	6.962	45.041	297	2.001	3,17
6	PR_2,2-2,5_>55-100	11.231	72.046	94	677	3,11
7	PA_2,2-2,5_25-100	3.772	30.342	688	6.988	3,43
8	PS_2,2-2,5_25-100	653	1.636	192	459	3,69
9	PR_2,7_25-100	10.826	97.397	739	5.370	3,21
10	PA_2,7_25-100	2.909	14.462	467	2.440	3,38
11	PS_2,7_25-100	2.146	6.197	147	509	3,21
TOTALES		94.322	766.299	3.823	32.691	3,10

Se obtiene la distribución de recintos y sus superficies totales para cada zona homogénea y estrato. Este diseño pretende minimizar la variabilidad dentro de los estratos y maximizarla entre ellos. Las zonas homogéneas estarán comprendidas entre 1,5 y 2 tm para la primera, entre 2,2 y 2,5 tm para la segunda y de 2,7 tm para la tercera.

¿Qué precisión se necesita de la muestra seleccionada para cada estrato? La variable a controlar es el CAP automático, que según el FEGA deberá adecuarse a una estratificación que tendrá unos intervalos a partir de 25% de CAP de 10 puntos porcentuales. Esto supone que aceptando un error de muestreo del $\pm 3\%$ asumimos un total del 6%, para el intervalo considerado del 10% entre valores continuos del CAP.

Para calcular el número de recintos a muestrear consideraremos, el peso de los recintos totales del estrato, la varianza de los mismos así como el error porcentual asumido 3% de la media del CAP del estrato. Con todo esto vemos que el número total a fotointerpretar es de 3.823 recintos y 32.691 ha, el error medio porcentual es de 3,10%. El método de muestreo ha sido aleatorio estratificado de asignación proporcional con muestras independientes en cada estrato.

$$n = \frac{Z_{1-\alpha}^2 N}{(N-1) e^2 + \sigma Z_{1-\alpha}^2}$$

n: Nº de recintos de cada estrato.

Z: t de student al 0,05% (1- α).

N: Nº de recintos totales del estrato.

e: Porcentaje de error del CAP medio del estrato.

σ : Varianza del CAP del estrato

Dentro de cada estrato se hace una selección aleatoria con SPSS para aquellas parcelas que van a ser fotointerpretadas. Los resultados de los estadísticos poblacionales y muestrales son los siguientes:

ESTADÍSTICOS POBLACIONALES Y MUESTRALES DEL CAP												
ESTRATOS (Uso + RenSec + CAP Auto)		Media		IC 95%				SD		CV		
		Población	Muestra	Población		Muestra		Población	Muestra	Población	Muestra	
PR 1,5-2 - 25-55	1	42,72	42,75	42,55	42,90	41,43	44,06	10,85	10,62	25,39	24,84	
PR 1,5-2 - >55-100	2	84,81	86,70	84,66	84,95	84,08	89,32	13,20	13,01	15,57	15,00	
PA 1,5-2 - 25-100	3	49,57	50,41	49,11	50,04	48,74	52,07	20,90	21,64	42,15	42,94	
PS 1,5-2 - 25-100	4	86,14	86,21	85,31	86,98	83,38	89,04	20,09	20,20	23,33	23,43	
PR 2,2-2,5 - 25-55	5	40,02	39,95	39,75	40,28	38,68	41,22	11,15	11,09	27,85	27,77	
PR 2,2-2,5 - >55-100	6	86,57	86,01	86,33	86,82	83,17	88,85	13,34	13,88	15,41	16,14	
PA 2,2-2,5 - 25-100	7	48,57	48,77	47,86	49,28	47,11	50,44	22,29	22,24	45,90	45,60	
PS 2,2-2,5 - 25-100	8	81,65	80,63	80,01	83,28	77,45	83,80	21,27	22,33	26,06	27,70	
PR 2,7 - 25-100	9	48,29	47,43	47,89	48,70	45,89	48,97	21,51	21,33	44,55	44,97	
PA 2,7 - 25-100	10	33,93	34,01	33,47	34,39	32,84	35,19	12,67	12,93	37,32	38,01	
PS 2,7 - 25-100	11	87,32	88,74	86,59	88,06	85,86	91,62	17,35	17,66	19,86	19,90	

De esta forma podemos comprobar la similitud de los estadísticos y como consecuencia su representatividad.

3º) Extracción de la muestra para la identificación del CAP en CAMPO:

De los 3.823 recintos fotointerpretados se extrae una muestra para ser visitados en campo, los resultados son:

ESTRATIFICACIÓN DE LA POBLACIÓN		POBLACIÓN OBJETO DE ESTUDIO		MUESTRA Fotointerpretada		MUESTRA Campo	
		Nº Recintos	Superficie	Nº Recintos	Superficie	Nº Recintos	Superficie
1	PR_1,5-2_25_55	14.856	131.745	253	2.541	33	283
2	PR_1,5-2_>55-100	30.991	254.655	97	730	12	136
3	PA_1,5-2_25_100	7.748	107.753	651	10.588	69	765
4	PS_1,5-2_25_100	2.228	5.023	198	389	21	38
5	PR_2,2-2,5_25-55	6.962	45.041	297	2.001	34	102
6	PR_2,2-2,5_>55-100	11.231	72.046	94	677	9	16
7	PA_2,2-2,5_25-100	3.772	30.342	688	6.988	82	680
8	PS_2,2-2,5_25-100	653	1.636	192	459	25	35
9	PR_2,7_25-100	10.826	97.397	739	5.370	79	401
10	PA_2,7_25-100	2.909	14.462	467	2.440	57	137
11	PS_2,7_25-100	2.146	6.197	147	509	15	20
TOTALES		94.322	766.299	3.823	32.691	436	2.613

El método de obtención de la muestra aleatoria de recintos para visitar en campo es, al igual que para los fotointerpretados, afijar la dimensión de la muestra de campo y con SPSS realizar una selección aleatoria del número de recintos por estrato.

4º) Estimación por Regresión:

Dentro de la estimación por razones, la estimación por regresión es un método que se emplea para correlacionar dos variables muestrales, obteniendo una función con su correspondiente R2 y significación de su coeficiente de regresión y ordenada en el origen. Cuanto mayor sea la

correlación (Coeficiente de Correlación de Pearson) entre las dos variables muestrales, mejor funcionará el modelo. Coeficientes de regresión próximos a la unidad y con Coeficiente de Correlación también próximo a uno, significará que la muestra fotointerpretada se corresponderá a la muestra campo.

En este caso se trata de comprobar la calidad de la estimación por métodos automáticos del CAP y para ello seleccionamos para cada estrato una muestra de recintos que analizaremos mediante procesos de fotointerpretación. Por otro lado, de la muestra a fotointerpretar seleccionaremos una submuestra aleatoria cuyos recintos serán visitados directamente en campo de forma que se calcule su C AP “verdad terreno”

Se ha utilizado para el cálculo de la estimación por regresión el paquete estadístico SPSS así como la bibliografía siguiente:

- *Muestreo: Diseño y Análisis. Sharon L. Lohr*
- *Técnicas de Muestreo Estadístico. Cesar Pérez López*
- *Problemas de muestreo. INE. Sánchez Crespo.*

Proceso de cálculo:

1º) de cada estrato se ha tomado los datos pareados de cada recinto fotointerpretados y de campo y se ha hecho con SPSS una regresión lineal simple

$$y = B_0 + B_1x$$

De donde la ordenada en el origen y el coeficiente de regresión se calculan por mínimos cuadrados según:

$$\hat{B}_1 = \frac{\sum_{i \in S} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i \in S} (x_i - \bar{x})^2} = \frac{rS_y}{s_x},$$

$$\hat{B}_0 = \bar{y} - \hat{B}_1\bar{x},$$

Se calcula el Coeficiente de correlación de Pearson (R) y el de Determinación R^2

Así como la tabla ANOVA para la regresión. Con todo esto obtenemos la bondad de ajuste de la regresión para las dos variables.

Se calcula el Error Cuadrático Medio para la estimación de la regresión de la siguiente forma:

$$ECM(\hat{y}_{reg}) = E\{[\bar{y} + \hat{B}_1(\bar{x}_U - \bar{x}) - \bar{y}_U]^2\}$$

$$\approx V[\bar{d}]$$

$$= \left(1 - \frac{n}{N}\right) \frac{S_d^2}{n}.$$

Si usamos la relación $B_1 = RS_y/S_x$, podemos mostrar que:

$$\left(1 - \frac{n}{N}\right) \frac{S_d^2}{n} = \left(1 - \frac{n}{N}\right) \frac{1}{n} \sum_{i=1}^N \frac{(y_i - \bar{y}_U - B_1[x_i - \bar{x}_U])^2}{N-1}$$

$$= \left(1 - \frac{n}{N}\right) \frac{1}{n} S_y^2 (1 - R^2).$$

El error cuadrático medio es pequeño cuando:

1. N es grande.
2. n/N es grande
3. S_y es pequeño
4. La correlación R es cercana a -1 ó +1

De esta forma podemos calcular el Error Estándar al determinar la varianza de los residuos para la muestra. Sea:

$$e_i = y_i - (\hat{B}_0 + \hat{B}_1 x_i)$$

entonces,

$$EE(\hat{y}_{reg}) = \sqrt{\left(1 - \frac{n}{N}\right) \frac{s_e^2}{n}}$$

2º) A partir de aquí aplicaremos este proceso de cálculo a cada uno de los 11 estratos de la siguiente forma. Ejemplo para el estrato nº 7:

- **Análisis SPSS:**

Estadísticos descriptivos

	Media	Desviación típ.	N
CAPfoto	46,16	22,199	82
CAPcampo	39,63	20,902	82

R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio					Durbin-Watson
				Sig. del cambio en F	Cambio en R cuadrado	Cambio en F	gl1	gl2	
,606(a)	0,368	0,360	17,764	0,368	46,487	1	80	0,000	1,941

ANOVA					
Modelo	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Regresión	14.669,664	1	14.669,664	46,487	,000(a)
Residual	25.245,275	80	315,566		
Total	39.914,939	81			

Coeficientes(a)					
	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
	B	Error tip.	Beta		
(Constante)	20,641	4,226		4,885	0,000
CAPcampo	0,644	0,094	0,606	6,818	0,000

Conclusiones: La F calculada por la ANOVA nos dice que la regresión es altamente significativa. La t de student (altamente significativa) lo corrobora para el coeficiente de regresión y la ordenada en el origen. El Coeficiente de correlación de Pearson es altamente significativo. El test de Durbin Watson para la independencia de los residuos se aproxima a 2, por tanto es válido.

- Cálculo XLS:

$$\hat{y}_{reg} = 51,847 \quad S_e^2 = 315,56 \quad S_y^2 = 22,2 \quad EE_{Yreg} = 1,8297$$

Intervalo de confianza (IC 95%): 53,934 – 46,773

Valor del CAP automático medio del estrato 7º: 48,574

Conclusión: El CAP automático medio se encuentra dentro del intervalo de confianza, luego su media es un valor verosímil que no hay que modificar.

NOTACIONES:

\hat{y}_{reg} : Valor de la función para la media de CAP de fotointerpretado.

S_e^2 : Cuadrado medio del error de la ANOVA de la regresión

S_y^2 : Varianza del CAP fotointerpretado.

EE_{Yreg} : Error estándar de la regresión.

RESULTADOS FINALES.-

Procediendo de la misma manera que para el estrato 7º en todos los estratos se obtiene:

ZONAS Y ESTRATOS	Nº Recintos del Universo	Nº Recintos Muestra		Superficie (ha)		CAP Automático medio (%)	CAP Estimado (%)			CAP Prevalente Aplicado medio (%)
		Fotointerpretados	Visitados en Campo	Catastral Media	Admisible Media		CAP Estimado medio	Intervalo 95%		
								Max	Min	
PR_1,5-2_25_55	14.856	253	33	8,56	3,68	42,725	36,59	42,144	31,045	34,670
PR_1,5-2_>55-100	30.991	97	12	11,33	8,50	84,804	77,79	86,327	69,247	52,915
PA_1,5-2_25_100	7.748	651	69	11,09	6,12	49,556	56,39	60,155	52,633	49,556
PS_1,5-2_25_100	2.228	198	21	1,82	1,62	86,156	89,90	95,252	84,556	86,156
Zona Geográfica 1	55.823	1199	135	8,96	5,78	68,760	58,78	61,989	55,568	48,920
PR_2,2-2,5_25-55	6.962	297	34	6,51	2,50	40,014	36,00	39,438	32,563	40,014
PR_2,2-2,5_>55-100	11.231	94	9	6,42	5,44	86,569	67,20	75,292	59,105	51,636
PA_2,2-2,5_25-100	3.772	688	82	8,13	4,38	48,574	50,36	53,945	46,773	48,574
PS_2,2-2,5_25-100	653	192	25	2,51	2,07	81,616	74,53	83,776	65,282	81,616
Zona Geográfica 2	22.618	1271	150	6,62	4,26	65,750	52,33	55,188	49,462	48,410
PR_2,7_25-100	10.826	739	79	8,99	4,85	48,301	48,70	52,577	44,833	48,301
PA_2,7_25-100	2.909	467	57	5,00	1,69	33,916	42,91	47,227	38,601	33,916
PS_2,7_25-100	2.146	147	15	2,88	2,42	87,337	91,39	93,144	89,633	87,337
Zona Geográfica 3	15.881	1353	151	7,43	3,94	45,070	51,96	54,633	49,288	45,070
TOTAL	94.322	3.823	436							

Conclusiones:

Estrato 1º.- CAP auto (42,72) superior al intervalo 95%. Disminución. CAP prevalente = 34,67% Exceso en la disminución.

Estrato 2º.- CAP auto (84,80) dentro del intervalo. CAP prevalente = 52,91% Exceso en la disminución.

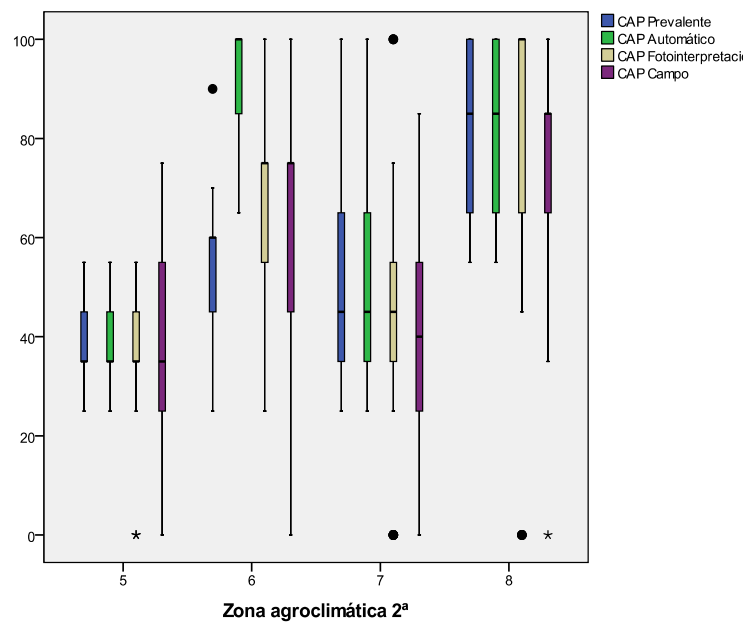
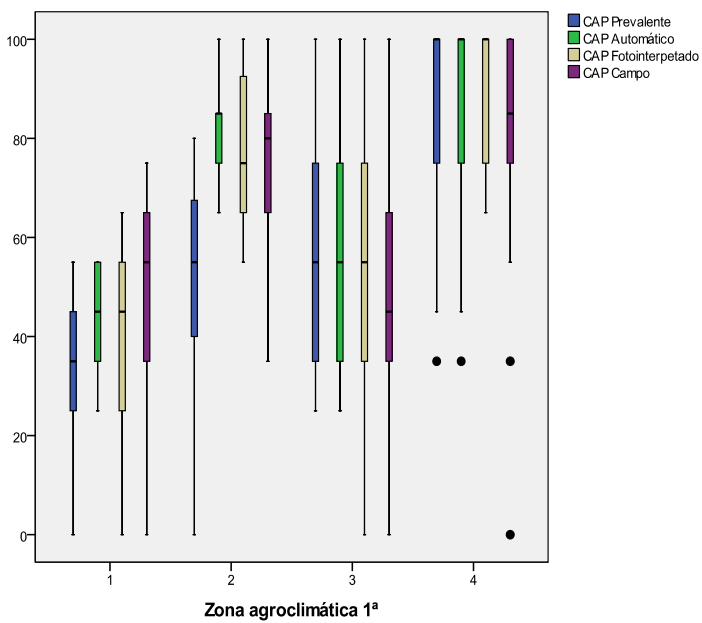
Estrato 3º, 4º: CAP auto en el intervalo de confianza, mantenimiento del CAP.

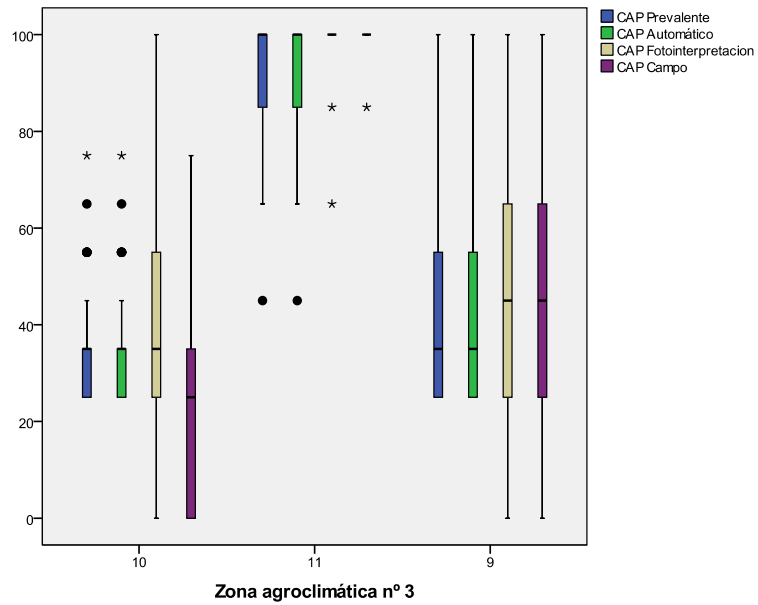
Estrato 5º: CAP auto fuera del intervalo. Disminuir.

Estrato 6º: Disminuir, el prevalente se ha disminuido en exceso.

Resto de estratos: CAP automático entre el intervalo. Correcto.

Diagrama de cajas para las diferentes CAP consideradas para las tres zonas agroclimáticas:





En los anteriores gráficos de cajas pueden verse las medias, distribuciones intercuartílicos y bigotes de la distribución de los CAP automáticos, fotointerpretados, campo y prevalente o modificado en 2015 para los 11 estratos considerados.

También pueden verse los datos aberrantes o atípicos, aquellos que están fuera del rango intercuartílico.