

CAI, Congreso Argentino de AgroInformática

Análisis de clustering en una región de La Pampa para clasificar sistemas ganaderos preponderantes y estimar emisiones de gases de efecto invernadero

Anabella Lozza^{1,2}, Yanina Bellini Saibene²¹ Consejo Nacional de Investigaciones Científicas y Técnicas² EEA INTA Anguil “Ing. Agr. Guillermo Covas”

Ruta Nac. N° , km 580

{lozza.anabella, bellini.yanina}@inta.gob.ar

Palabras Claves: Huella de carbono, clasificación, data mining, microrregión

10

1 Introducción

El ganado contribuye a la Emisión de Gases de Efecto Invernadero (EGEI) en un 14.5%, de manera directa (a partir de la fermentación entérica y el aprovechamiento del estiércol) e indirecta (actividades de producción y conversión de bosques en pastizales) [1]. La cadena de carne bovina en La Pampa está presente en toda la provincia, convirtiéndose en un gran contribuidor de GEI a escala provincial. Para realizar estimaciones regionales es necesario conocer los sistemas ganaderos preponderantes, los cuales son distintas combinaciones de actividades agropecuarias que más se utilizan en una zona [2]. El objetivo de este trabajo es identificar y describir sistemas ganaderos preponderantes en la microrregión 10 de La Pampa utilizando técnicas de agrupamiento para clasificarlos, y a partir de sus variables promedio, estimar las EGEI. La microrregión 10 se extiende en todo el sur de la provincia de este a oeste bordeando el Río Colorado [3]. No se incluyen en este estudio las actividades bajo riego.

2 Materiales y Métodos

A partir de los resultados reportados por [4] se calcularon 12 variables con los datos del 2014 del Registro Provincial Agropecuario de La Pampa, las cuales fueron estandarizadas (Tabla 1). Sobre esas variables, la proporción de cosecha gruesa y de fina no se utilizó ya que tenía el mismo valor (0%) para todos los casos. Esas actividades agrícolas no se realizan en esta zona, a menos que sea bajo riego. Se aplicaron dos técnicas de agrupamiento con el software R: k-means y k-medoids y buscando 2 y 3 grupos respectivamente. Esta cantidad de grupos se determinó teniendo en cuenta: a) las características productivas de la zona permiten presuponer que pueden existir, a lo

sumo, dos sistemas preponderantes [2], [3], b) se generaron 2 a 10 grupos (kmeans) y se graficó la suma de cuadrados dentro de cada grupo. Los gráficos generados sugieren tres grupos o menos y c) se utilizó la función de R pamk, que estima la cantidad de grupos, dando como resultado un sólo grupo para los tres dataset. Ante esta situación se decidió correr ambos algoritmos con 2 y 3 grupos. Se evaluó la estabilidad de los clusters generados con la función clusterboot() del paquete R fpc. Este algoritmo utiliza el coeficiente Jaccard, una medida de similitud entre conjuntos. Dicho coeficiente se comporta de la siguiente manera: valores $<0,6$ manifiestan grupos inestables; entre $0,6$ y $0,75$ indican un patrón en los datos, pero no hay certeza sobre qué casos deben agruparse; finalmente los valores $>0,85$ revelan grupos altamente estables. También se midió la Silueta que indica que tan bien está agrupado un caso y estima la distancia promedio entre los grupos. Un valor de silueta cercano a 1 indica que los casos están bien agrupados, un valor cercano a cero indica que los casos se encuentran en el límite entre un grupo y otro y un valor negativo indica que las observaciones probablemente están en el grupo incorrecto. Se realizaron 100 corridas con cada algoritmo-cantidad de clusters (400 corridas totales).

Tabla 1 Variables de proporciones e índices.

Variable y Fuente	Cálculo
Porcentaje de cultivos de cosecha fina ([4])	(Trigo+Avena+Cebada+Centeno+Otros) /Total de Hectáreas
Porcentaje de cultivos de cosecha gruesa ([4])	(Girasol+Maíz+Sorgo Granífero+Soja+Maní+Otros) /Total de Hectáreas
Porcentaje de forrajeras anuales ([4])	(Maíz+Mijo+Sorgo+Avena+Cebada+Centeno+otros verdeos de invierno+otros verdeos de verano) /Total de Hectáreas
Porcentaje de forrajeras perennes ([4])	(Alfalfa Pura +Alfalfa Consociada + Otras Leguminosas Puras + Otras Leguminosas Consociadas + Pasto Llorón +Digitaria + Panicum + Gramineas Puras) /Total de Hectáreas
Porcentaje de Caldenal ([4])	(Caldenal+Jarillar+Renoval)/Total de Hectáreas
IndiceOrientacionGanadera ([4])	Total Bovinos/(Total bovinos+ Total ovinos+ Total porcinos+ Total caprinos+ Total equinos) Total ovinos/(Total bovino s+ Total ovinos+ Total porcinos+ Total caprinos+ Total equinos) Total porcinos/(Total bovinos+ Total ovinos+ Total porcinos+ Total caprinos+ Total equinos) Total caprinos/(Total bovinos+ Total ovinos+ Total porcinos+ Total caprinos+ Total equinos) Total equinos/(Total bovinos+ Total ovinos+ Total porcinos+ Total caprinos+ Total equinos)
Indice Novillo ([4])	(Total Novillo de 1 a 2 años + Total Novillos >2años)/Total Cabezas
Índice de Cría ([4])	(Total Vacas + Total Vaquillonas de 1 a 2 años + Total Vaquillonas >2años) + Total Toros + Total Toritos)/Total Cabezas

La estimación de EGEI se hizo siguiendo las Directrices del IPCC (2006). Dicha metodología brinda la marcha de cálculo, basada en ecuaciones, que permiten estimar las emisiones de metano a partir de la fermentación entérica del ganado y emisiones de metano y óxido nitroso de la gestión del estiércol. Para bovinos se usó el nivel 2 del método, que minimiza el uso de datos por defecto, es decir, requiere una caracterización detallada de la población bovina, generando un buen nivel de desagregación a efectos de obtener factores de emisión ajustados a las circunstancias locales. En el resto de las especies ganaderas fue utilizado el nivel 1, que utiliza como input sólo datos de stock y regionalización con factores de emisión por defecto.

3 Resultados y Discusión

El mejor modelo se obtuvo con 2 grupos con k-medoids (más robusto al ruido y a los valores extremos, dos características de los datos analizados). Los grupos contienen 203 y 47 casos y son muy estables (Jaccard de 0,98 y 0,93 respectivamente) y presentan una adecuada distancia entre clusters (Silueta de 0,51 y 0,22). La tabla 2 presenta la media de los grupos obtenidos (n° de cabezas por cada categoría animal) y el cálculo de emisiones correspondientes (EGEI) en tnCO₂Eq. El Grupo 1 presenta un sistema de cría bovina sobre forraje natural con una superficie promedio de 4000 ha y el Grupo 2 un sistema de producción caprina predominantemente, sobre monte natural con una superficie promedio de 3700 ha. En este grupo la importancia del número de equinos supera las necesidades para el manejo ganadero. Los resultados de ambos grupos se condicen con el trabajo [2], tanto en los tipos de sistemas, las hectáreas involucradas y las particularidades del número de cabezas equinas.

En el Grupo 1 las emisiones son explicadas en un 98% por los bovinos, con más del 67% en la categoría vacas. En el grupo 2, sólo el 29% de las emisiones son bovinas mientras que un 65% corresponden a caprinos y equinos. La paridad en la proporción de emisiones de caprinos y equinos (33% y 34% respectivamente), se explica en el mayor peso de los equinos (510 kg más) respecto de los caprinos; aunque a nivel de “kilo vivo” los caprinos emiten más que los equinos por ser rumiantes.

Tabla 2. Características de los dos grupos y estimaciones de EGEI (tnCO₂Eq).

	Bov.	A	B	C	D	E	F	G	H	I	Ovi.	Por.	Cap	Equ
G1	344	51	13	28	162	19	19	3	7	42	6	1	2	5
EGEI	300,1	0,0	10,8	21,5	200,1	25,5	25,7	5,0	11,3	0,0	1,3	0,0	0,4	5,0
G2	17	3	0	0	11	0	0	0	0	3	13	1	93	22
EGEI	18	0	1	1	14	0	1	1	0	0	2,8	0,0	21,0	21,2

A: Tras < 1 año, B: Vaq. 1 a 2 años, C: Vaq.>2 años, D: Vacas, E: Nov.1 a 2 años, F: Nov. > 2 años, G: Tor 1 a 2 años, H: Toros > 2 años, I: Tros <1 año.

4 Conclusiones

Los grupos resultantes presentan un desbalance en la cantidad de casos y una silueta con valores medio y bajo que puede explicarse por las características agroecológicas de la zona de estudio, en los cuales los límites administrativos no conciben con el contexto natural que configura las actividades productivas, a pesar de esto; los métodos reconocen la existencia de un sistema productivo caprino característico del oeste pampeano, del cual la microrregión 10 sólo contiene una baja proporción. En este estudio, no se pudieron evaluar las diferencias de stock de carbono contenido en el suelo.

Al igual que en trabajos previos ([4]), la descripción sintética de ambos grupos permitió realizar un primer inventario de estimación de emisiones para el seguimiento de la huella de carbono en sistemas ganaderos. Nuevamente el método k-medoids presentó los mejores resultados.

Referencias

- [1] FAO, “Soluciones ganaderas para el cambio climático,” 2018.
- [2] J. Caviglia, H. Lorda, and J. Lemes, “Caracterización de las unidades de producción agropecuarias en la provincia de La Pampa,” *Boletín Divulg. técnica*, no. 99, pp. 10–15, 2010.
- [3] Ministerio de la Producción. “Ley 2.461,” no. 1, 2008.
- [4] A. Lozza, Y. Bellini Saibene, and H. Lorda, “Clasificación de Sistemas Productivos Preponderantes utilizando técnicas de agrupamiento para la estimación de emisiones de Gases de Efecto Invernadero,” *Contrib. al CAI 2017*, pp. 151–164, 2017. ISSN: 2525-0949. 7 de septiembre de 2017.