1-3-2023

Clustering climate and management practices to define environmental challenges affecting gastrointestinal parasitism in Katahdin sheep

Brian C. Arisman

Joan M. Burke

James L. M. Morgan

Ronald M. Lewis

Journal of Animal Science, 2023, **101**, 1–10
https://doi.org/10.1093/jas/skad002
Advance access publication 5 January 2023
Animal Genetics and Genomics

Clustering climate and management practices to define environmental challenges affecting gastrointestinal parasitism in Katahdin sheep

Brian C. Arisman,[†] Joan M. Burke,[‡] James L. M. Morgan,[§] and Ronald M. Lewis[†,1]

[†]Department of Animal Science, University of Nebraska-Lincoln, P.O. Box 830908, 68588, Lincoln, NE, USA
[‡]USDA, ARS, Dale Bumpers Small Farms Research Center, 6883 South State Highway 23, 72927, Booneville, AR, USA
[§]Round Mountain Consulting, 18235 Wildlife Road, 72701, Fayetteville, AR, USA

[1]Corresponding author: rlewis5@unl.edu

Abstract

Gastrointestinal nematodes (GIN) negatively affect the performance and well-being of sheep. Due to anthelmintic resistance, GIN are difficult to control leading producers to choose breeds that can exhibit resistance to parasitism. An example is Katahdin sheep. Katahdins are raised in various climates and management systems in the United States. These environmental factors can be combined to form eco-management groupings or clusters. We hypothesized that GIN challenge varies predictably based on the characteristics of these environmental clusters. Forty Katahdin producers from across the United States were surveyed for management information, with body weights (BW), fecal egg counts (FEC), and FAMACHA scores (FAM) available from 17 of the 40 flocks. The performance data included 3,426 lambs evaluated around 90 d of age. Management and climate data were combined into clusters using multiple correspondence and principal component (PC) analysis. Performance data were aligned with their corresponding cluster. Depending on the trait, eco-management cluster, birth-rearing type, sex, and, as a covariate, dam age, were fitted as systematic effects with ANOVA. Clusters also were formed based on climate or management data alone. When compared with fitting the eco-management clusters, they defined less variation in each of the traits based on Akaike and Bayesian information criterion, and adjusted r^2 values. To further examine variation defined by eco-management clusters, residuals from an ANOVA model excluding eco-management cluster were retained, and their correlation with PC loadings calculated. All PC loadings were included as potential independent variables and tested for significance using backward stepwise regression. The PC loadings with a correlation |≥0.49| explained significant variation in each trait and were included in the final models chosen; adjusted r^2 values for BW, FEC, and FAM were 0.90, 0.81, and 0.97, respectively. When analyzing GIN challenge, eco-management clusters corresponding with hotter temperatures and greater rainfall, and with pasture-born lambs, suffered greater parasitism. Conversely, the eco-management clusters with lambs turned out to pasture at older ages benefited from reduced parasitism. Through the formation of eco-management clusters, an environmental variable can be defined to study interactions of genotypes to their environment, providing a potentially useful tool for identifying parasite-resistant sheep.

Lay Summary

Katahdin sheep are a popular maternal hair breed that can exhibit resistance to gastrointestinal nematodes (GIN). Still, the consequences of GIN infection on performance levels, even in this breed, depend on the climatic and management conditions in which they are raised. Information on management practices in 40 U.S. Katahdin flocks was collected with an online survey. Climate data corresponding with these flock's locations were gathered from the National Weather Service. Using multivariate analysis to combine these data, nine distinct eco-management groupings or clusters were identified. These clusters differed in temperature, rainfall, grain supplementation, and the age at which the lambs were introduced to pasture. In 17 of these flocks, traits indicative of GIN parasitism—body weight, fecal egg count, and FAMACHA score—were measured in 90-d old Katahdin lambs. Eco-management cluster explained more variation in performance in all three traits than climate or management alone. Based on fecal egg counts, eco-management clusters corresponding with hotter temperatures and greater rainfall, and with pasture-born lambs, suffered greater parasitism. Conversely, eco-management clusters with lambs turned out to pasture at older ages benefited from reduced parasitism. Eco-management clusters provide a holistic approach to combine environmental factors that predispose lambs to parasitism.

Key words: eco-management clusters, fecal egg counts, gastrointestinal nematodes parasitism, multivariate analyses, sheep

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; BR, birth and rearing type; BW, body weight; CG, contemporary groups; FAM, FAMACHA score; FAMD, factor analysis on mixed data; FEC, fecal egg counts; GIN, gastrointestinal nematodes; HCPC, hierarchical clustering on principal components; MCA, multiple correspondence analysis; NSIP, National Sheep Improvement Program; PC, principal component; PCA, principal component analysis

Introduction

The sheep industry contributed $5.8 billion to the U.S. economy in 2017 (Shiflett, 2017). Given this industry's monetary importance, ensuring that management and selection practices are sustainable is vital. A priority in sheep production is genetic improvement of growth because of its economic importance (Borg et al., 2007). However, a major factor negatively affecting the growth and well-being of sheep managed in forage-based systems is gastrointestinal nematodes (GIN) infection (Coop and Holmes, 1996). Signs of GIN infection includes weight loss, decreased appetite, anemia, and even

Received August 17, 2022 Accepted January 3, 2023.

death (Bowman, 1999). In the European Union, the estimated economic cost of GIN resistant to anthelmintics in sheep was ₤42 million or roughly US$41 million (Charlier et al., 2020). The exact value of the economic impact in the United States, however, is unknown.

Since their introduction, the main way to control GIN infection was through the administration of anthelmintics. Anthelmintic resistance, however, has led to the increased popularity of breeds with higher levels of genetic resistance to parasites. Katahdin is a relatively prolific maternal composite hair breed comparable to other medium-sized maternal breeds in adult bodyweight (BW) and lamb growth rates (Ngere et al., 2018). This breed also potentially exhibits greater resilience and tolerance to GIN infection than other breeds (Burke and Miller, 2004). Resilience is defined as the productivity of an animal in the face of infection, whereas tolerance is defined as the net impact on performance at a specified level of infection (Bishop and Woolliams, 2014).

One approach to defining tolerance is to characterize the sensitivity of individuals or families to environments varying in GIN infection levels. The environment can be defined by differences among flocks in geographic locations, coinciding with distinct climates and management practices. Management differences among flocks, however, are often not considered due to the difficulty in quantifying husbandry practices. Such may be a misstep in precision breeding, where the environment is critical to characterizing complex interactions with the underlying biology of animals (Rexroad et al., 2019). Still, by combining information on climatic conditions and management systems, the extent of environmental challenge may be better classified. Henceforth, such categories will be referred to as eco-management clusters.

Katahdin producers from across the United States were surveyed for management information. These were combined with data from the National Weather Service (2021) that captured rainfall, snowfall, and temperatures for the different flock locations to form eco-management clusters. Performance data, specifically BW and measures indicative of GIN parasitism, were also available on a subset of these flocks. The objective of this study was to test if eco-management clusters explained variation in these economically relevant traits, and thereby was useful in the design of management and selection programs. Our specific focus was to test whether environmental conditions deemed coincident with a greater GIN challenge were predictive of parasitism in Katahdin lambs.

Materials and Methods

Animal handling and sample collection was conducted by the animal's owner and, therefore, did not require institutional animal care and use approval. Survey data were collected and stored in accordance with the University of Nebraska–Lincoln Institutional Review Board approval and standards.

Data collection

Over 3 yr (2017–2019), data were collected on 4,645 lambs from 142 sires and 1,855 dams spanning 17 Katahdin flocks from across the United States. These flocks were members of the National Sheep Improvement Program (NSIP; Notter, 1998). These data included pedigree, BW (kg), fecal egg counts (FEC, eggs/g), and FAMACHA scores (FAM), an indicator of anemia associated with GIN parasitism (Bath et al., 1996), which were collected in lambs at around 90-d of age (90.5 [SD 6.5] d). FAMACHA was subjectively assessed 1 to 5 with respect to the color of the membrane within the eyelid, where 1 was red or healthy and 5 was pale or anemic (Bath et al., 1996). Drenching events were also available. The FEC were determined from stool samples collected directly from the rectum using the modified McMaster technique (Whitlock, 1948).

Using the tseries package in R (Trapletti and Hornik, 2020), distributional properties of the data were evaluated using the Shapiro-Wilk normality test (Shapiro and Wilk, 1965) and Q-Q plots. Skewedness and kurtosis were assessed using the Jarque-Bera test (Jarque and Bera, 1980). The FEC were not normally distributed, and were transformed as log(FEC+25) (Ngere et al., 2018). The FAM were positively skewed. Several transformations were considered (log, square root, reciprocal) but none significantly improved the distributional properties of the trait. Therefore, for simplicity and to facilitate interpretations, a Gaussian distribution for FAM was assumed.

Only FEC and FAM from animals for which no anthelmintic was given 30 d prior to measurement were used. For those animals with a valid FEC, and with a corresponding BW and FAM record, data were further edited for outliers in FEC. Observations ± 4 SD from the mean transformed FEC were removed. Contemporary groups (CG) were assigned based on flock, birth year, management group, record date, and age slice. Age slice, as defined by NSIP, was in 35-d intervals from the birth of the first lamb in a lambing season. Any CG with fewer than 5 individuals was discarded. The final data consisted of 3,426 lambs with BW, FEC, and FAM (Table 1).

Climate data were obtained from the weather station closest to the geographical location of each of the 40 flocks completing the survey using the U.S. Climate Normals (1991–2020) published by the National Weather Service (2021). The weather data captured were yearly averages for rainfall, snowfall, and minimum and maximum temperature. Site elevation was also available. The means of the yearly averages over the 30-yr timeframe were used as the climatic values associated with each flock location.

A Qualtrics online survey was distributed to all active Katahdin breeders in NSIP. Of the 100 producers contacted, 40 participated. The general locations of those producers' flocks that responded to the survey are summarized in Figure 1. This map subdivided the United States into eco-regions based on temperature, precipitation, and elevation (Rowan et al., 2021). Of the 40 flocks completing the survey, 17 provided performance data. As shown in Figure 1, those flocks were representative geographically of the NSIP Katahdin flocks responding to the survey. Those 40 flocks themselves

Table 1. Means, SD, CV, minimum, and maximum for the 3,426 observations for body weight (kg), log transformed fecal egg count (log[eggs/g + 25]), and FAMACHA score

Trait	Mean	SD	CV	Min	Max
Body weight	22.3	6.4	28.9	7.3	49.0
Log transformed fecal egg count	3.2	0.7	24.1	1.4	4.8
FAMACHA score[1]	1.8	0.8	44.9	1	5

[1] FAMACHA was scored from 1, for red or healthy, to 5, for pale or anemic, with respect to the color of the membrane within the eyelid (Bath et al., 1996).

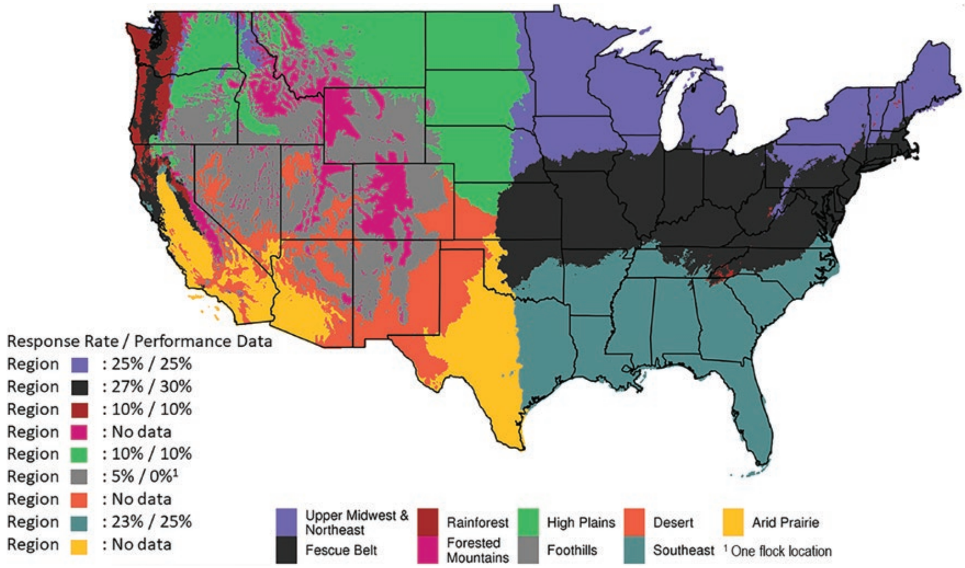

Figure 1. Distribution of 40 Katahdin flocks in the National Sheep Improvement Program that responded to the survey (response rate), and the corresponding 17 flocks providing performance information (performance data), relative to geographical region based on the eco-region map described in Rowan et al. (2021).

closely reflected the distribution of performance recording Katahdin flocks in NSIP.

The intent of the survey was to quantify differences in management. Practices of interest were grazing system, GIN impact, selection strategies to mitigate parasitism, feeding regime, and management strategies for the ewe flock and for lambs. The topics considered in the survey are provided in Supplementary Table S1. The producers' responses to the survey were coupled with the climate data associated with their flocks' geographic location.

Multivariate analysis

Using the combined management and climate data collated on each of the 40 flocks responding to the survey, a factor analysis on mixed data (FAMD) followed by hierarchical clustering on principal components (HCPC) was conducted. The FAMD with HCPC approach allows the grouping of flocks with similar management practices and climate into eco-management clusters that define an environment. With FAMD, qualitative and quantitative variables can be jointly analyzed. Principal component analysis (PCA) was applied to the quantitative data (climate and feed regime data) while multiple correspondence analysis (MCA) was applied to the qualitative data (obtained from the other survey topics described in the Supplementary Table S1). The MCA involved a pre-processing step in which categorical variables were transformed to a continuous scale with a mean and SD, allowing their integration as quantitative variables in the PCA (Aluja et al., 2018).

The FAMD analysis was run using the FactoMineR package in R (Lê et al., 2008). The HCPC was then utilized to form the eco-management clusters starting with a Ward's hierarchical classification (Lebart et al., 1995). Due to the potential nested structure of survey data, the Ward's hierarchical classification alone is not always optimal (Argüelles et al., 2014). Therefore, K-means clustering was applied using the centers of the classes obtained from the Ward's hierarchical classification (Pardo and Del Campo, 2007). After each iteration of the HCPC (Ward's hierarchical classification followed by K-means clustering), a chi-square test was used to determine which principal components (PC) significantly contributed to a cluster (α = 0.05; Nyairo et al., 2020). The resulting eco-management clusters were plotted using the visualization software Factoextra (Mundt and Kassambara, 2020). Significant PC loadings were used to define the climatic conditions and management practices that characterized the individual clusters. The PC loadings, otherwise known as scaled eigenvectors, are the covariances or correlations between the original variables and the unit-scaled components. In addition to the formation of the eco-management clusters, climate and management were clustered separately following the same strategy.

Model selection

The performance data collected from the 17 flocks were aligned with the corresponding eco-management cluster for each flock. Body weights were adjusted to 90-d equivalents as:

$$aY_i = (Y_i \times \ a_a)/a_{o_i}$$

where aY_i was the age adjusted BW of animal i, Y_i was the observed BW, a_a was the target age of recording of 90 d, and a_{o_i} was the observed age.

The importance of systematic (fixed) effects in defining variation in each trait was tested using ANOVA (Fox and Weisberg, 2019) with α = 0.05. One such fixed effect was birth and rearing type (BR). It was defined as a concatenated variable with six levels: single–single, twin–single, twin–twin, triplet plus–single, triplet plus–twin, and triplet plus–triplet plus, where the first adjective was the birth type, and the second adjective was the rearing type. A triplet plus indicated a triplet or higher birth or rearing type. The CG, lamb's sex (male, including castrates, or female), sex by BR interaction, and the linear and quadratic effect of its dam's age (d) were also considered. Residual error was included as the random effect.

For all three traits, CG, BR, and the linear and quadratic effects of dam age defined significant variation in performance based on stepwise regression (Venables and Ripley, 2002). For BW and FEC, the lamb's sex also was important but not its interaction with BR.

Cluster-type model fitting

The importance of cluster type (climate, management, or eco-management) in defining variation in each trait was tested by ANOVA (Fox and Weisberg, 2019) with $\alpha = 0.05$. Since the definition of CG included flock, they in part were confounded with cluster type. Therefore, in this analysis, cluster was included in lieu of CG in the models fitted. The other systematic effects included for a trait were left unchanged. Using R, the Akaike information criterion (AIC; Akaike, 1973), Bayesian information criterion (BIC; Schwarz, 1978), and adjusted r^2 values (Steel and Torrie, 1960) were obtained and used to compare models differing in cluster type.

Eco-management cluster validation

Residuals were obtained from fitting ANOVA models including all significant fixed effects defined for a trait excluding CG and eco-management cluster. Two different approaches were used to determine which PC loadings were significant. The first was based on estimating the correlations between the residuals and the 40 PC loadings. Stronger correlations would reveal which of the factors defining eco-management clusters were more strongly associated with a trait. The second approach involved including all 40 loadings as independent variables in a backward stepwise regression on residuals for each trait (Venables and Ripley, 2002). It was hypothesized that both approaches would help identify PC loadings that explained significant variation in the eco-management clusters.

Results

Multivariate analysis

Flocks were sampled from across the continental U.S. spanning a wide array of climates and 40 individual flock locations. Elevation varied from 2 to 790 m above sea level at these locations. Rainfall and snowfall ranged from 229 to 1600 mm and 25 to 1524 mm per yr, respectively. The average high temperature varied from 12 to 28 °C while the average low temperature varied from −1 to 12 °C. As with the climate data, management practices also differed. The greatest dissimilarities among farms were in feeding regime (from 0% to 100% pasture based) and sheep density per acre (from 2 to 28 ewes per acre). However, there were also pronounced similarities in management systems. Ninety-two percent of flocks lambed in the spring, and 91% of the producers used a combination of estimated breeding values and phenotypic performance records to select for GIN resistance.

Based on analysis of the combined climate and survey data, 9 eco-management clusters were identified. The PC 1 and PC 2 captured 11.2 % and 8.5 % of the variation, respectively (Table 2). The eco-management clusters that contained each of the 17 flocks with performance data are highlighted in Figure 2 with the predominant environmental descriptors associated with each cluster shown in Table 3. Six of the eco-management clusters were comprised of multiple flocks. The 3 flocks that were not located within a multiple-flock cluster (ellipse) formed separate eco-management clusters. All 9 of the clusters contained flocks with performance data. The differences in the environmental descriptors were primarily based on temperature, rainfall, grain supplementation while on pasture, and the age at which the lambs were turned out to pasture. Clusters defining similar combinations of climatic and management variables grouped together. For instance, eco-management clusters 1, 2, and 5 lined up vertically along the PC 1 axis; each of these clusters reflected warmer and wetter climates with lambs born on pasture. Clusters 4, 5, and 6 lined up horizontally along the PC 2 axis. These 3 clusters characterized eco-management clusters with moderate to high levels of grain supplementation.

Table 2. Contributions of the principal components to eco-management clusters

Principal component	Eigen value	Percent contribution	Cumulative percent contribution[1]
1	7.75	11.2	11.2
2	5.86	8.5	19.7
3	4.29	6.2	25.9
4	4.14	6.0	31.9
5	3.96	5.7	37.7
10	2.62	3.8	60.0
15	1.74	2.5	74.9
20	1.22	1.8	85.3
25	0.87	1.2	92.6
30	0.46	0.7	97.2
35	0.19	0.3	99.3
40	0.05	0.1	100.0

[1] Cumulative percent contribution is the sum of the percent contribution of its individual principal component and those that preceded it.

To compare the effect of eco-management cluster to other cluster delineations, clusters based upon climate and management data alone also were formed. Six multi-flock climate clusters were defined (Figure 3) in which PC 1 and PC 2 captured 48.4% and 24.3% of the variation, respectively. Any flock not in an ellipse formed a separate climate cluster. Alternatively, 4 multi-flock management clusters were formed (Figure 4), in which PC 1 and PC 2 captured 10.8% and 8.4% of the variation, respectively. As with eco-management clusters, the number of separate clusters was greater than the number of multiple-flock clusters. However, there were more flocks combined into multiple-flock clusters than standalone clusters. Flocks with performance data were assigned to 9 climate clusters and 6 management clusters.

Cluster-type model fitting

The results of fitting each clustering type (climate, management, or eco-management) in the predictive model for each trait are summarized in Table 4. The AIC, BIC, and adjusted r^2 values are provided. For BW and FAM, including any of the clustering types defined variation in performance ($P < 0.05$). However, the "best model" was the one with the eco-management cluster. For FEC, only the model in which eco-management cluster was fitted explained significant variation in performance levels ($P < 0.05$). The solutions (least squares means) by eco-management cluster are given for each trait in Table 3.

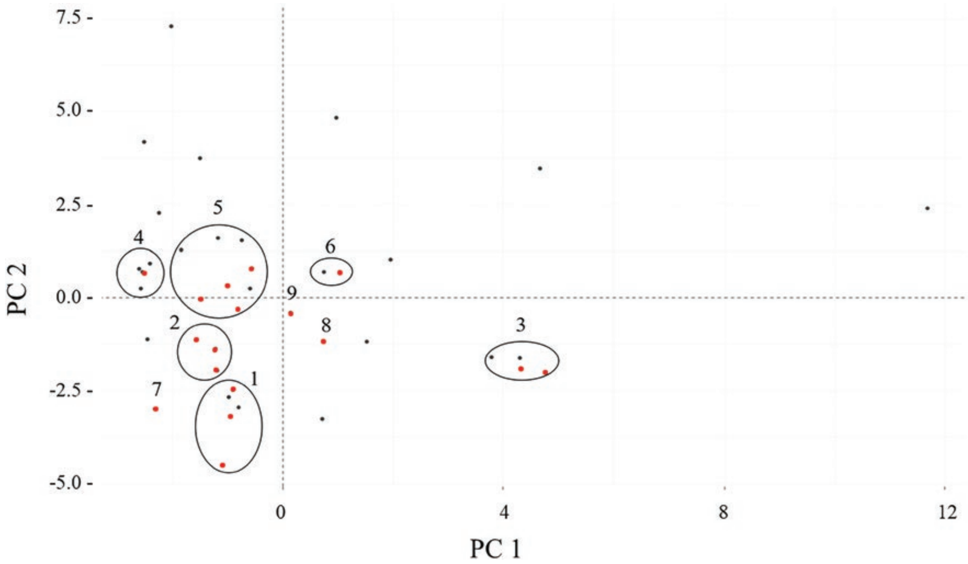

Figure 2. Distributions of flocks with performance data (red) within the eco-management clusters formed from the climate and management data. Numbered ellipses identify eco-management clusters that contained flocks with performance data, which are described in Table 3.

Table 3. The environmental delineators for the eco-management clusters (CL) with the corresponding least square means (standard error) for body weight (BW, kg), log transformed fecal egg count (FEC, log[eggs/g + 25]), and FAMACHA score (FAM)

CL No.[1]	CL Descriptors[2]	No.	BW	FEC	FAM[3]
1	Hotter, low grain supplementation, wet, pasture born	31	4.76 (0.97)[d]	0.50 (0.01)[a]	0.66 (0.16)[bc]
2	Warm, no grain supplementation, wetter, pasture born	531	11.65 (0.41)[a]	0.28 (0.05)[b]	1.18 (0.06)[a]
3	Hot, low grain supplementation, wet, pasture born	315	−0.96 (0.31)[e]	0.18 (0.04)[c]	0.39 (0.05)[d]
4	Cool, higher grain supplementation, mild, pasture born	122	8.80 (0.40)[bc]	0.04 (0.02)[d]	1.13 (0.06)[a]
5	Hottest, higher grain supplementation, wettest, pasture born	591	7.36 (0.26)[c]	−0.02 (0.01)[d]	0.18 (0.04)[e]
6	Coldest, moderate grain supplementation, mild, pasture born	460	7.07 (0.27)[c]	−0.05 (0.03)[d]	0.17 (0.04)[e]
7	Warm, no grain supplementation, driest, 30-d pasture turn out	377	7.49 (0.33)[c]	−0.14 (0.04)[e]	0.24 (0.05)[de]
8	Colder, moderate grain supplementation, mild, 45-d pasture turn out	291	4.76 (0.97)[d]	−0.24 (0.04)[f]	0.58 (0.07)[c]
9	Colder, no grain supplementation, drier, 30-d pasture turn out	259	11.65 (0.41)[a]	−0.40 (0.04)[g]	0.78 (0.06)[b]

Means within a column with different superscripts differ ($P < 0.05$).
[1] Clusters ordered by least square means for log transformed FEC.
[2] Descriptors were temperature, grain supplementation, rainfall, and whether pasture born or age at turnout to pasture.
[3] FAMACHA was scored from 1, for red or healthy, to 5, for pale or anemic, with respect to the color of the membrane within the eyelid (Bath et al., 1996).

Cluster validation

To further explore the variation explained by the eco-management clusters, residuals were obtained from the models fitted for the three traits. The correlation between the PC and the residuals are provided in Table 5. Only those PC loadings that had a correlation greater or equal to the absolute value of 0.49 defined variation ($P < 0.05$) in a trait and were selected to be retained in a model. The final regression models selected using backward stepwise regression, along with adjusted r^2 values, are given in Table 6. The descriptions of the PC loadings that described significant variation in a trait are displayed in Table 7.

Discussion

The combination of management practices with climatic variables into eco-management clusters improved the description of BW and severity of GIN parasitism (FEC, FAM). Formerly, environmental differences had been delineated based on climatic conditions alone using K-means clustering, as recently described in beef cattle (Rowan et al., 2021). This approach was deemed beneficial in defining the environmental challenges experienced by range animals that were constantly exposed to the climatic conditions. However, differences in management practices also impact the environment in which animals are raised, and in some case can account for more variation in phenotype (Rexroad et al., 2019). For instance, in small ruminant production systems, some flocks are raised in total confinement or dry lotted; individuals therefore are not continuously exposed to environmental challenges such as GIN infection.

In Intermountain West sheep populations, the diversity of management practices was comprehensively quantified using a survey (Stewart et al., 2020). The survey captured information on the management strategies currently in place, and the ways those management practices were adapted to address GIN infection. However, those enterprises were located over a limited geographical range (Stewart et al., 2020). In the current study, management practices of Katahdin sheep flocks spread across the United States were surveyed. This allowed for more

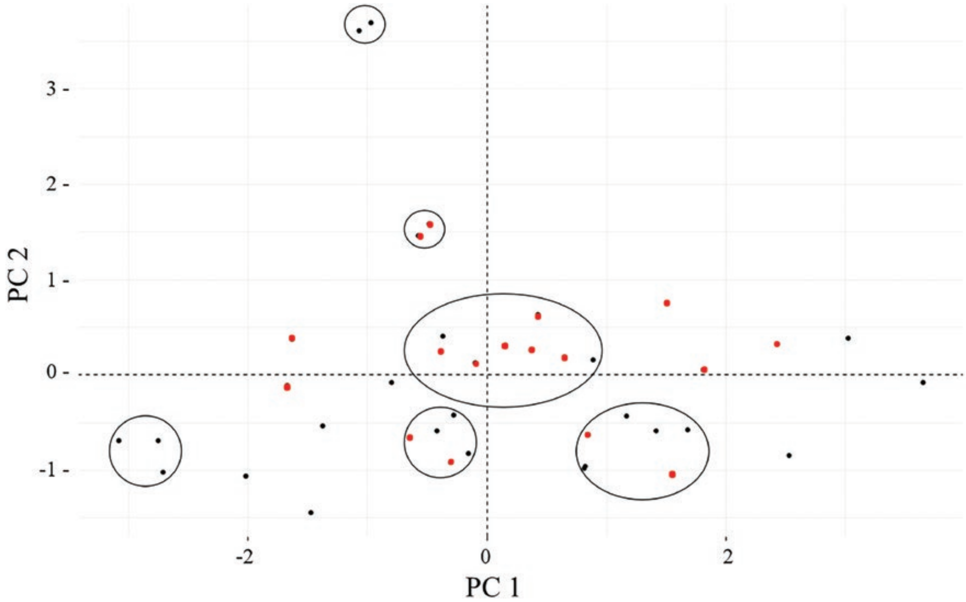

Figure 3. Distributions of flocks with performance data (red) within climate clusters formed by using strictly climate data. Each ellipse identifies flocks within different climate clusters.

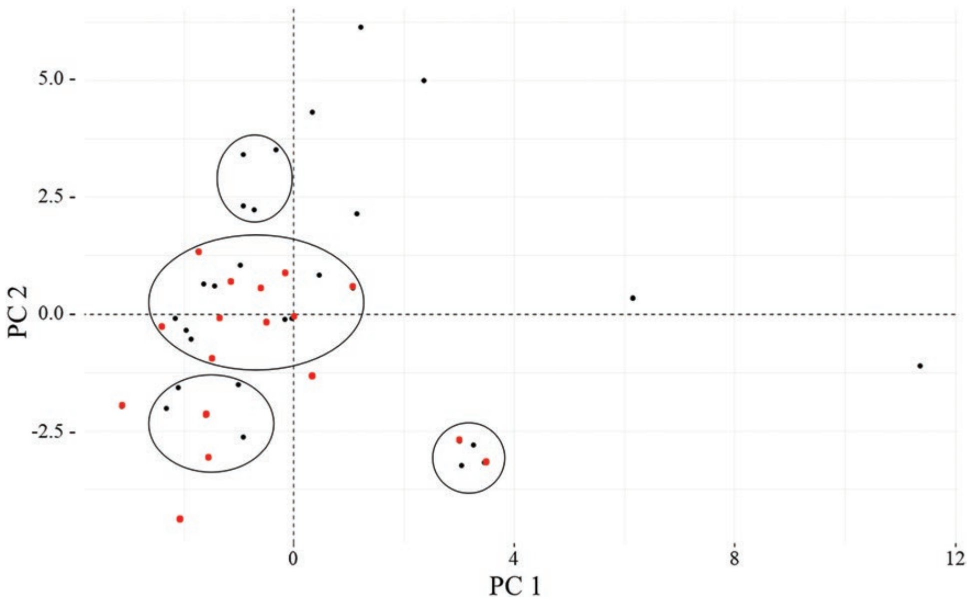

Figure 4. Distributions of flocks with performance data (red) within management clusters formed by using strictly management data. Each ellipse identifies flocks within different management clusters.

diverse climates and management practices to be considered. An outcome was a more precise definition of the environment to which animals were exposed across the varying regions.

With the Katahdin flocks spread across the eastern half of the United States, with a few located in the northwest, climatic differences were anticipated. Although some management practices were common across enterprises, such as pasture rotation, others differed, such as the extent of supplementation. These differences may be due to producers tailoring their management practices to better match their animals' phenotypes to their environments (Rexroad et al., 2019). A value of forming eco-management clusters was to capture those differences among operations, particularly if they were located within the same climatic region. Still, as management practices may change over time, there likely is a need to reclassify eco-management clusters periodically.

The idea of including an environmental cluster in predictive models is not new. However, previously, it has been studied using environments solely based on climatic variables (McManus et al., 2021; Rowan et al., 2021). Although the methods used to define climatic clusters differed, their delineations of climatic regions aligned with those observed in this study. Such was the case even with a different environmental variable—McManus et al. (2021) used the temperature humidity index—and another pastoral species (beef cattle). The consistency in our findings suggest that the approaches adopted to form clusters of at least the climate variables were effective.

Table 4. Fit of cluster-type to define variation in body weight, log transformed fecal egg count, and FAMACHA score

		Statistics[1]			
Trait	**Cluster-type**	**AIC**	**BIC**	**Adj. r^2**	***P*-value**
Body weight	Climate	−1398.3	−1367.9	0.26	0.043
	Management	−13445.5	−1714.2	0.33	0.048
	Eco-management	0	0	0	0.034
Log transformed fecal egg count	Climate	−168.1	−138.1	0.06	0.213
	Management	−156.5	−132.6	0.05	0.654
	Eco-management	0	0	0	0.021
FAMACHA score	Climate	−127.4	−97.5	0.04	0.008
	Management	−184.9	−160.9	0.06	0.011
	Eco-management	0	0	0	0.002

[1] Akaike information criterion (AIC), Bayesian information criterion (BIC), and adjusted r^2 values are given as deviations from "best" fit model highlighted in bold. *P*-values for cluster-type from ANOVA are also provided.

Table 5. Correlations between principal component (PC) loadings and residuals for body weight (BW), log transformed fecal egg count, and FAMACHA score (FAM)

Trait	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
BW Residual	−0.17	**−0.86**[1]	−0.43	**−0.51**	−0.37	−0.02	0.47	0.18	0.11	0.04
FEC Residual	0.03	0.23	0.01	0.12	0.09	−0.13	**−0.53**	−0.34	0.46	−0.25
FAM Residual	−0.42	**−0.59**	−0.29	−0.23	**−0.47**	−0.43	**0.54**	−0.13	0.22	−0.36
	PC11	PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19	PC20
BW Residual	−0.27	−0.03	**0.49**	0.02	0.21	0.28	0.06	0.36	0.32	0.31
FEC Residual	**0.85**	0.07	**−0.62**	0.01	−0.15	0.00	−0.30	**−0.55**	−0.39	−0.10
FAM Residual	0.06	−0.06	0.12	**0.51**	−0.08	**0.68**	−0.29	0.07	0.05	−0.37
	PC21	PC22	PC23	PC24	PC25	PC26	PC27	PC28	PC29	PC30
BW Residual	**0.51**	0.37	0.25	0.18	**0.54**	0.33	0.45	0.07	0.25	−0.40
FEC Residual	−0.18	−0.46	**−0.60**	−0.08	−0.02	0.44	**−0.59**	0.20	**−0.65**	0.16
FAM Residual	0.16	0.46	0.20	0.36	−0.03	0.27	0.20	0.36	0.22	−0.26
	PC31	PC32	PC33	PC34	PC35	PC36	PC37	PC38	PC39	PC40
BW Residual	0.27	−0.26	0.11	−0.07	−0.17	0.08	−0.45	−0.41	−0.11	0.00
FEC Residual	−0.01	−0.05	0.24	0.47	−0.29	0.12	−0.11	0.41	−0.01	0.33
FAM Residual	−0.49	−0.21	0.61	0.12	−0.05	0.46	−0.36	0.01	−0.42	0.42

[1] PC loadings defining variation in a trait are highlighted in bold ($P < 0.05$).

Table 6. Final model selected from regression of body weight (BW, kg), log transformed fecal egg count (FEC, log[eggs/g + 25]), and FAMACHA score (FAM) residuals on principal components (PC) loadings, and corresponding adjusted r^2 value

Variable	Model	Adjusted r^2
BW	−0.799×PC2 − 0.129×PC4 + 0.868×PC13 + 0.694×PC21 + 0.841×PC25 − 1.477	0.90
FEC	0.012×PC7 + 0.109×PC11 + 0.002×PC13 − 0.037×PC18 + 0.059×PC23 − 0.073×PC27 − 0.062×PC29 + 0.027	0.81
FAM	−0.223×PC2 − 0.051×PC5 − 0.106×PC7 + 0.108×PC14 − 0.121×PC16 − 0.089×PC31 + 0.014×PC33 + 0.878	0.97

When management is considered independent of the environment, the management clusters were not useful in defining variation in the 3 traits evaluated. Such was likely to be expected. Management practices are often employed to overcome environmental stressors. For instance, ewes housed during the lambing season are not exposed to the prevailing climatic conditions. It was not surprising, therefore, that less variation was explained with the management clusters alone when compared to the other clustering types.

Of the combined climatic and management factors considered, temperature, rainfall, grain supplementation, and the age at which the lambs were turned out to pasture were key determinants of GIN infection levels. Based on FEC, eco-management clusters corresponding with hotter temperatures and greater rainfall, and pasture lambing, suffered from a greater parasite challenge. Conversely, the eco-management clusters with confinement or shed lambing, with lambs turned out to pasture at older ages, benefited from a reduced challenge.

Table 7. Dominant three factors (climatic conditions or management practices) associated with principal components (PC) defining variation in residual body weight (BW), log transformed fecal egg count (FEC). and FAMACHA score (FAM)

PC	Significant Variables	BW[1]	FEC	FAM
2	Age at which the lambs are out to pasture, active pasture grazed April to June, lambing location	−0.86		−0.60
4	Lambs by management groups, castration of males, elevation	−0.51		
5	Grazing system, lambing season, priority of selection against GIN in replacement ewes			−0.50
7	Cover crops grazed October to December, lambing drenching protocol, ewe drenching protocol		−0.53	0.54
11	Active pasture grazed April to June, grain fed April to June, castration of males		0.85	
13	Lambing season, lamb Mortality due to GIN, age at which the lambs are turned out to pasture	0.49	−0.62	
14	Were the sheep grazed with other livestock, elevation, active pasture grazed July to September			0.51
16	Decrease in lamb performance, rainfall, grain fed January to March			0.68
18	October to December actively grazed pasture, thermal min, lambing season		−0.55	
21	Harvested forage fed April to June, harvested forage fed July to September, deworming rate of replacement animals	0.51		
23	Lamb drenching protocol, lambing location, snowfall		−0.60	
25	Priority of GIN selection in rams, lambs by management groups, grazing system	0.54		
27	Were the sheep grazed with other livestock, decrease in lamb performance due to GIN, deworming rate of replacement animals		−0.59	
29	Cover crops grazed April to June, selection tool utilized against GIN, priority of GIN selection in adult ewes		−0.65	
31	Grazing system, lamb mortality due to GIN, cover crops grazed April to June			−0.50
33	Cover crops grazed January to March, cover crops grazed July to September, castration of males			0.61

[1] Correlation coefficients for each PC defining variation in a trait ($P < 0.05$).

These clusters were also characterized by lower levels of grain supplementation of lambs at pasture.

The importance of temperature and precipitation on GIN was expected, since humidity accelerates the development and reproduction of the nematode larvae (Manfredi, 2006). The major GIN of concern in the U.S. sheep population is *Haemonchus contortus* (Howell et al., 2008; Stewart et al., 2020). When using FEC as a primary indicator of GIN burden, *H. contortus* is expected to be the predominant nematode species due to its prevalence and high reproductive capacity in most of the United States (Bowman, 1999).

O'Connor et al. (2006) summarized the necessary temperature ranges for infection for prominent GIN. For *H. contortus*, the range was 11 to 40 °C; for *Trichostrongylus* spp. and *Teladorsagia* spp. those ranges were 6 to 39 °C and 1 to 35 °C, respectively. Among the eco-management clusters on which FEC data were available, nearly all met the temperature requirements for *H. contortus* development for at least 100 d each year. McCulloch et al. (1984) indicated that *H. contortus* thrive in wetter, hotter climates. The hotter and wetter climates surveyed had a greater FEC, probably due to greater *H. contortus* infection obtained by potentially grazing infected pasture for longer. Conversely, the colder and drier climates surveyed had a lower FEC, probably due to the wide variation in daily temperatures and lower relative humidity (Smith, 1990). However, *Trichostrongylus* and *Teladorsagia,* which may predominant under such conditions, produce fewer eggs than *H. contortus* potentially resulting in lower FEC. Although the distribution of GIN species present in the fecal samples was not quantified in this study, variability in the GIN population may explain differences observed in FEC among eco-management clusters.

The ages at which lambs were turned out to pasture influenced the age and period they were exposed to GIN infection. Lambs born on pasture expressed a higher GIN burden. Lambs often begin to minimally graze at 20 d of age, initiating their exposure to GIN. Early naïve exposure coupled with an immature immune system response to GIN may explain a greater GIN load in lambs turned out to pasture at younger as compared to older ages. Also, with grazing during the periparturient period, more susceptible ewes and lambs increase pasture infestation resulting in a greater environmental challenge (Notter et al., 2017).

The level of grain supplementation while on pasture varied among flocks. However, according to Wood et al. (2018), the higher the grain supplementation the lower the FEC due to an improved nutritional status. An increase in, particularly, protein intake increases tolerance to GIN infection (Coop and Holmes, 1996; Steel, 2003). Furthermore, if supplement replaces forage intake, infection level also presumably is reduced. Such benefits were not observed in the current study. Information on forage quality and the type or amount of supplement was not sought in the survey for brevity. However, sheep could still consume substantial amounts of infected forage even while being supplemented. Although suppositions, such possibilities may help explain the ambiguous relationship of grain supplementation with FEC and FAM observed.

The eco-management clusters were ranked by their least squares means for BW, FEC, and FAM. FAMACHA is a predictor of anemia, which is anticipated with infection by *H.*

contortus although not with other GIN. The clusters with a higher mean for FEC had a higher mean for FAM suggesting a *H. contortus* challenge. However, their correspondence (rankings) was not absolute. This again may reflect differences in GIN species contributing to the infection. There was also little concordance between parasite infection levels (FEC, FAM) and BW. This may reflect genetic resistance to GIN and other management interventions implemented (e.g., supplementation practices, grazing systems, drenching events).

The environment was defined by a flock's cluster, which was derived using PCA. Some PC were highly correlated—greater or equal to the absolute value of 0.5—with the residuals for a trait. McNicol et al. (1993) reported similar observations. When such was the case, the values of the climatic variables, and levels of the management variables, sensibly corresponded with the expression of that trait. For BW, the correlated PC were mainly related to feeding practices, grazing system, castration of males, and housing of the lambs (i.e., competition level), all of which influence growth rates in lambs. For FEC, the correlated PC were mainly related to selection pressure for GIN, incidence of drenching, weather factors (temperature, rainfall, and snowfall), and whether lambs were grazed with other livestock, all of which influence GIN infection rates in lambs. For FAM, like FEC, the correlated PC were mainly related to selection pressure for GIN, grazing regime, weather factors, and whether lambs were grazed with other livestock. Again, such factors are associated with GIN infection rates in lambs. When the PC were used to predict residuals for each trait, those that were more highly correlated were unsurprisingly selected to include in the models providing best fit (high adjusted r^2 values).

The formation and use of the eco-management clusters to define the environment animals inhabit explained variation in BW, FEC, and FAM. This is valuable in several ways. In genetic evaluation, the clusters could be used as the environmental variable when evaluating a genotype by environment interaction. Assuming sufficiently strong familial relationships across clusters, a genotype by environment interaction could be fitted to identify superior individuals or families able to tolerate variable environmental challenges, including GIN infection.

Eco-management clusters could be further utilized in precision breeding strategies. Although not yet widely implemented in the U.S. sheep industry (McMillan et al., 2022), genomic selection strategies can be used to tailor selection to fit a specific environment, even those created by management (Rexroad et al., 2019). Another utility would be to identify replacement individuals from a similar climate and management background. Such animals may be better adapted to the new production environment and, moving forward, optimize performance based on the environment and management inputs. Importantly, the formation of eco-management clusters can be implemented in any pastural-based species, such as beef cattle, where more performance data are currently available. The successful formation of eco-management clusters, therefore, could contribute to implementing precision breeding practices by producers.

Supplementary Data

Supplementary data are available at *Journal of Animal Science* online.

Acknowledgements

The authors wish to thank the National Sheep Improvement Program, and its member Katahdin sheep producers, for their contributions to this research. This project was supported by the Organic Agriculture Research and Extension Initiative (2016-51300-25723), and by the Agriculture and Food Research Initiative Competitive (2022-67015-36073), from the United States Department of Agriculture (USDA) National Institute of Food and Agriculture. The USDA is an equal opportunity provider and employer. The mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.

Conflict of interest statement

The authors declare no real or perceived conflict of interest.

Literature Cited

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. In: B. N. Petrov and F. Csaki, editor. *Internatl. Symp. Information Theory*. Springer-Verlag, p. 267–281. doi:10.1007/978-1-4612-1694-0_15

Aluja, T., A. Morineau, and G. Sanchez. 2018. Principal component analysis for data science. Available from https://pca4ds.github.io

Argüelles, M., C. Benavides, and I. Fernández. 2014. A new approach to the identification of regional clusters: hierarchical clustering on principal components. *Appl. Econ*. 46:2511–2519. doi:10.1080/00036846.2014.904491

Bath, G. F., F. S. Malan, and J. A. Van Wyk. 1996. The "FAMACHA" Ovine Anemia Guide to assist with the control of haemonchosis. In: *Proc. 7th Ann. Congr. Livest. Health and Prod*. Port Elizabeth, South Africa: Group of the South African Vet. Assoc. p. 5.

Bishop, S. C., and J. A. Woolliams. 2014. Genomics and disease resistance studies in livestock. *Livest. Sci*. 166:190–198. doi:10.1016/j.livsci.2014.04.034

Borg, R. C., D. R. Notter, L. A. Kuehn, and R. W. Kott. 2007. Breeding objectives for Targhee sheep. *J. Anim. Sci*. 85:2815–2829. doi:10.2527/jas.2006-064

Bowman, D. D. 1999. *Georgis' parasitology for veterinarians*. 7th ed. Philadelphia, PA: Saunders.

Burke, J. M., and J. E. Miller. 2004. Relative resistance to gastrointestinal nematode parasites in Dorper, Katahdin, and St. Croix lambs under conditions encountered in the southeastern region of the United States. *Small Rumin. Res*. 54:43–51. doi:10.1016/j.smallrumres.2003.10.009

Charlier, J., L. Rinaldi, V. Musella, H. W. Ploeger, C. Chartier, H. R. Vineer, B. Hinney, G. von Samson-Himmelstjerna, B. Băcescu, M. Mickiewicz, et al. 2020. Initial assessment of the economic burden of major parasitic helminth infections to the ruminant livestock industry in Europe. *Prev. Vet. Med*. 182:105103. doi:10.1016/j.prevetmed.2020.105103

Coop, R. L., and P. H. Holmes. 1996. Nutrition and parasite interaction. *Int. J. Parasitol*. 26:951–962. doi:10.1016/s0020-7519(96)80070-1

Fox J., and S. Weisberg. 2019. *An R companion to applied regression*. Thousand Oaks, CA: SAGE Publications, Inc.

Howell, S. B., J. M. Burke, J. E. Miller, T. H. Terrill, E. Valencia, M. J. Williams, L. H. Williamson, A. M. Zajac, and R. M. Kaplan. 2008. Prevalence of anthelmintic resistance on sheep and goat farms in the southeastern United States. *J. Am. Vet. Med. Assoc*. 23:1913–1919. doi:10.2460/javma.233.12.1913

Jarque, C. M., and A. K. Bera. 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Econ. Letters* 6:255–259. doi:10.1016/0165-1765(80)900024-5

Lê, S., J. Josse, and F. Husson. 2008. FactoMineR: An R Package for Multivariate Analysis. *J. Stat. Softw.* 25:1–18. doi:10.18637/jss.v025.i01

Lebart, L., A. Morineau, and M. Piron. 1995. *Statistique exploratoire multidimensionnelle*. Paris: Dunod.

Manfredi, M. T. 2006. Biology of gastrointestinal nematodes of ruminants. *Parasitology* 48:397–401.

McCulloch, B., H. G. Kühn, and R. R. Dalbock. 1984. The relation of climate and topography to worm egg counts of gastro-intestinal nematodes of sheep in the Eastern Cape. *Onderstepoort J. Vet. Res.* 5:223–238.

McManus, C., P. M. Hermuche, R. S. Paiva, R. F. Guimarães, O. A. Carvalho, Jr., and H. D. Blackburn. 2021. Gene bank collection strategies based upon geographic and environmental indicators for beef breeds in the United States of America. *Livest. Sci.* 254:104766. doi:10.1016/j.livsci.2021.104766

McMillan, A. J., D. J. Brown, J. M. Burke, J. L. Morgan, and R. M. Lewis. 2022. Crossvalidation of single-step genetic evaluation in US Katahdin sheep. In: *Proc. 12th World Cong. Genet. Appl. Livest. Prod*. Wageningen Academics Publishers. Rotterdam, Netherlands.

McNicol, J., D. Hirst, and R. Kempton. 1993. *Graphical methods for multivariate data. 2.1*. Edinburgh Scotland: Scottish Agric. Stat. Service.

Mundt, F., and A. Kassambara. 2020. Factoextra: extract and visualize the results of multivariate data analysis. Available from https://rpkgs.datanovia.com/factoextra/index.html

National Weather Service. 2021. U.S. climate normals. Available from https://www.ncei.noaa.gov/products/land-based-station/us-climate-normals

Ngere, L., J. M. Burke, J. L. M. Morgan, J. E. Miller, and D. R. Notter. 2018. Genetic parameters for fecal egg counts and their relationship with body weights in Katahdin lambs. *J. Anim. Sci.* 96:1590–1599. doi:10.1093/jas/sky064

Notter, D. R. 1998. The U.S. National Sheep Improvement Program: across-flock genetic evaluations and new trait development. *J. Anim. Sci.* 76:2324–2330. doi:10.2527/1998.7692324x

Notter, D. R., J. M. Burke, J. E. Miller, and J. L. M. Morgan. 2017. Factors affecting fecal egg counts in periparturient Katahdin ewes and their lambs. *J. Anim. Sci.* 95:103–112. doi:10.2527/jas2016.0955

Nyairo, R., T. Machimura, and T. Matsui. 2020. A combined analysis of sociological and farm management factors affecting household livelihood vulnerability to climate change in rural Burundi. *Sustainability* 12:4296. doi:10.3390/su12104296

O'Connor, L. J., S. W. Walkden-Brown, and L. P. Kahn. 2006. Review: ecology of the free-living stages of major trichostrongylid parasites of sheep. *Vet. Parasitol.* 142:1–15. doi:10.1016/j.vetpar.2006.08.035

Pardo, C. E., and P. C. Del Campo. 2007. Combination of factorial methods and cluster analysis in R: the FactoClass package. *Colombian J. Stat*. 30:231–245. doi:10.15446/rce

Rexroad, C., J. Vallet, L. K. Matukumalli, J. Reecy, D. Bickhart, H. Blackburn, M. Boggess, H. Cheng, A. Clutter, N. Cockett, et al. 2019. Genome to phenome: improving animal health, production, and well-being – a new USDA blueprint for animal genome research 2018–2027. *Front. Genet.* 10:4–10. doi:10.3389/fgene.2019.00327

Rowan, T. N., H. J. Durbin, C. M. Seabury, R. D. Schnabel, and J. E. Decker. 2021. Powerful detection of polygenic selection and evidence of environmental adaptation in US beef cattle. *PLoS Genet.* 17:e1009652. doi:10.1371/journal.pgen.1009652

Schwarz, G. E. 1978. Estimating the dimension of a model. *Ann. Stat.* 6:461–464. doi:10.1214/aos/1176344136

Shapiro, S. S., and M. B. Wilke. 1965. An analysis of variance test of normality (complete samples). *Biometrika* 52:591–611. doi:10.1093/biomet/52.3-4.591

Shiflett, J. S. 2017. U.S. Sheep industry economic impact study. *ASIA*. Available from https://d1cqrq366w3ike.cloudfront.net/http/DOCUMENT/SheepUSA/ASI%202016%20Sheep%20Industry%20Economic%20 Impact%20Analysis.pdf

Smith, G. 1990. The population biology of the free-living phase of *Haemonchus contortus*. *Parasitology* 101:309–316. doi:10.1017/s003118200006337x

Steel, J. W. 2003. Effects of protein supplementation of young sheep on resistance development and resilience to parasite nematode. *Aust. J. Exp. Agric*. 43:1469–1476. doi:10.1071/ea03004

Steel, R. G. D., and J. H. Torrie. 1960. *Principles and procedures of statistics with special reference to the biological sciences*. New York, NY: McGraw-Hill Book Company Inc.

Stewart, W. C., D. M. Scott, S. B. Howell, R. M. Kaplan, B. L. Roeder, and T. W. Murphy. 2020. Anthelmintic resistance in gastrointestinal nematodes and associated management factors in Intermountain West sheep flocks. *Sheep Goat Res. J.* 35:30–37. Available from https://www.sheepusa.org/wp-content/uploads/2021/01/SGRJ_V35_30-37_Stewart_11-20.pdf

Trapletti, A., and K. Hornik. 2020. Tseries: time series analysis and computational finance. Available from https://CRAN.R-project.org/package=tseries

Venables, W. N., and B. D. Ripley. 2002. *Modern applied statistics with S*. 4th ed. New York, NY: Springer.

Whitlock, H. V. 1948. Some modifications of the McMaster helminth egg-counting technique and apparatus. *J. Counc. Sci. Ind. Res* 21:177–180.

Wood, E. L., J. M. Burke, P. Dirk, M. Acharya, J. E. Miller, and K. P. Coffey. 2018. Gastrointestinal nematode infection and growth of grass-based Katahdin lambs. *Sheep Goat Res. J.* 33:110. Available from https://d1cqrq366w3ike.cloudfront.net/http/DOCUMENT/SheepUSA/SGRJ_V33_1-10_Wood_11-18.pdf