



**HAL**  
open science

## Diversity of dairy goat lactation curves in France

Mathieu Arnal, Christèle Robert-Granié, Helene Larroque

► **To cite this version:**

Mathieu Arnal, Christèle Robert-Granié, Helene Larroque. Diversity of dairy goat lactation curves in France. *Journal of Dairy Science*, 2018, 101 (12), pp.11040-11051. 10.3168/jds.2018-14980 . hal-02621009

**HAL Id: hal-02621009**

**<https://hal.inrae.fr/hal-02621009>**

Submitted on 26 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



## Diversity of dairy goat lactation curves in France

M. Arnal,\*†<sup>1</sup> C. Robert-Granié,\* and H. Larroque\*

\*GenPhySE, Université de Toulouse, INRA, INPT, ENVT, 31326 Castanet-Tolosan, France

†Institut de l'élevage (Idele), 31326 Castanet-Tolosan Cedex, France

### ABSTRACT

A high level of production at the peak of lactation may be associated with animal health disorders, high feeding costs, and reduced milk supply throughout the year. The objective of this study was to typologize the lactation curves in French dairy goats and analyze the influence of environmental and genetic factors on these curves. The data set consisted of 2,231,720 monthly test-day records of 213,534 French Saanen and Alpine goats recorded between September 2008 and June 2012. First, principal component analysis classified the shape of the lactation curves into 3 principal components: the first component accounted for milk yield level throughout lactation, the second component accounted for lactation persistency, and the third component accounted for milk yield in mid-lactation. Then, from the principal component scores, the lactations were clustered into 5 different groups. Most lactations had a similar shape to the mean curve, except 30% of the lactations that fell into 3 clusters that had a high production level at the peak and then a different persistency according to cluster. Estimated breeding value for milk yield and home region of breeding were the factors most related to lactation production level. Month of kidding, breed, and gestation stage had the biggest effect on persistency. Month of kidding was the factor most strongly linked to mid-lactation production. A herd effect was observed on all 3 principal components.

**Key words:** French dairy goat, test-day milk record, classification, lactation curve

### INTRODUCTION

A goat's level of milk production is often evaluated by its daily average yield or by its total production throughout lactation. However, neither approach factors in the change in milk production according to DIM (i.e., lactation curve). Shape of the lactation curve can affect an animal's dietary needs, and consequently its

health, as well as the time-course scale of milk production. Indeed, a high level of production at the peak can lead to an energy deficit for the animal if it is unable to ingest enough food to compensate. Goats, like dairy cattle, will then begin drawing on their body reserves, which may lead to metabolic and reproductive disorders (Gipson and Grossman, 1990). A goat in energy deficit has to be fed expensive supplement concentrates and high nutrition value forage, whereas a persistent goat can be fed lower-quality forage and fewer concentrates (Sölkner and Fuchs, 1987), which is an asset to farm systems aiming to reduce inputs. The ability to maintain lactation in a partially season-cycled production context meets the processing industry demand for more evenly spread milk production, in line with the needs of the market (Kearney, 2010), while also avoiding the need to freeze cheeses, which can change their flavor (Van Hekken et al., 2005).

Research on shape of the lactation curve has mainly addressed milk persistency, trying to find indicator criteria of the slope of the curve after the peak of lactation (Gengler, 1996), or modeling of overall curve shape. Some authors have focused on finding mathematical models that describe the biological processes of milk production by mammary gland cells (Pollott, 2000; Elvira et al., 2013a). Others have proposed a statistical approach by modeling the shape of the curve from elementary controls such as random regression test-day models including genetic and nongenetic effects (Menéndez-Buxadera et al., 2010; Mucha et al., 2014; Brito et al., 2017) or by using methods such as principal component analysis (PCA) to sum up the information contained in the data (Carta et al., 2014). Principal component analysis is a dimension-reduction tool that reduces a large set of variables to a small set containing most of the original information. Macciotta et al. (2006) used a PCA based on rank of bovine test-days where the first principal component (PC) summarized the averaged milk yield level and the second component summarized persistency. van der Werf et al. (1998) and Druet et al. (2003) performed data dimension reduction on a covariance genetic matrix derived from a random regression test-day model and also obtained 2 PC with the same biological interpretation.

Received April 26, 2018.

Accepted July 23, 2018.

<sup>1</sup>Corresponding author: [mathieu.arnal@idele.fr](mailto:mathieu.arnal@idele.fr)

Bouloc (1991) showed in French dairy goats that lactation curve shapes may depend on breed and genetics of the animal. Menéndez-Buxadera et al. (2010) showed in Murciano-Granadina goats that persistency is heritable (0.20). Curve shape may also vary with breeding and physiological factors such as parity, age at kidding, kidding season, dry-period length, gestation stage, and herd environment factors (Gipson and Grossman, 1990; Bouloc, 1991).

To our knowledge, no recent study has been done on lactation curve shapes using test-day records in French dairy goats. The latest study, performed 27 yr ago, was by Bouloc (1991), who classified protein yield curves based on different input features: lactation duration, average daily protein yield, persistency of the protein yield calculated as the ratio of d 50–d 200-period protein yield to d 31–d 60-period protein yield.

Here we typologized lactation curve shapes based on PCA scores taking into account test-day lactation stage in the 2 main French dairy goat breeds (Alpine and Saanen), using a recent large data set of test-day milk production records from the French genetic national database. We went on to study the influence of different genetic and environmental factors on these curves.

## MATERIALS AND METHODS

### Data

The data were extracted from the genetic database used for the national French genetic evaluations. France has been running a goat milk yield recording scheme since 1989 that applies the standard A, D, AT, AZ, or CZ methods registered by ICAR (2017) with sampling at 4- or 5-wk intervals. The data retained here was collected from the A, D, and AZ protocols with precise measurement of daily production, taking into account amount of milk at the 2 daily milkings. About 60% of the overall database comes from these protocols. The weights of milk produced by a goat in the morning and evening milkings were added together to obtain the weight of daily milk.

Lactations had to have a conventional duration between 180 and 350 d to be conserved. Test-day records before 7 and beyond 300 DIM were removed to have a sufficient number of records per DIM. Within a lactation, a goat was required to have at least 4 test-day records, with the first test day recorded before d 81 of lactation to assess production at the beginning of lactation, the peak of lactation occurring around 40 DIM. Only goats from herds having at least 30 conserved lactations per year were kept in the analysis.

The final data set consisted of 2,231,720 test-day records collected from 319,975 lactations of 213,534

French Saanen and Alpine goats in 910 French herds spread throughout France.

As described by Pulina et al. (2018), in France goats production systems present a great diversity according to feeding (grazing or not, concentrate quantity), reproduction (out of season or not), transformation of milk (cheese maker/deliverer), herd size, or even breeding goal (milk yield/composition). However, the database used recording animal milk productions and did not contain detailed information on production system of herds. Almost 60% of the lactations retained were on Alpine goats, and mainly (77.5%) in the north-west of France. The final data set contained 77% of lactations in parity 1 to 3, with other lactations up to parity 8 (Table 1). Lactations lasted 288 d on average, for an averaged total production of 964.1 kg of milk (minimum: 132.5 kg, maximum: 2,615 kg, SD: 253).

Goats were measured on average 7 times per lactation (minimum: 4, maximum: 11, SD: 1), at an average measure-to-measure interval of 39 d (minimum: 10, maximum: 158, SD: 10.9). Goats were not all measured at the same DIM, and production varied from 0.5 to 11.4 kg per test-day record.

For each lactation, parturition date served to calculate age at kidding and define kidding month. For goats in their second lactation or more, “dry-period length” was defined as the length of the period between the end of the previous lactation and parturition date. Breeders give the end date of the previous lactation; otherwise, it is a deducted date by adding to the date of the last test day 14 d (or 28 d if the interval with the following lactation is greater than 56 d). Gestation stage at 300 DIM was calculated for goats having subsequent lactation.

### Analysis

All statistical analyses were performed using R software (RCore Team, R Foundation for Statistical Computing, 2014, R: A Language and Environment for Statistical Computing, Version 3.0.2, Vienna, Austria, <http://www.R-project.org>).

The general approach of the analyses was to perform (1) a PCA on test-days after smoothing the data, then (2) a classification of the curves, and finally, (3) a study of the relationships between environmental and genetic factors and curve shapes.

**Principal Component Analysis.** The PCA was performed on the test-day milk yield data, for each lactation, using the R package *fdapace* (Dai et al., 2016). *Fdapace* was well-adapted to the structure of our data set: sparse data, irregular measurement stage, irregular number of measurements per lactation, small number of repeated measurements available per lactation. First,

**Table 1.** Percentage of data for each level of different factors in French dairy goats

Variable	%
Parity	
1	32.9
2	26.5
3	17.8
4	10.9
5	6.1
6	3.3
7	1.7
8	0.8
Breed	
Alpine	59.1
Saanen	40.9
Milk EBV (kg)	
-349 to -79	25.1
-79 to -33	25.3
-33 to 13	24.7
13 to 301	24.9
Region <sup>1</sup>	
NE	17.7
NW	77.5
S	4.9
Kidding month	
1	24.51
2	27.89
3	14.24
4	3.50
5	1.13
6	0.19
7	0.03
8	1.06
9	9.02
10	8.21
11	6.83
12	3.40
Kidding age <sup>2</sup> (mo)	
Other	12
P4+	22.8
Normal	50.8
Young	2.9
Old	11.6
Dry-period length <sup>3</sup> (d)	
1-60	19.5
60-90	32.7
90-120	9.1
Other lengths	5.8
Parity 1	32.9
SCC EBV <sup>4</sup> (base 100)	
101-105	23.8
105-131	21.2
65-97	26.2
97-101	25
EBV missing	3.8
Gestation stage <sup>5</sup> (d)	
120-150	5.2
150-180	1.3
1-60	5.9
60-90	32.5
90-120	24
Other stages	1.3
No consecutive lactation	29.9

<sup>1</sup>French herd region (NW = north-west; S = south; NE = north-east).

<sup>2</sup>Age at kidding for the first 3 parturitions (young = 9-10 mo at first kidding, 16-22 mo at second kidding, or 22-34 mo at third kidding; normal = 11-13 mo at first kidding, 23-25 mo at second kidding, or 35-37 mo at third kidding; old = 14-30 mo at first kidding, 26-56 mo at second kidding, or 38-77 mo at third kidding; P4+ = parity 4 and more; other = other ages at kidding for the first 3 parturitions).

<sup>3</sup>Dry-period length before last kidding (d; for goats in second and more lactation).

<sup>4</sup>Somatic cell count EBV (base 100): <100: increase in SCC, >100: decrease in SCC.

<sup>5</sup>Gestation stage (d) at 300 DIM.

fdapace calculated the smoothed mean curve using local linear Gaussian kernel regression between 51 equidistant support points (default option) aggregating all the measurements together. Then, the raw covariance for each curve was calculated, and all covariances were aggregated to generate the sample raw covariance. The smoothed covariance was estimated from the off-diagonal of sample raw covariance. Eigen analyses performed on the smoothed covariance matrix gave the estimated eigenfunctions, eigenvalues, and PC scores (Yao et al., 2005).

To evaluate the accuracy of the smoothed curves using local linear regression, the true and predicted milk yields were compared by studying the coefficient of determination ( $R^2$ ). This analysis can be done according to the number of PC.

**Classification.** Using the PC scores (**SPC**) of the lactations, an unsupervised classification, also called cluster analysis, was performed using the R package *rmixmod* (Langrognon et al., 2016). *Rmixmod* is a tool to fit a mixture model of multivariate components to a data set (Biernacki et al., 2006). A new feature of the *Rmixmod* software is that it considers a parameterization of the variance matrix of a cluster through its eigenvalue decomposition, leading to many meaningful models for clustering (Celeux and Govaert, 1995). Each variance matrix is decomposed as the product of one parameter determining the volume of the cluster, a matrix of eigenvectors determining its orientation, and a diagonal matrix determining its shape (Lebrete et al., 2015). By allowing some but not all of these parameters and matrices to vary between clusters, parsimonious and easily interpreted models can be obtained. As SPC are continuous variables, we tested 14 different Gaussian models with different variance-covariance structures as proposed by Celeux and Govaert (1995), and we selected the best one based on the Bayesian information criterion (**BIC**). The number of components of the mixture model is generally known or fixed beforehand. If this number is not known, it is possible to compare models with different numbers of components and to use criteria proposed in the *Rmixmod* package to make the choice. Here, the optimal number of components was researched between 2 and 5 clusters, and the BIC was used to determine it.

The mixture parameter was estimated through maximization of the log-likelihood. In a first step, a stochastic expectation maximization algorithm was used 100 times with 1,000 iterations each time. Then, an expectation maximization algorithm was performed 100 times with 1,000 iterations each time. The stochastic expectation maximization algorithm used random drawing at each iteration, which avoided converging to

stationary points of the log-likelihood (Biernacki et al., 2003).

**Relationships Between Environmental and Genetic Factors and PC.** The effects of environmental and genetic factors on the first 3 PC were studied according to the following model:

$$y_{ijklmnopqr} = \mu + B_i + R_j + N_k + A_l + M_m + D_n + G_o + I_p + SCI_q + \varepsilon_{ijklmnopqr}, \quad [\text{model 1}]$$

where  $y_{ijklmnopqr}$  is the dependent variable (SPC) and  $\mu$  is overall mean. The fixed effects tested were  $B_i$ , breed (Saanen or Alpine);  $R_j$ , French herd region (NW = north-west; S = south; NE = north-east);  $N_k$ , parity (1 to 8);  $A_l$ , age at kidding for the 3 first parturitions (5 levels: 1 = 9–10 mo at first kidding, 16–22 mo at second kidding, or 22–34 mo at third kidding; 2 = 11–13 mo at first kidding, 23–25 mo at second kidding, or 35–37 mo at third kidding; 3 = 14–30 mo at first kidding, 26–56 mo at second kidding, or 38–77 mo at third kidding; 4 = parity 4 and more; 5 = other ages at kidding for the first 3 parturitions);  $M_m$ , kidding month (12 levels: 1 = January, . . . , 12 = December);  $D_n$ , class of dry-period length before the last kidding (d) for goats in second or more lactations (5 levels: 1 = [1–60], 2 = [60–90], 3 = [90–120], 4 = other lengths, 5 = 1st parity);  $G_o$ , class of gestation stage (d) at 300 DIM (7 levels: 1 = [1–60], 2 = [60–90], 3 = [90–120], 4 = [120–150], 5 = [150–180 d], 6 = other stages, 7 = no consecutive lactation);  $I_p$ , class of genetic EBV based on total-lactation milk yield (kg) from official French genetic evaluation (4 levels: 1 = [–349 to –79], 2 = [–79 to –33], 3 = [–33 to 13], 4 = [13 to 301]);  $SCI_q$ , class of SCC EBV (base 100): <100: increased SCC, >100: decreased SCC) from official French genetic evaluation (5 levels: 1 = [65–97], 2 = [97–101], 3 = [101–105], 4 = [105–131], 5 = missing information); and  $\varepsilon_{ijklmnopqr}$  are the normally distributed residuals. Model 1 using  $SPC_x$  with  $x = \{1, \dots, X\}$  is noted  $M_{SPC_x}$ . Percentages of data for each class of the different factors are presented in Table 1. Despite the low percentages of data, particularly for the kidding month effect in summer, the lowest number of data in a class was 87.

Each effect of model 1 was successively removed to study its relative effect using a Fisher test and comparing the  $R^2$  of the model.

Least squares means were calculated using proportions of the population as weights to take into account the unbalanced number of data in each level of effects, using the R package lsmeans (Lenth, 2016).

The influence of 2-way interactions between herd and year of lactation (2,145 classes) on SPC was studied

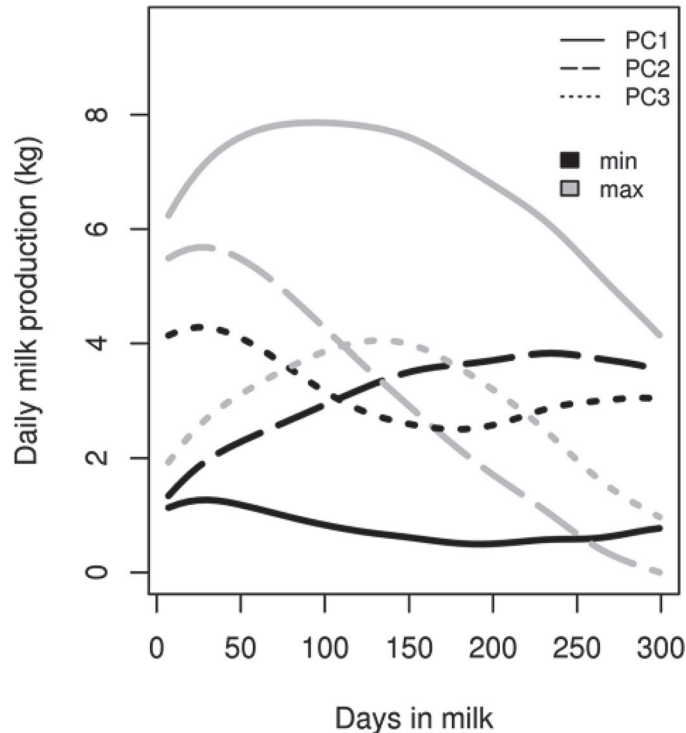
from the residuals of model 1, due to the total confusion between herd number and region.

## RESULTS AND DISCUSSION

### Principal Component Analysis

The first 3 components represented 99.5% of the original variance: 81% for PC1, 15%, for PC2, and 3.5% for PC3. It was not considered useful to further study PC due to the total variance already explained. The  $R^2$ , used to evaluate the precision of the fit by local linear regression, was 0.92 with 3 PC and fell to 0.88 with only the first 2 PC. Clearly, adjustment was better with all 3 components, especially considering the smoothed lactation curves of extreme goats after peak lactation (Figure 1). Therefore, 3 PC were retained.

Plotting extreme lactation curves, with SPC extremes for one PC and other SPC close to 0, showed that each PC characterized a component of curve shape (Figure 1). Principal component 1 was an indicator of level of milk production throughout lactation, where a low SPC1 corresponds to low milk production (mean  $\pm$  SD: SPC1:  $-1.4 \pm 144.4$ , total lactation production (kg):  $964.1 \pm 253.3$ ; correlation between SPC1 and lactation



**Figure 1.** Local linear regressions (curves) for lactations having a minimum (min) or a maximum (max) score for 1 principal component (PC; first, PC1; second, PC2; or third, PC3) and scores close to 0 for the 2 remaining principal components.

**Table 2.** Mean (SD) per cluster for the first, second, and third principal component (PC1, PC2, and PC3, respectively), and proportion of lactations per cluster

Cluster number	PC1	PC2	PC3	Proportion (%)
1	-76.5 (94.4)	107.1 (40.7)	-39.5 (17.7)	4.2
2	-121.7 (80.9)	7.2 (40.8)	9.7 (18.2)	39
3	221.9 (120.5)	-41 (69.3)	-23.8 (30.5)	8.9
4	101.1 (108.7)	64.2 (59.5)	29.4 (26.5)	16.2
5	41.7 (78)	-29 (42.1)	-4.9 (19)	31.7

production: 1); PC2 was an indicator of the difference in milk production between the peak and end of lactation, where a low SPC2 was equivalent to high milk production at the end of the lactation (mean  $\pm$  SD: SPC2:  $4.9 \pm 61.9$ , production difference (kg) between DIM 40 and 240:  $1.11 \pm 0.9$ ; correlation between SPC2 and production difference: 0.99); finally, PC3 was an indicator of milk production after the peak of lactation at around 120 DIM: a low SPC3 equated to low production around 120 DIM (mean  $\pm$  SD: SPC3:  $3.3 \pm 27.2$ , production difference between DIM 120 and the sum between the first and the last DIM (kg):  $-1.6 \pm 0.7$ ; correlation between SPC3 and production difference: 0.97).

On cow data, some authors have defined persistency as a degree of maintenance of peak milk yield (Grossman et al., 1999), others as the degree of decline in milk yield after the peak, using indicators of the speed of decline such as the ratio of milk yield in the first 100 DIM to milk yield in the last 100 DIM (Sölkner and Fuchs, 1987). Based on these definitions of persistency, PC2 and PC3 could both be considered persistency indicators, where “complete” persistency could be a minimal value for PC2 and a maximal value for PC3. For the rest of the study, we defined persistency as a difference in milk production between the peak and end of lactation. Consequently, PC2 was associated with persistency and PC3 was associated with milk production in mid-lactation.

Macciotta et al. (2006) performed a PCA on a correlation matrix from dairy cow test-day milk yields. They used not DIM but rank of the test-day during lactation and obtained 2 PC: the first representing total-lactation production, and the second representing persistency. Carta et al. (2014), with a similar approach, obtained the same results in dairy sheep.

### Classification of Curve Shapes

The best Gaussian mixture model for our data based on the BIC was defined by free volume, equal shape, and equal orientation between the different clusters. The number of clusters tested varied from 2 to 5. The

BIC criteria were 10,775,614, 10,767,801, 10,763,784, and 10,760,253 for 2, 3, 4, and 5 clusters, respectively. The best model was then obtained by considering 5 clusters. To obtain clusters with sufficient sizes, it was not desirable to have a greater number of clusters. On the whole data file, all iterative procedures used to build the best model were run several times, and the same clusters were found each time.

Table 2 reports mean and standard deviation per cluster of the lactation scores for each PC. Figures 2a and 2b represent each lactation according to their PC scores and their cluster membership. Each ellipse represents the internal covariance between 2 PC within a cluster.

Clusters 3 and 4, with the biggest ellipses in Figures 2a and 2b, had the largest SPC variances, with standard deviations of 120.5 and 108.7, respectively, for PC1. In contrast, clusters 2 and 5, although including proportionally more lactations (39 and 31.7%, respectively), had the lowest SPC variances (80.9 and 78, respectively, for PC1, for example).

The averages of daily milk production were plotted for all the data as well as according to cluster number (Figure 3). In each cluster, average daily production did not fluctuate between 2 consecutive DIM. Each shape of average daily production looked like a smooth function.

We estimated the probability of each lactation belonging to a cluster. To challenge the robustness of the clustering, lactations with a cluster probability higher than 0.7 were conserved ( $n = 161,676$  lactations; i.e., 51% of the data set). In this case, the clusters looked well separated (the ellipses did not overlap) and there was no substantial change in the means curve profiles (results not shown).

According to BIC, the repeatability of the classification results, the smoothness of the mean curve of each cluster and the similarity of the mean curves using a probability higher than 0.7, the choice of 5 clusters appeared robust.

Principal component 3 (representing only 3.5% of the original variance) contributed significantly to the diversity of the shape of curves highlighted, which is

easily understandable as PC3 was responsible for the flexibility in mid-lactation. Indeed, not taking into account PC3 in the classification would make cluster 1 disappear and change the shape of clusters 3 and 4 (results not shown).

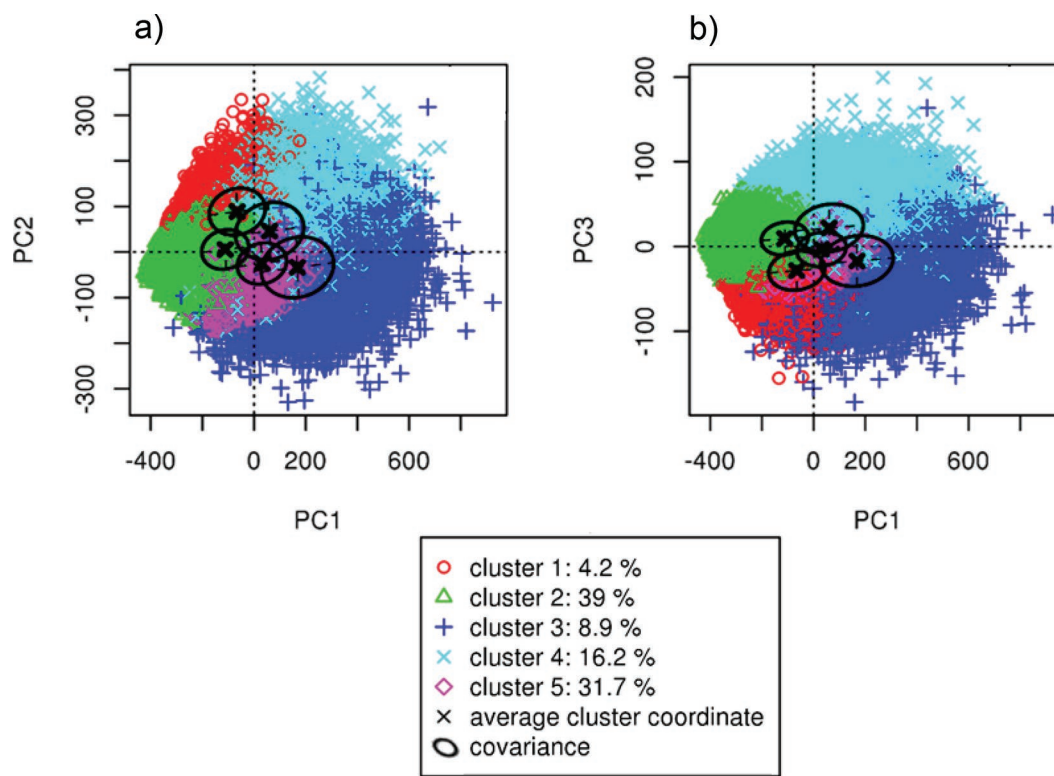
The average lactation curve could be decomposed into 3 phases according to daily milk yield: (1) an initial rapid increase until the 50th DIM, (2) a stabilization for 50 d, and finally, (3) a slow decrease through to the end of lactation.

Day of peak lactation yield differed according to shape of the lactation curve. The lactation peak, defined as the day of the maximal daily milk production, was before 40 DIM for clusters 1 and 5 but was 10 d later for clusters with the highest milk production level (clusters 3 and 4). Goats in cluster 4 maintained peak milk production for a far longer time than goats in cluster 1.

Cluster 2 (the green curve in Figure 3), representing 39% of lactations, was characterized by the lowest milk yield level, intermediate means for PC2 and PC3, and a curve shape similar to the mean curve (the black curve in Figure 3). Cluster 5 (the pink curve in Figure 3), representing 31.7% of lactations, was defined by intermediate means for PC1 and PC3 and the second-lowest

mean for PC2 corresponding to a “flat” curve (Figure 3). The mean curves of 3 others clusters (nearly 30% of the data) were very different from the all-data set mean curve. Indeed, cluster 3 (the dark blue curve), characterizing 8.9% of the data, had the highest mean milk production (PC1) and highest persistency (PC2) but the second-lowest value for PC3. Cluster 1 (the red curve), which ranked fourth for milk production level, had the lowest persistency and the highest decrease in mid-lactation, resulting in a very specific lactation peak (Figure 3), and constituted the smallest group with 4.2% of the lactations. Cluster 4 (the light blue curve in Figure 3), which was characterized by the second-highest total milk production level and ranked second for nonpersistency and first for maintaining milk in mid-lactation, giving an early-bulging convex curve, represented 16.2% of lactations.

On a sample of 146,280 French goats of Saanen breed (31.8%), Alpine breed (56.1%), and other breeds (12.1%) milked in 1987, Bouloc (1991) realized a classification based on 3 criteria: (1) persistency of protein yield calculated as the ratio of d 50–d 200-period protein yield to d 31–d 60-period protein yield, (2) average daily protein yield, and (3) duration of lactation. The author identified 8 distinct clusters and was able to



**Figure 2.** Plot of the lactations according to (a) the first 2 principal components (PC1 and PC2), and (b) the first and third principal components (PC1 and PC3).

**Table 3.** Coefficient of determination ( $R^2$ ) of model 1 where the lactation scores of each of the 3 principal components (PC) are explained by all the environmental factors, and of model 1 successively reduced by 1 environmental factor

Item	PC1	PC2	PC3
Model 1	0.402	0.225	0.135
Model 1 without parity	0.399	0.215	0.134
Model 1 without kidding age	0.399	0.222	0.133
Model 1 without kidding month	0.395	0.199	0.034
Model 1 without dry-period length	0.400	0.222	0.135
Model 1 without gestation stage	0.395	0.208	0.130
Model 1 without breed	0.394	0.201	0.135
Model 1 without milk EBV	0.161	0.224	0.135
Model 1 without SCS EBV	0.401	0.219	0.135
Model 1 without region	0.385	0.223	0.134

highlight: (1) lactations with atypical forms (i.e., low production during the first weeks, which returned to an average level thereafter), (2) persistent curves with a difference of 1.5 kg of milk yield between the peak and end of lactation, (3) curves with marked peaks and a difference of more than 2.0 kg in milk yield between the peak and end of lactation. In our data set, only the mean lactation curve of cluster 2 (low milk yield level and medium persistency) was similar to what Bouloc found. Bouloc's study did not find the profiles of clus-

ters 3 and 5 (high production and high persistency), probably due to lower levels of genetics and management 27 yr ago, nor the profiles of clusters 4 and 1, whereas the atypical curves found by Bouloc (1991) were not found here, probably due to the use of different phenotypes for the classification.

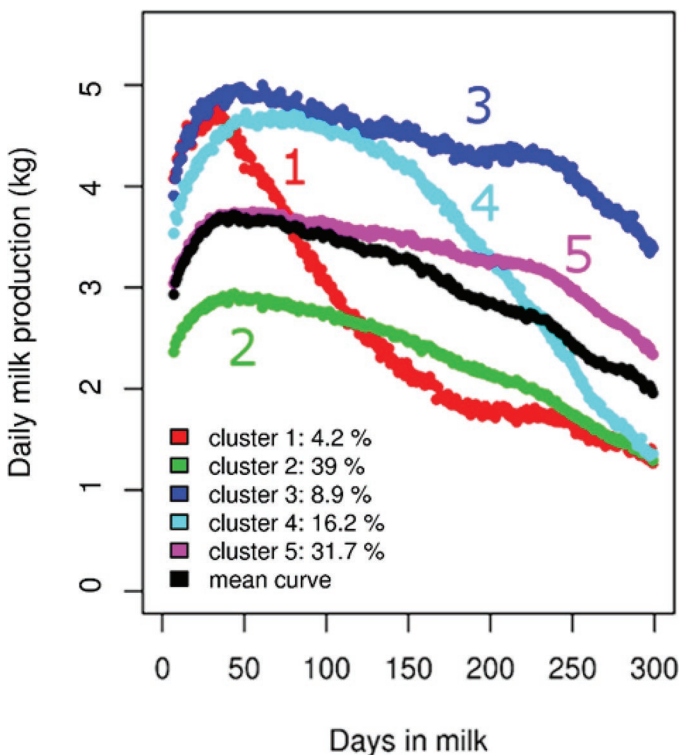
#### Relationship Between Environmental and Genetic Factors and PC

Model 1 was used to estimate the relative influence of each environmental and genetic factor on each PC. For each PC, Table 3 presents the  $R^2$  of model 1 and the  $R^2$  of model 1 iterative reduced by one of the factors. The Fisher's tests between model 1 and the different sub-models were all significant ( $P < 10^{-5}$ ) except for breed on PC3 ( $P = 0.08$ ).

The  $R^2$  of model 1 was 0.402 for PC1. The biggest decrease in  $R^2$  ( $R^2 = 0.161$ ) occurred when EBV of total milk yield was removed from model 1, meaning that it was the factor most related to PC1, far ahead of region ( $R^2 = 0.385$ ) and breed ( $R^2 = 0.394$ ). For PC2, the  $R^2$  of the model 1 was 0.225. Principal component 2 was not strongly related to environmental and genetic factors, as the biggest association was found for kidding month with  $R^2$  decreased to 0.199, followed by breed ( $R^2 = 0.201$ ), gestation stage ( $R^2 = 0.208$ ), parity ( $R^2 = 0.215$ ), and SCS EBV ( $R^2 = 0.219$ ). The  $R^2$  of model 1 for PC3 was 0.135. The only factor strongly linked to this PC was kidding month, with  $R^2$  reduced to 0.034.

Figures 4, 5, 6, 7, and 8 chart the effects on PC1 and PC2 of factors having the greatest effects.

**Milk EBV and Region.** Clearly, PC1 score increased with EBV milk yield classes (Figure 4). Furthermore, the correlation between milk EBV and PC1 was +0.51, meaning that animals with the highest genetic breeding value for total milk yield had the highest level of daily milk yield. Milk EBV was not linked with PC2, meaning that genetic selection on total-lactation



**Figure 3.** Average lactation curves for all data (in black) and according to cluster number (each point corresponds to the average milk output of the animals).

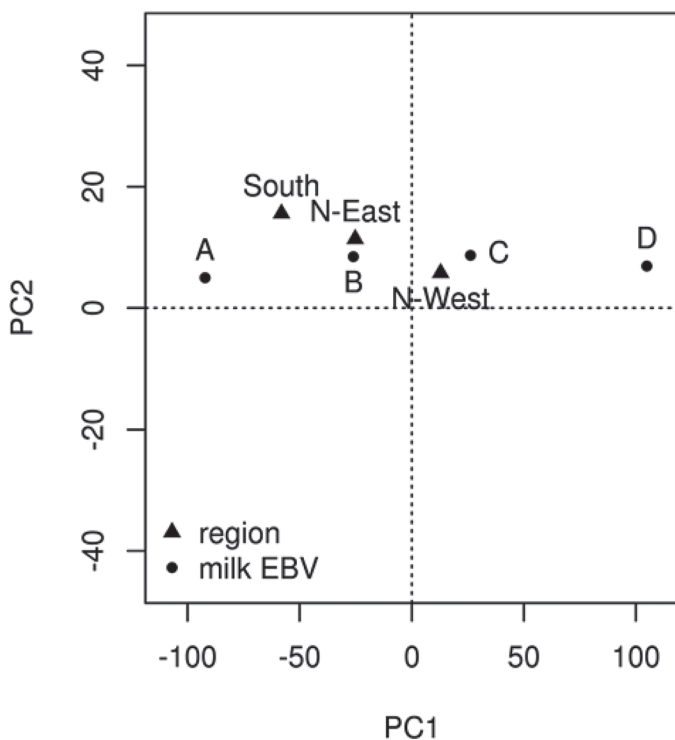


milk production was not linked to lactation persistency. The region factor showed a gradient of increasing production from south < north-east < north-west France.

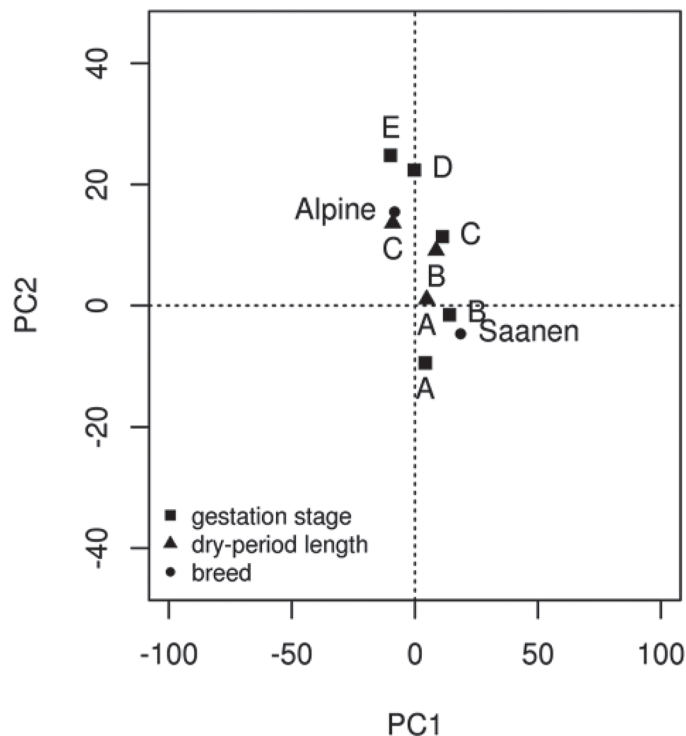
Bouloc (1991) observed differences in milk production between regions: goats from Pays-de-la-Loire (north-west in our study) were more numerous in high-producing groups whereas goats from Midi-Pyrenees (South in our study) were relatively more numerous in lower-producing groups.

**Gestation Stage, Dry-Period Length, and Breed.** Goats had more persistent lactations when their gestation was less advanced at 300 DIM and when they were dried for a short period lower than 90 d before their last kidding. Saanen goats were more productive and more persistent than Alpines (Figure 5).

Like the current study, Bouloc (1991) also observed that Saanen goats were more numerous in the groups of persistent lactations. Furthermore, Rupp et al. (2011) observed in the French goat population that Saanens were more productive than Alpines. Like the current study, Knight and Wilde (1988) reported that gestation causes a decrease in milk yield compared with nonges-

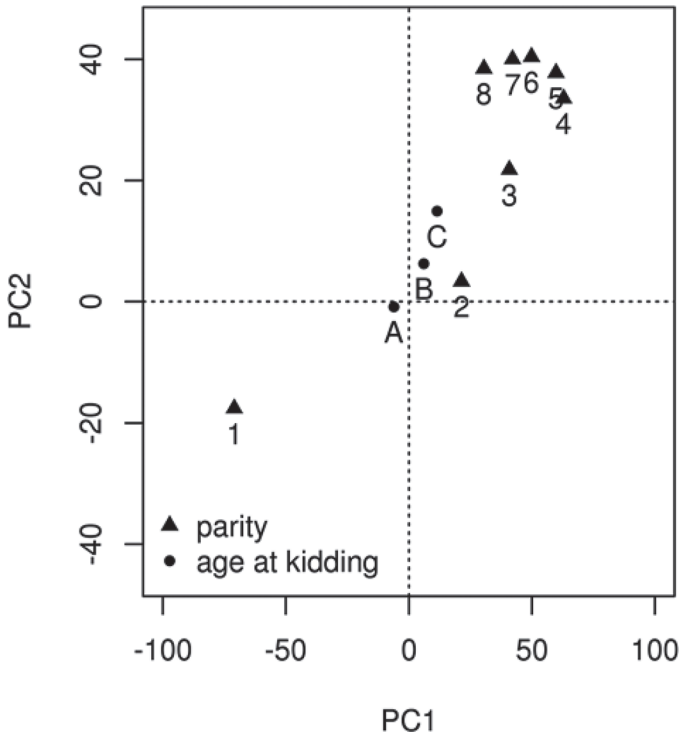


**Figure 4.** Least squares means on the first 2 principal components (PC1 and PC2) of the environmental factors most strongly linked to the first principal component: milk EBV (kg) from the French genetic evaluation model (A: -349 to -79, B: -79 to -33, C: -33 to 13, D: 13 to 301), and home region of the herd (N-West = north-west; S = south; N-East = north-east).



**Figure 5.** Least squares means on the first 2 principal components (PC1 and PC2) of the environmental factors strongly linked to the second principal component: gestation stage (d) at 300 DIM (A: 1-60, B: 60-90, C: 90-120, D: 120-150, E: 150-180), dry-period length (d) (A: 1-60, B: 60-90, C: 90-120), and breed (Saanen or Alpine).

tating does. Although their study counted a limited number of goats, they found that the influence of gestation was significant after 8 wk of gestation. Salama et al. (2005), on goats milked once a day, also found that the gestation caused a decrease of milk yield on gestating goats compared with goats in extended lactation. In the dairy cow, Druet et al. (2003) also showed a decrease of milk yield according to the progress of gestation stage. The effect of dry-period length on lactation curve has been understudied. Knight and Wilde (1988) specified that an absence of dry period cause a higher peak and lower persistency. Caja et al. (2006) studied the effect of dry-period length on 17 goats and found that no-dry-period goats produced less than dry-period goats. Atashi et al. (2013) studied the influence of the dry period length on lactation curve shape, in dairy cows. As in our study, they noticed the best persistency, the smallest peak, and the higher milk lactation production for the cow dried previously during a short time. In contrary, they observed also that animals previously dried for a long period produced higher milk yield at the peak than animals dried during an average period. Similarly, Elvira et al. (2013b) showed that the longer dry periods were not associated with high milk yield



**Figure 6.** Least squares means on the first 2 principal components (PC1 and PC2) of the environmental factors linked to the first 2 principal components: parity (1 to 8) and age at kidding (A: 9–10 mo at first kidding, 16–22 mo at second kidding or 22–34 mo at third kidding; B: 11–13 mo at first kidding, 23–25 mo at second kidding or 35–37 mo at third kidding; C: 14–30 mo at first kidding, 26–56 mo at second kidding or 38–77 mo at third kidding).

productions throughout the lactation, on their sheep data.

**SCS EBV.** Goats had a more persistent lactation when they had a high SCS EBV (less SCC). The relationship between high SCC index (indicator of genetically low SCC levels) and persistency in goats remains unknown. One hypothesis could be that goats with genetically low SCC levels are more resistant to mammary infections and thus better able to maintain high milk production. Indeed, Leitner et al. (2004) reported that infected goats produced less milk than the noninfected goats.

In bovines, Appuhamy et al. (2009) showed a negative genetic correlation ( $-0.24$ ) between persistency and a mastitis occurring in the second half of lactation. Cole and Null (2009) found a negative correlation between SCS persistency and milk persistency. Note that Nayeri et al. (2017) found a common QTL between persistency and mastitis.

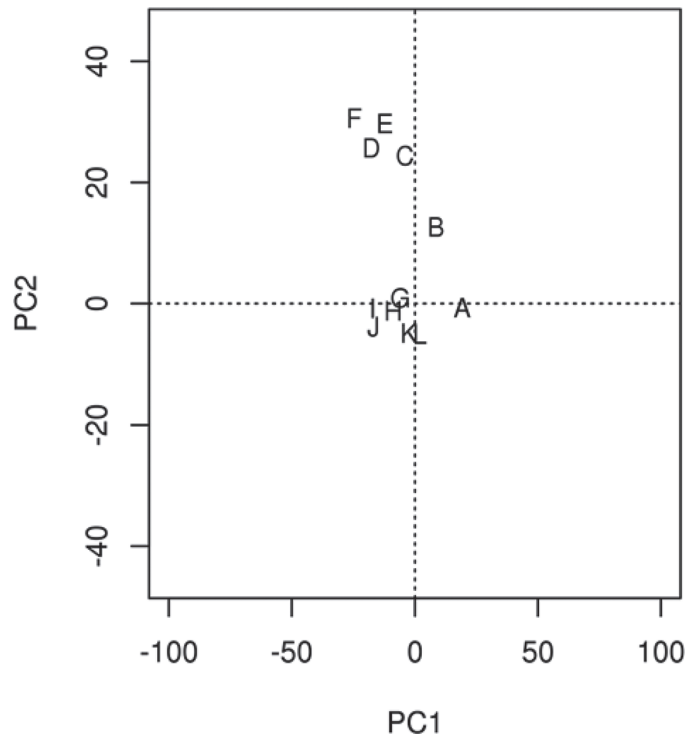
**Age at Kidding and Parity.** Goats that kidded youngest (9–10 mo at first kidding, 16–22 mo at second kidding, or 22–34 mo at third kidding) had a lower milk production levels but higher persistency (Figure

6). From first to fourth parity, the goats produced more milk on average but their lactations were less persistent. Particularly, this highlights the important difference in lactation production potential between goats in their first parity and others parities. After the fourth parity, level of milk production decreased more and more whereas lactations were a little less persistent than in the fourth lactation.

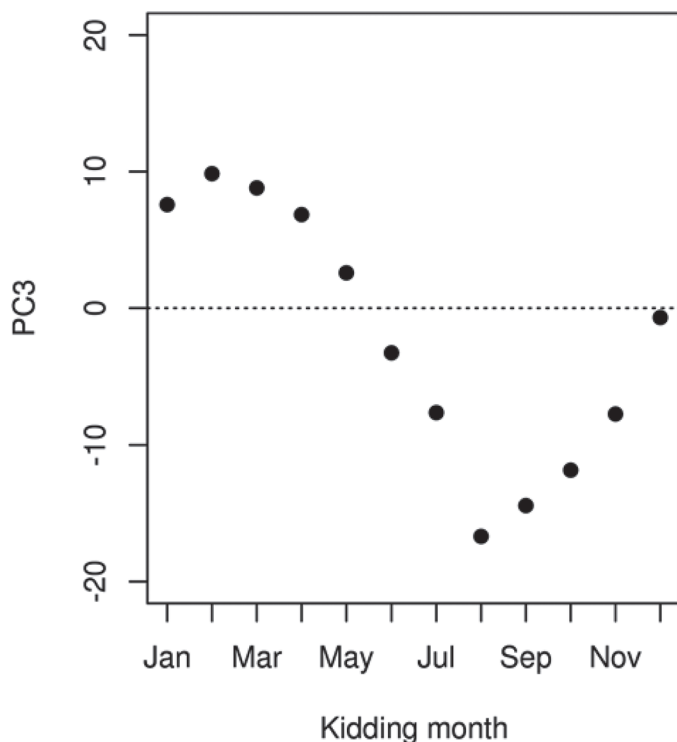
León et al. (2012) analyzed the production of 38,000 Murciano-Granadina breed goats following their first 5 lactations. Like in the current study, they showed that persistency was greater in first lactation than second lactation, which was in turn greater than third lactation and more.

In the French goat population, Boichard et al. (1989) showed that age at kidding had a positive near-linear relationship with milk yield. Bouloc (1991) found the same effect of age at kidding on shape of the lactation curve as the current study: goats kidding before 12 mo had a higher lactation persistency. French dairy cattle showed a similar pattern (Leclerc et al., 2008).

**Kidding Month.** Lactations were less persistent in goats kidding during the spring but more persistent in goats kidding in autumn (Figure 7). Note that the end



**Figure 7.** Least squares means on the first 2 principal components (PC1 and PC2) of the environmental factor the most linked to the second principal component: kidding month (A = January; B = February; C = March; D = April; E = May; F = June; G = July; H = August; I = September; J = October; K = November; L = December).



**Figure 8.** Least squares means on the third principal component (PC3) of kidding month, which is the environmental factor most linked to the PC3.

of the more persistent lactation occurred when days were long (at the beginning of summer), whereas the end of the less persistent lactation occurred when days were short (at the beginning of winter).

The effect of kidding month on PC3 showed broad variability, with a maximal value in February and a minimum value in August (Figure 8). Principal component 3 represented the production at approximately 120 DIM, and variations of the least squares means followed the evolution of the photoperiod duration with 120 d difference. For example, a goat that kids in February will be at its 120 DIM in June when photoperiod is maximal, whereas a goat that kids in August will be at its 120 DIM in December when photoperiod is minimal.

The goat producing more during the longer days will yield more milk from the middle to the end of the lactation according to kidding month. The influence of kidding month on persistency had already been highlighted in studies by Bouloc (1991), Montaldo et al. (1997), and León et al. (2012). In these 3 studies, the greater persistency in goats kidding between October and January was explained by the thriving spring grass, which occurs after the peak of lactation and maintains milk production. Here, 80% of the data were recorded in the north-west (Loire-Atlantique, Poitou-Charente)

where 84% of the goats did not pasture (Caillat et al., 2016). Therefore, the spring grass did not explain the persistency or the effect on PC3. However, abundant literature is available on the effect of duration of the day, which has a significant effect on goat lactation curves (Chemineau et al., 2007; Russo et al., 2013). These studies showed that long exposure to artificial light (16 h per day) has a big effect on milk yield throughout the lactation stage (up to +33% in France; Delouis and Mirman, 1984). Here, we found this same effect of photoperiod, independent of other environmental factors, on a large data set of goats raised in commercial herds.

A kidding in January seems interesting (high lactation yield, average persistency, and high production around DIM120) but has to be put in relation to the milk price, which fluctuates a lot seasonally, and to justify a kidding out of season.

**Herd Effect.** The influence of herd was studied on the residuals of model 1. The remaining variance explained by the herd effect was relatively significant, as illustrated by the  $R^2$  of 0.397, 0.178, and 0.162 for residuals of  $M_{SPC1}$ ,  $M_{SPC2}$ , and  $M_{SPC3}$ , respectively. The dispersion of herd effects on residuals was large, ranging from  $-243.8$  to  $314.5$  for PC1, from  $-90.43$  to  $101$  for PC2, and from  $-35.82$  to  $43.21$  for PC3. Bouloc (1991) demonstrated that herd effect had the biggest effect, in comparison to the other factors, on milk production and persistency.

When the herd effect was replaced by the estimated herd effect on total-lactation milk yield from the official genetic evaluation (4 classes),  $R^2$  was 0.329 for PC1 and 0 for PC2 and PC3. This showed that the herd effect from the official genetic evaluation well explained the average of all-herd milk production level. However, as expected, the herd effect did not take into account the effect of the herd on persistency and production at 120 DIM.

Lactation length was not studied as a factor influencing lactation curves because of its links with breeder decisions (according to the potential of the animal, the day of insemination, and breeder strategies). León et al. (2012) specified that number of kids affected the production and shape of the lactation curve. A goat with more kids had more milk in the beginning of the lactation and was less persistent. Zumbach et al. (2008), like Fernández et al. (2002), used the “number of born kids” factor in their genetic evaluation model. This factor was not tested here as the information was unavailable. Likewise, we did not have more precise data on the farm regarding, for example, feeding system, system of sales, breeding goal, or climatic conditions, to better understand the large influence of herd on the different SPC values.

## CONCLUSIONS

This study, based on a big recent data set, typologized lactation curve shapes in the current French goat population. It demonstrated a very real diversity of lactation curve shapes in France: 30% of the lactations with a shape very different from the mean. This great diversity of curves is mainly due to the global level of milk production and persistency of the lactation, but not only this. Indeed, we showed that the curve bends in mid-lactation, at 120 DIM, which causes particular lactation curve shapes tied to duration of the photo-period. Furthermore, this trait could be viewed as a complementary indicator of milk persistency. Lactation persistency plays a key role in the discrimination of lactation curve types, highlighting a certain variability. However, none of the environmental and genetic factors tested in our study had a strong effect on this trait. Some well-known factors could not be tested here, which may explain these results. This trait also has a genetic variability that we were unable to factor here. Other approaches will have to be tested to investigate the genetic variability of the lactation curves. The use of random regression models taking into account environmental and genetic effects that vary throughout lactation will have to be explored. The present study, showing the effect of some environmental factors on the lactation curve shape, will help us to model the environmental part of the random regression model. Once the genetic evaluation is carried out, the correlations between lactation shape curve and current or future selected characters (on milk quality, longevity, or health) will have to be evaluated to ensure the relevance of the selection. The effects of the environmental factors could also be used in management advice in the dairy goat husbandry.

## ACKNOWLEDGMENTS

This project receives financial support from Ministry of Agriculture (Compte d'affectation Spécial au Développement Agricole et Rural, CASDAR, Rustic). The first author received financial support from APIS-GENE (Paris, France) and Association nationale de la recherche et de la technologie (ANRT, Paris, France).

## REFERENCES

- Appuhamy, J. A., B. G. Cassell, and J. B. Cole. 2009. Phenotypic and genetic relationships of common health disorders with milk and fat yield persistencies from producer-recorded health data and test-day yields. *J. Dairy Sci.* 92:1785–1795.
- Atashi, H., M. J. Zamiri, and M. Dadpasand. 2013. Association between dry period length and lactation performance, lactation curve, calf birth weight, and dystocia in Holstein dairy cows in Iran. *J. Dairy Sci.* 96:3632–3638.
- Biernacki, C., G. Celeux, and G. Govaert. 2003. Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Comput. Stat. Data Anal.* 41:561–575.
- Biernacki, C., G. Celeux, G. Govaert, and F. Langrognet. 2006. Model-based cluster and discriminant analysis with the MIXMOD software. *Comput. Stat. Data Anal.* 51:587–600.
- Boichard, D., N. Bouloc, G. Ricordeau, A. Piacere, and F. Barillet. 1989. Genetic parameters for first lactation dairy traits in the Alpine and Saanen goat breeds. *Genet. Sel. Evol.* 21:205.
- Bouloc, N. 1991. Analyse de la forme de la courbe de lactation. Application à l'étude des modalités d'allègement du contrôle laitier et de prévision précoce de la production dans l'espèce caprine. PhD thesis, Institut national agronomique (INA) Paris Grignon, Paris, France.
- Brito, L. F., F. G. Silva, H. R. Oliveira, N. O. Souza, G. C. Caetano, E. V. Costa, G. R. O. Menezes, A. L. P. Melo, M. T. Rodrigues, and R. A. Torres. 2017. Modelling lactation curves of dairy goats by fitting random regression models using Legendre polynomials or B-splines. *Can. J. Anim. Sci.* 98:73–83.
- Caillat, H., B. Ranger, E. Bruneteau, C. Paraud, and R. Delargarde. 2016. Use of grazing in a dairy goat farm to design sustainable production systems in France. Page 794 in Book of Abstract of the 67th Annual Meeting of the European Federation of Animal 67. Annual Meeting of the European Association for Animal Production (EAAP). Wageningen Academic Publishers, Wageningen (Pays-Bas), Belfast, GBR.
- Caja, G., A. A. K. Salama, and X. Such. 2006. Omitting the dry-off period negatively affects colostrum and milk yield in dairy goats. *J. Dairy Sci.* 89:4220–4228.
- Carta, A., S. Casu, M. G. Usai, and S. Salaris. 2014. Heritability of persistency traits and their genetic correlations with milk yield and udder morphology in dairy sheep. Page 39th ICAR Session, Berlin (Germany). <https://www.icar.org/wp-content/uploads/2016/07/Berlin-2014-Carta-Casu-Usai-Heritability-of-persistency-traits-and-their-genetic-correlations.pdf>.
- Celeux, G., and G. Govaert. 1995. Gaussian parsimonious clustering models. *Pattern Recognit.* 28:781–793.
- Chemineau, P., B. Malpoux, J. P. Brillard, and A. Fostier. 2007. Seasonality of reproduction and production in farm fishes, birds and mammals. *Animal* 1:419–432.
- Cole, J. B., and D. J. Null. 2009. Genetic evaluation of lactation persistency for five breeds of dairy cattle. *J. Dairy Sci.* 92:2248–2258.
- Dai, X., P. Z. Hadjipantelis, H. Ji, H.-G. Mueller, and J.-L. Wang. 2016. Fdapace: Functional Data Analysis and Empirical Dynamics. R Package Version 0.2.5. Accessed Jan. 18, 2017. <https://cran.r-project.org/web/packages/fdapace/index.html>.
- Delouis, C., and B. Mirman. 1984. Influence de la durée quotidienne d'éclairage sur la production laitière de la chèvre. *Compte-Rendu 9e Journ. Rech. Ovine Caprine* 352–360.
- Druet, T., F. Jaffrézic, D. Boichard, and V. Ducrocq. 2003. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. *J. Dairy Sci.* 86:2480–2490.
- Elvira, L., F. Hernandez, P. Cuesta, S. Cano, J.-V. Gonzalez-Martin, and S. Astiz. 2013a. Accurate mathematical models to describe the lactation curve of Lacaune dairy sheep under intensive management. *Animal* 7:1044–1052.
- Elvira, L., F. Hernandez, P. Cuesta, S. Cano, J.-V. Gonzalez-Martin, and S. Astiz. 2013b. Factors affecting the lactation curves of intensively managed sheep based on a clustering approach. *J. Dairy Res.* 80:439–447.
- Fernández, C., A. Sánchez, and C. Garcés. 2002. Modeling the lactation curve for test-day milk yield in Murciano-Granadina goats. *Small Rumin. Res.* 46:29–41.
- Gengler, N. 1996. Persistency of lactation yields: A review. *Interbull Bull.* 12:87–96.
- Gipson, T. A., and M. Grossman. 1990. Lactation curves in dairy goats: A review. *Small Rumin. Res.* 3:383–396.
- Grossman, M., S. M. Hartz, and W. J. Koops. 1999. Persistency of lactation yield: A novel approach. *J. Dairy Sci.* 82:2192–2197.

- ICAR. 2017. ICAR guidelines. Accessed Mar. 8, 2018. <https://www.icar.org/index.php/icar-recording-guidelines/>.
- Kearney, J. 2010. Food consumption trends and drivers. *Philos. Trans. R. Soc. B Biol. Sci.* 365:2793–2807.
- Knight, C. H., and C. J. Wilde. 1988. Milk production in concurrently pregnant and lactating goats mated out of season. *J. Dairy Res.* 55:487–493.
- Langrognet, F., R. Lebre, C. Poli, S. Iovleff, B. Auder, C. Biernacki, G. Celeux, and G. Govaert. 2016. Rmixmod: Supervised, Unsupervised, Semi-Supervised Classification with MIXture MODELing (Interface of MIXMOD Software). R Package Version 2.1.1. Accessed Jan. 18, 2017. <https://cran.r-project.org/web/packages/Rmixmod/index.html>.
- Lebre, R., S. Iovleff, F. Langrognet, C. Biernacki, G. Celeux, and G. Govaert. 2015. Rmixmod: The r package of the model-based unsupervised, supervised and semi-supervised classification mixmod library. *J. Stat. Softw.* 67:241–270.
- Leclerc, H., D. Duclos, A. Barbat, T. Druet, and V. Ducrocq. 2008. Environmental effects on lactation curves included in a test-day model genetic evaluation. *Animal* 2:344–353.
- Leitner, G., U. Merin, and N. Silanikove. 2004. Changes in milk composition as affected by subclinical mastitis in goats. *J. Dairy Sci.* 87:1719–1726.
- Lenth, R. V. 2016. Least-Squares Means: The R Package lsmeans. *J. Stat. Softw.* 69:1–33.
- León, J. M., N. P. P. Macciotta, L. T. Gama, C. Barba, and J. V. Delgado. 2012. Characterization of the lactation curve in Murciano-Granadina dairy goats. *Small Rumin. Res.* 107:76–84.
- Macciotta, N. P. P., D. Vicario, and A. Cappio-Borlino. 2006. Use of multivariate analysis to extract latent variables related to level of production and lactation persistency in dairy cattle. *J. Dairy Sci.* 89:3188–3194.
- Menéndez-Buxadera, A., A. Molina, F. Arrebola, M. J. Gil, and J. M. Serradilla. 2010. Random regression analysis of milk yield and milk composition in the first and second lactations of Murciano-Granadina goats. *J. Dairy Sci.* 93:2718–2726.
- Montaldo, H., A. Almanza, and A. Juárez. 1997. Genetic group, age and season effects on lactation curve shape in goats. *Small Rumin. Res.* 24:195–202.
- Mucha, S., R. Mrode, M. Coffey, and J. Conington. 2014. Estimation of genetic parameters for milk yield across lactations in mixed-breed dairy goats. *J. Dairy Sci.* 97:2455–2461.
- Nayeri, S., M. Sargolzaei, M. K. Abo-Ismael, S. Miller, F. Schenkel, S. S. Moore, and P. Stothard. 2017. Genome-wide association study for lactation persistency, female fertility, longevity, and lifetime profit index traits in Holstein dairy cattle. *J. Dairy Sci.* 100:1246–1258.
- Pollott, G. E. 2000. A biological approach to lactation curve analysis for milk yield. *J. Dairy Sci.* 83:2448–2458.
- Pulina, G., M. J. Milán, M. P. Lavín, A. Theodoridis, E. Morin, J. Capote, D. L. Thomas, A. H. D. Francesconi, and G. Caja. 2018. Invited review: Current production trends, farm structures, and economics of the dairy sheep and goat sectors. *J. Dairy Sci.* 101:6715–6729.
- Rupp, R., V. Clément, A. Piacere, C. Robert-Granié, and E. Manfredi. 2011. Genetic parameters for milk somatic cell score and relationship with production and udder type traits in dairy Alpine and Saanen primiparous goats. *J. Dairy Sci.* 94:3629–3634.
- Russo, V. M., A. W. N. Cameron, F. R. Dunshea, A. J. Tilbrook, and B. J. Leury. 2013. Artificially extending photoperiod improves milk yield in dairy goats and is most effective in late lactation. *Small Rumin. Res.* 113:179–186.
- Salama, A. A., G. Caja, X. Such, R. Casals, and E. Albanell. 2005. Effect of pregnancy and extended lactation on milk production in dairy goats milked once daily. *J. Dairy Sci.* 88:3894–3904.
- Sölkner, J., and W. Fuchs. 1987. A comparison of different measures of persistency with special respect to variation of test-day milk yields. *Livest. Prod. Sci.* 16:305–319.
- van der Werf, J. H. J., M. E. Goddard, and K. Meyer. 1998. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. *J. Dairy Sci.* 81:3300–3308.
- Van Hekken, D. L., M. H. Tunick, and Y. W. Park. 2005. Effect of frozen storage on the proteolytic and rheological properties of soft caprine milk cheese. *J. Dairy Sci.* 88:1966–1972. [https://doi.org/10.3168/jds.S0022-0302\(05\)72872-1](https://doi.org/10.3168/jds.S0022-0302(05)72872-1).
- Yao, F., H.-G. Müller, and J.-L. Wang. 2005. Functional data analysis for sparse longitudinal data. *J. Am. Stat. Assoc.* 100:577–590.
- Zumbach, B., S. Tsuruta, I. Misztal, and K. J. Peters. 2008. Use of a test day model for dairy goat milk yield across lactations in Germany. *J. Anim. Breed. Genet.* 125:160–167.