

Scotland's Rural College

## Analyses of lambing dates in sheep breeds using von Mises distribution

Id-Lahoucine, Samir; Schaeffer, Larry R.; Cánovas, Angela; Casellas, Joaquim

*Published in:*

Journal of Animal Breeding and Genetics

*DOI:*

[10.1111/jbg.12664](https://doi.org/10.1111/jbg.12664)

Print publication: 01/05/2022

*Document Version*

Peer reviewed version

[Link to publication](#)

*Citation for published version (APA):*

Id-Lahoucine, S., Schaeffer, L. R., Cánovas, A., & Casellas, J. (2022). Analyses of lambing dates in sheep breeds using von Mises distribution. *Journal of Animal Breeding and Genetics*, 139(3), 271-280. <https://doi.org/10.1111/jbg.12664>

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1 **Analyses of lambing dates in sheep breeds using von Mises distribution**

2 Short title: Analyses of circular data

3

4 S. Id-Lahoucine<sup>1,2,\*</sup>, L. R. Schaeffer<sup>1</sup>, A. Cánovas<sup>1</sup> & J. Casellas<sup>3</sup>

5

6 <sup>1</sup>Department of Animal Biosciences, University of Guelph, Guelph N1G 2W1, Ontario,  
7 Canada.

8 <sup>2</sup>Scotland's Rural College, Edinburgh EH9 3JG, Scotland, UK.

9 <sup>3</sup>Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona,  
10 Bellaterra 08193, Barcelona, Spain.

11 \*Corresponding author: S. I.-L. Samir.IdLahoucine@sruc.ac.uk

12

13 **ABSTRACT**

14 Regular changes in the environment and biological responses generate seasonal patterns  
15 in the reproduction in small ruminants. Breeding seasonality is a significant constraint  
16 impacting efficiency of lamb production. However, seasonality related traits present a  
17 special peculiarity from a statistical point of view being circular data (day of year running  
18 1:365). Recently circular mixed models have been developed on the basis of the von  
19 Mises distribution and were applied to analyze lambing day distribution recorded from  
20 five major Canadian sheep breeds (Rideau Arcott, Romanov, Dorset, Suffolk and  
21 Polypay). In a simulation study, the linear model was not able to capture the variance  
22 components simulated under the circular paradigm, however, the von Mises model  
23 evidenced its ability to infer the variance components of simulated circular records.  
24 Using real data of sheep, mostly negligible variances were observed for additive genetic  
25 effect when using a linear model on circular data values. In contrast, when using the von  
26 Mises model, genetic variances were different across breeds and it raises the possibility  
27 to delay the peak of reproduction and to change the seasonality of the ewes. However, a  
28 large variance was captured by flock-year effects emphasizing the strong influence of  
29 management in lambing seasons for Canadian sheep populations. Finally, the results  
30 suggest the potential of using the von Mises model to analyze circular data and further  
31 research is needed for better understand the complexity of this trait and the von Mises  
32 models.

33 **Keywords:** circular data, lambing distribution, seasonality, sheep, von Mises

## 34 1 INTRODUCTION

35 Regular changes in the environment and biological responses generate seasonal  
36 patterns (Battey, 2000). In agriculture, seasonality causes fluctuations in the availability  
37 of plant and animal products over the year. Seasonal reproduction patterns are recognized  
38 in small ruminants which significantly constrain efficiency of production (Notter and  
39 Cockett, 2005). Seasonal distribution of births in small ruminants results from seasonal  
40 variations in their sexual behaviour, spermatogenic activity, fertilization rates, and  
41 ovulation frequency, and in turn are consequences of the natural endogenous rhythmicity  
42 in combination with environmental factors such as photoperiod (Chemineau et al., 2010).  
43 Seasonality must be viewed as an inherited characteristic from the ancestral natural  
44 populations where animals give birth at the most favourable period for the progeny to  
45 survive (Chemineau et al., 2007). In fact, different degrees of breeding seasonality are  
46 observed from southern to northern latitudes in both sheep and goat species (Chemineau  
47 et al., 2010). Specifically, breeds in the tropics generally present a very low seasonality  
48 (Yenikoye 1984; Arroyo et al., 2007) whereas more marked seasonality is expressed in  
49 breeds in the temperate climates of northern countries with large photoperiodic variations  
50 (Chemineau et al., 2004, 2010). The variability between and within breeds indicates a  
51 genetic basis of reproductive seasonality; thus, genetic selection could be a feasible  
52 strategy to decrease seasonality, opening opportunities for farmers to adjust the breeding  
53 season of their flock all year-round.

54 Seasonality related traits present a special peculiarity from a statistical perspective.  
55 Lambing dates (as day of year), for example, are distributed around a circular parametric  
56 space of a yearly basis (i.e., from January to December). Circular data is common in

57 biological studies (Landler et al., 2018) such as processes within time-related biological  
58 rhythms (Cancho-Candela et al., 2007; Gustafson and Partch 2015), diseases/infections  
59 (Mooney et al., 2003; Gao et al., 2006), and cardinal points (Begall et al., 2008;  
60 Shimatani et al., 2012), among others. They differ by existing in a periodical closed space  
61 in a real or virtual circle from  $0^\circ$  to  $360^\circ$  (0 to  $2\pi$  radians). With their periodic nature,  
62 measurements at  $0^\circ$  ( $0\pi$ ) and  $360^\circ$  ( $2\pi$ ) are overlapped and not located at opposite ends  
63 of a linear scale.

64 In animal breeding, circular data have not received attention because of complicated  
65 statistical analysis and the general propensity of animal breeders to transform non-linear  
66 to linear data (Erdem and Shi, 2011). Recently, Casellas et al. (2019a,b) derived the  
67 circular mixed model generalized from the Mulder and Klugkist (2017) circular model  
68 based on the von Mises distribution (Fisher, 1993). The circular model was adapted to  
69 include random sources of variations under the homogeneous variance model and also  
70 extended to account for partly genetic effects underlying the heterogeneity of  
71 environmental variance (i.e., canalization model) as proposed by SanCristobal-Gaudy et  
72 al. (1998) and Ibáñez-Escriche et al. (2008). In the context of canalization models, as  
73 seasonality is associated with a concentration of lambing in specific periods, the increase  
74 of the residual variance could be a hypothetical plausible strategy to reduce seasonality  
75 as proposed by Casellas et al. (2019b). In contrast, selection for reducing the residual  
76 variance may potentially generate even more concentrated lambing peaks. Thus, the  
77 canalization circular model, with genetic sources of variation affecting the residual  
78 effects, offers opportunities for selection to increase residual variance and reduce  
79 seasonality (Casellas et al. 2019b). However, the circular model has not been tested with

80 simulated data or compared with traditional linear models yet. The objectives of this  
81 study were (1) to compare von Mises (circular) and Gaussian linear models using  
82 simulated data and (2) to estimate variance components of lambing distribution in  
83 Canadian commercial sheep populations, including five major breeds.

## 84 **2 MATERIAL AND METHODS**

### 85 **2.1. Phenotypes**

86 Lambing date is normally recorded as day/month/year or as a Julian day relative  
87 to some start date. These can be converted to circular data by ignoring year. The lambing  
88 distribution over the year in sheep breeds represents the day of lambing rescaled from 0  
89  $\pi$  (January 1<sup>st</sup>) to  $2\pi$  (December 31<sup>st</sup>). The lambing dates phenotypic records were  
90 collected from the Canadian commercial sheep population, including five major breeds:  
91 Rideau Arcott (99,716 records), Romanov (52,577 records), Dorset (118,070 records),  
92 Suffolk (96,054 records) and Polypay (37,136 records). The phenotypes included  
93 historical records of lambing from 1986 to 2019. The pedigrees contained between  
94 12,807 (733 and 5,204) and 47,818 (2,908 and 15,554) animals (with known sires and  
95 ewes, respectively) depending on the breed (Table 1). The number of generations ranged  
96 between 19 (Dorset) and 24 (Romanov; Table 1). The number of flocks ranged between  
97 80 (Polypay) and 736 (Suffolk; Table 1).

### 98 **2.2. Simulated data**

99 To evaluate the performance of the circular model, simulated data were generated  
100 using the real pedigree of the Polypay breed with 37,136 lambing records and 12,807  
101 individuals (including 733 rams and 5,240 ewes). The data were simulated following the  
102 model described in section 2.4. (Material and Methods). Thus, the phenotypes included

103 two fixed effects (age at lambing with 6 levels and lambing interval effects with 5 levels)  
104 and three random factors (animal additive genetic, animal permanent environmental, and  
105 flock-year effects). The two fixed factors were simulated by sampling effects from a  
106 uniform distribution of  $[\pi-0.25, \pi+0.25]$  and  $[-0.25, 0.25]$ , respectively. Given our current  
107 lack of knowledge about circular data, arbitrary variance components were chosen to  
108 guarantee phenotypic records ranging approximately between  $0\pi-2$  to  $2\pi+2$  (i.e.,  
109 covering the circular parametric space  $[0\pi, 2\pi]$  and allowing to simulate a cyclical pattern  
110 effect). Thus, random effects were assumed to be drawn from independent Gaussian  
111 distribution with mean 0 and variance equal to 1.00, 0.35, 0.75 and 1.50 for animal  
112 additive genetic, animal permanent environmental, flock-year and residual effects,  
113 respectively (Table 2). The phenotypes were simulated under a linear assumption (i.e.,  
114 continuous from  $-\infty$  to  $+\infty$ ) and transformed to a circular space (i.e., 0 to  $2\pi$ ). This later  
115 were performed under the following conditions; when  $y < 0$  then  $y^* = y+2\pi$  and when  
116  $y > 2\pi$  then  $y^* = y-2\pi$ ; where  $y$  and  $y^*$  were the linear and circular records, respectively.  
117 A total of 10 replicates were generated under stochastic simulation.

### 118 **2.3. von Mises distribution**

119 The von Mises distribution is a continuous probability distribution on the circle  
120 developed by von Mises (1918) to address ‘cyclical error theory’ (Siegmond-Schultze,  
121 2006). This distribution is characterized by being wrapped around a circular parametric  
122 space from 0 to  $2\pi$ . It is a close approximation to the circular normal distribution (the  
123 circular analogue of the normal distribution) and is characterized by its appealing  
124 mathematical properties (Pewsey, 2013). The von Mises probability density function is  
125 given by:

126 
$$f(\mathbf{y}|\boldsymbol{\mu},\kappa)=\frac{e^{\kappa\times\cos(\mathbf{y}-\boldsymbol{\mu})}}{2\pi I_0(\kappa)},$$

127 where;  $y_i$  is the circular record,  $\mu$  is a measure of location (i.e., the mean),  $\kappa$  is the  
 128 concentration parameter and  $I_0(\kappa)$  is the modified Bessel function of order 0, i.e.,

129 
$$I_0(\kappa)=\sum_{i=0}^{\infty}\frac{(0.25\kappa^2)^i}{(i!)^2}.$$

130 The concentration parameter,  $\kappa$ , is analogous to the inverse of the variance of a normal  
 131 distribution. If  $\kappa=0$ , the distribution is uniform, and if  $\kappa$  is large, the distribution becomes  
 132 very concentrated about the mean (i.e., small variation). The circular variance of von  
 133 Mises distribution is equal to:

134 
$$var(y)=\frac{1-I_1(\kappa)}{I_0(\kappa)}$$

135 where;  $I_1(\kappa)$  is the modified Bessel function of order 1.

136 **2.4. Analytical models**

137 The analyses of lambing distribution were fitted as a trait of the ewe using the  
 138 circular mixed model developed by Casellas et al. (2019a,b) within the context of BLUP  
 139 (Best Linear Unbiased Prediction). This model relies on the von Mises distribution  
 140 (Fisher, 1993) and where the circular phenotype was parameterized as:

141 
$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_p\mathbf{p} + \mathbf{Z}_f\mathbf{f} + \mathbf{e}$$

142 where;  $\mathbf{y}$  was the vector of phenotypic records,  $\mathbf{b}$ ,  $\mathbf{a}$ ,  $\mathbf{p}$  and  $\mathbf{f}$  were fixed effects, additive  
 143 genetic, permanent environmental of the ewe and flock-year random effects with  
 144 pertinent incidence matrices ( $\mathbf{X}$ ,  $\mathbf{Z}_a$ ,  $\mathbf{Z}_p$  and  $\mathbf{Z}_f$ ), and  $\mathbf{e}$  was the vector of the residual  
 145 effects. Within the circular paradigm, the  $\mathbf{e}$  refers to the circular residual term and, under  
 146 a standard Bayesian approach, the joint posterior distribution of all unknown parameters  
 147 in the model was defined as:



148  $p(\mathbf{b}, \mathbf{p}, \mathbf{a}, \mathbf{f}, \kappa | \mathbf{y}) \propto p(\mathbf{y} | \mathbf{b}, \mathbf{p}, \mathbf{a}, \mathbf{f}, \kappa) p(\mathbf{b}) p(\mathbf{a} | \mathbf{A}, \sigma_a^2) p(\mathbf{A}, \sigma_a^2) p(\mathbf{p} | \sigma_p^2) p(\sigma_p^2) p(\mathbf{f} | \sigma_f^2) p(\sigma_f^2) p(\kappa),$

149 where;  $\mathbf{A}$  was the numerator relationship matrix. The likelihood of circular data becomes:

150 
$$p(\mathbf{y} | \mathbf{b}, \mathbf{a}, \mathbf{p}, \mathbf{f}, \kappa) = \prod_{i=1}^n \frac{e^{\kappa \cos(y_i - (\mathbf{x}_i \mathbf{b} + \mathbf{z}_{a;i} \mathbf{a} + \mathbf{z}_{p;i} \mathbf{p} + \mathbf{z}_{f;i} \mathbf{f}))}}{2\pi I_0(\kappa)},$$

151 where;  $\mathbf{x}_i$ ,  $\mathbf{z}_{a;i}$ ,  $\mathbf{z}_{p;i}$  and  $\mathbf{z}_{f;i}$  were appropriate row vectors of incidence inherent to the  $i_{th}$   
 152 circular record ( $y_i$ ). Moreover, random effects,  $\mathbf{a}$ ,  $\mathbf{p}$  and  $\mathbf{f}$  follow a multivariate normal  
 153 (MVN) distribution:

154 
$$p(\mathbf{a} | \mathbf{A}, \sigma_a^2) = \text{MVN}(0, \mathbf{A} \sigma_a^2),$$

155 
$$p(\mathbf{p} | \sigma_p^2) = \text{MVN}(0, \mathbf{I} \sigma_p^2) \text{ and}$$

156 
$$p(\mathbf{f} | \sigma_f^2) = \text{MVN}(0, \mathbf{I} \sigma_f^2),$$

157 where;  $\mathbf{I}$  was an identity matrix. The variances (i.e.,  $\sigma_a^2$ ,  $\sigma_p^2$  and  $\sigma_f^2$ ) are scaled inverse  
 158  $\chi^2$  distributions with hyper-parameters  $\nu$  and  $S^2$ . For  $\kappa$  and  $\mathbf{b}$  parameters, a priori  
 159 distributions can be stated as uniform between appropriate bounds:

160 
$$p(\kappa) = \frac{1}{\gamma} \text{ (if } \kappa \in [0, \gamma] \text{ and 0 otherwise) and}$$

161 
$$p(\mathbf{b}) = \frac{1}{2\pi} \text{ (if } \mathbf{b} \in [0, 2\pi] \text{ and 0 otherwise).}$$

162 where;  $\gamma$  is a large enough boundary for  $\kappa$ .

163 An additional model to analyze circular records in-depth is the extended circular  
 164 mixed model with heterogeneous variance (canalization circular model) developed by  
 165 Casellas et al. (2019a,b). This alternative model parameterized the dispersion parameter  
 166 of von Mises distribution,  $\kappa$ , as follows:

167 
$$\kappa_i = e^{\mathbf{x}_i \boldsymbol{\beta} + \mathbf{z}_{a;i} \boldsymbol{\alpha} + \mathbf{z}_{p;i} \boldsymbol{\rho} + \mathbf{z}_{f;i} \mathbf{f}},$$

168 where;  $\boldsymbol{\beta}$ ,  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\rho}$  and  $\mathbf{f}$  were fixed effects, additive genetic, permanent environmental of the  
 169 ewe and flock-year factors affecting the dispersion parameter of the residual term,

170 respectively. Thus, this model assumes the presence of additive genetic effects (as fixed  
171 and other random effects) affecting the residual term (i.e., effects originated/controlling  
172 heterogeneous environmental variance; see Casellas et al. (2019a,b) for details). For  
173 lambing date, the ram may be a significant source of variation reflecting when the breeder  
174 put the ram with the ewes, however, analysed as a trait if the ewe, the ram effect will  
175 remain in the residual.

## 176 **2.5. Analytical process**

177 The analytical model accounted for year effect (from 1986 to 2019, i.e., 45 years),  
178 ewe age at lambing (<3, 3, 4, 5, 6 and >6 years) and lambing interval (first lambing, <210  
179 days, between 210 and 239 days, between 240 and 329 days, and  $\geq 330$  days). The number  
180 of levels for random effects for each sheep breed are presented in Table 1. The unknown  
181 parameters in the model were sampled from their marginal posterior distributions by  
182 Metropolis-Hastings (Metropolis et al., 1953) and Gibbs sampling (Gelfand and Smith,  
183 1990). Details about the Markov chain Monte Carlo algorithm are described in Casellas  
184 et al. (2019a,b). Three independent Monte Carlo Markov chains with 600,000 iterations  
185 were run for each analysis, where the first 100,000 iterations were discarded as burn-in.  
186 In addition to the mixed circular model (with homogenous variance) and canalization  
187 circular model (with heterogeneous variance), a Gaussian linear model was implemented  
188 in the analyses of sheep data. Models were compared in terms of goodness of fit by the  
189 deviance information criterion (DIC; Spiegelhalter et al., 2002).

## 190 **3 RESULTS AND DISCUSSION**

### 191 **3.1. Phenotypic distribution of lambing date**

192           The distributions of lambing dates of the purebred Canadian commercial  
193 populations are presented in Figure 1. The five breeds analyzed displayed a large peak  
194 of lambings between 0 and 1 radians (January–February on a calendar basis), with the  
195 exception of Rideau Arcott breed which presented more concentrated lambing records  
196 between 1 and 2 radians (March–April; Figure 1). In addition, most breeds showed  
197 lambings across the whole year although within a particular flock, different pattern were  
198 observed such as lambings being concentrated on specific periods depending on the  
199 flock, year and breed. The main difference was observed in the Suffolk breed where the  
200 incidence of parturitions was reduced (mostly null) between 3 and 4 radians (June–July–  
201 August). The Suffolk breed showed more concentrated peaks of lambing (Figure 1),  
202 which is also evidenced by the small estimated circular variance (0.33) compared to other  
203 breeds (0.83, 0.75, 0.70 and 0.67 for Romanov, Polypay, Dorset and Rideau Arcott  
204 breeds, respectively). It is important to mention that a circular variance (Fisher, 1993)  
205 close to 0 corresponds to small variation, and a value of 1 implies maximum variation  
206 (i.e., uniform distribution). The low variance of the Suffolk population is a breed  
207 characteristic itself, as it is known by its generally poor out-of-season fertility (Dufour,  
208 1974; Jeffcoate et al., 1984) when compared to Polypay, for example (Notter, 2002),  
209 which has better out-of-season fertility (Hulet et al., 1984). In general, lambing  
210 distribution in the purebred Canadian commercial populations showed different ranges  
211 of phenotypic variances.

### 212 **3.1. Simulated data**

213           Results from analysis of simulated data showed the potential for using the von  
214 Mises model to infer effects and variance components of circular data. In Table 2, the

215 estimated genetic variance and the two environmental variances with the circular model  
 216 were similar to the simulated values. For example, among the 10 replicates, the genetic  
 217 variance ranged between 0.94 and 1.15, being close to the simulated genetic variance  
 218 (i.e., 1 radians<sup>2</sup>). Similar behavior was observed for the linear model when continuous  
 219 records were analyzed (i.e., the non transformed dates). However, when the linear model  
 220 was applied to the transformed dates, the model was not able to capture the variance  
 221 components simulated; all the parameters were underestimated with respect to the  
 222 simulated values (i.e., 0.09 versus 1.00 for  $\sigma_a^2$ , 0.08 versus 0.35 for  $\sigma_p^2$  and 0.07 versus  
 223 0.75 for  $\sigma_f^2$ ). Moreover, the residual variance from the linear model compared with the  
 224 circular model was overestimated (2.49 versus 1.50; Table 2). The poor performance of  
 225 the linear model was given to the assumed normal distribution of the effects in a linear  
 226 scale. The use of von Mises distribution in the circular model, overcame these problems.  
 227 Notice that the accuracy of predicting values was moderately high, e.g., the Pearson  
 228 correlation coefficients between simulated and estimated genetic effects were ~80%. The  
 229 circular model relies on using trigonometric functions in the von Mises distribution (e.g.,  
 230  $\cos(x)=\cos(x+2\pi)$ ).

231 In addition, the circular variance based on Fisher (1993;  $1-I_1(\kappa)/I_0(\kappa)$ ) was lower  
 232 than the simulated one. However, using a more linear approach, i.e.,  $\text{error} =$   
 233  $\arccos\left(\cos\left(y_i - (\mathbf{x}_i\mathbf{b} + \mathbf{z}_{a,i}\mathbf{a} + \mathbf{z}_{p,i}\mathbf{p} + \mathbf{z}_{f,i}\mathbf{f})\right)\right)$ ; the residual variance was equal to the  
 234 simulated one. Thus, both estimates must be taken with caution when interpreting the  
 235 results under a circular paradigm. On the other hand, the circular model had better  
 236 goodness-of-fit in terms of DIC when compared to the linear model. Thus, reductions up  
 237 to 4,739.60 and 18,268.20 DIC units were observed for the circular model respect the

238 linear model with circular and continuous records, respectively (Table 2). Notice that  
239 differences between models  $> 3$  DIC units are assumed as statistically relevant and  
240 suggest a better model fit and a lesser degree of complexity (Spiegelhalter et al., 2002).  
241 All the results support the importance of using the von Mises model due to its realistic  
242 framework for circular data.

### 243 **3.1. Variance components of lambing distribution in purebred Canadian** 244 **populations**

245 Variance components of the lambing distribution using both circular and linear  
246 models for the purebred Canadian populations are summarized in Table 3. First of all,  
247 the circular model gave better goodness-of-fit than the linear model for the circular data,  
248 thus, displaying smaller values of DIC units with reductions ranging from 22,412.40  
249 (Polypay breed) to 92,962.14 (Dorset breed). Interestingly, no significant additive  
250 genetic variance was captured by the linear model in any of the sheep breeds. More  
251 specifically, a  $h^2 \leq 0.01$  was observed for Suffolk and Romanov breeds and  $\leq 0.003$  for  
252 Rideau Arcott, Dorset and Polypay breeds. However, different genetic variances were  
253 captured by the circular model, being 1.26 ( $h^2=0.19$ ), 0.33 ( $h^2=0.12$ ), 0.10 ( $h^2=0.03$ ), 0.05  
254 ( $h^2=0.02$ ) and 0.001 radians<sup>2</sup> ( $h^2=0.001$ ) for Romanov, Rideau Arcott, Polypay, Dorset  
255 and Suffolk breeds, respectively (Table 3). First, these results showed the importance of  
256 using the von Mises model for circular data to estimate the genetic background of  
257 seasonality of lambing date in sheep Canadian populations. Second, the different genetic  
258 variances observed across breeds suggest a genetic predisposition to advance or postpone  
259 the lambing season (or its distribution pattern) with different degrees in some sheep  
260 breeds. In addition, the variances of the permanent ewe effects showed similar

261 magnitudes of genetic variance, where the largest and null variance was observed in  
262 Romanov and Suffolk, respectively (Table 3). The repeatabilities observed were 0.35,  
263 0.22, 0.05, 0.03 and 0.001 for Romanov, Rideau Arcott, Dorset, Polypay and Suffolk  
264 breeds, respectively.

265 On the other hand, the largest variances were captured by the flock-year effect, was  
266 3.34 (Romanov breed; Table 3), corresponding to 51.06% of the total phenotypic  
267 variance. This result showed the strong influence of management of the flock on lambing  
268 season for Canadian sheep populations. Moreover, permanent ewe and additive genetic  
269 sources of variation were mostly null under a linear model, the variance for flock-year  
270 effect ranged from 20.81% (Rideau Arcott breed;  $\sigma_f^2=0.57$  and  $\sigma_e^2=2.15$ ) to 31.95%  
271 (Dorset breed;  $\sigma_f^2=1.25$  and  $\sigma_e^2=2.78$ ) of the total phenotypic variance. It is important to  
272 mention that the use of flock-year-season effect instead of flock-year **may** generate a  
273 confounding factor between the observations and the flock-year-season effect when  
274 analyzing this trait. Notice that the seasons (corresponding to the date of lambing) and  
275 the observations (date of lambing) in some way are the same, thus, the flock-year-season  
276 effect **should be taken with caution when analyzing with this** type of traits.

277 Regarding the canalization circular model (heterogeneous variance model), this  
278 model did not converge properly. In general, the results presented here provide a genetic  
279 component to delay the peak of reproduction, and taking into consideration the strong  
280 correlation of direct and residual effect reported by Casellas (et al. 2019b) using  
281 canalization circular model, this genetic **variance opens** opportunities for selecting for  
282 seasonality as well. The genetic basis of seasonality was different across breeds, the  
283 genetic variance **being negligible** in Suffolk and ranging up **to** a heritability of 0.19 for

284 Romanov breed. Nevertheless, these reported variances and heritabilities represent novel  
285 results when compared to the traditional linear model, to further study and better  
286 understand this trait and the von Mises models. **Despite the linear model fails to identify**  
287 **additive genetic variance with the definition of lambing date (1:365) used in this study,**  
288 **we did not discard the existent of other transformations that may increase the**  
289 **performance of the linear model for this trait.** Finally, the results obtained in this study  
290 emphasize the relevance of implementation of appropriate methods - von Mises model -  
291 for circular data to accurately estimate the variance components of circular records.

## 292 **ACKNOWLEDGEMENTS**

293 The authors acknowledge the data provided by the member organizations of the  
294 Canadian Sheep Genetic Evaluation System. This study was funded by the Ontario Agri-  
295 Food Innovation Alliance [Ontario Ministry of Agri-culture, Food and Rural Affairs  
296 (OMAFRA)] (Guelph, Canada) and the Ontario Sheep Farmers (Guelph, Canada).

## 297 **CONFLICT OF INTEREST STATEMENT**

298 The authors declare no conflict of interest.

## 299 **DATA AVAILABILITY STATEMENT**

300 Data that support the findings of this study were provided by the Canadian Sheep  
301 Genetic Evaluation System. A subset of real data is available at  
302 <https://github.com/SamirIdLahoucine/CircularDATA>

## 303 **REFERENCES**

304 Arroyo, L. J., Gallegos-Sanchez, J., Villa-Godoy, A., Berruecos, J. M., Perera, G., &  
305 Valencia. J. (2007). Reproductive activity of Pelibuey and Suffolk ewes at 19 degrees  
306 north latitude. *Anim Reprod Sci*, 102, 24–30.doi:10.1016/j.anireprosci.2006.09.025

307 Battey, N. H. (2000). Aspects of seasonality. *J. Experimental Botany*, 51, 1769–  
308 1780.doi:10.1093/jexbot/51.352.1769

309 Begall, S., Červený, J., Neef, J., Vojtěch, O., & Burda, H. (2008). Magnetic alignment  
310 in grazing and resting cattle and deer. *Proc Natl Acad Sci USA*, 105, 13451–  
311 13455.doi:10.1073/pnas.0803650105

312 Cancho-Candela, R., Andrés-de Llano, J. M., & Ardura-Fernández, J. (2007). Decline  
313 and loss of birth seasonality in Spain: analysis of 33421731 birth over 60 years. *J Epidem  
314 Community Health*, 61, 713-718.doi:10.1136/jech.2006.050211

315 Casellas, J., Id-Lahoucine, S., & Martín de Hijas-Villalba, M. (2019b). Analysis of  
316 lambing distribution in the Ripollesa breed. II. Systematic and genetic sources of  
317 variation. *Animal*, 13, 2140–2145.doi:10.1017/S1751731119000375

318 Casellas, J., Martín de Hijas-Villalba, M., & Id-Lahoucine, S. (2019a). Analysis of  
319 lambing distribution in the Ripollesa breed. I. development and comparison of circular  
320 von Mises models. *Animal*, 13, 2133–2139.doi:10.1017/S1751731119000363

321 Chemineau, P., Bodin, L., Migaud, M., Thiéry, J. C., & Malpoux, B. (2010).  
322 Neuroendocrine and genetic control of seasonal reproduction in sheep and goats. *Reprod  
323 Domest Anim.*, 45, 42-49.doi:10.1111/j.1439-0531.2010.01661.x

324 Chemineau, P., Daveau, A., Cognié, Y., Aumont, G., & Chesneau, D. (2004). Seasonal  
325 ovulatory activity exists in tropical Creole female goats and Black Belly ewes subjected  
326 to a temperate photoperiod. *BMC Physiol*, 4, 12. Published 2004 Aug  
327 27.doi:10.1186/1472-6793-4-12



328 Chemineau, P., Malpoux, B., Brillard, J. P., & Fostier, A. (2007). Seasonality of  
329 reproduction and production in farm fishes, birds and mammals. *Animal*, 1, 419-  
330 423.doi:10.1017/S1751731107691873

331 Notter, D. R. (2002). Opportunities to Reduce Seasonality of Breeding in Sheep by  
332 Selection. *Sheep & Goat Research J.*, 17:3.

333 Dufour, J. J. (1974). The duration of the breeding season of four breeds of sheep. *Can J*  
334 *Anim Sci.*, 54, 389-392.

335 Erdem, E., & Shi, J. (2011). ARMA based approaches for forecasting the tuple of wind  
336 speed and direction. *Appl Energy*, 88, 1405-1414.doi:10.1016/j.apenergy.2010.10.031

337 Fisher, N. I. (1993). *Statistical analysis of circular data*. Cambridge University Press,  
338 Cambridge, United Kingdom.

339 Gao, K. S., Chia, F., Krantz, I., Nordin, P., & Machin, D. (2006). On the application of  
340 the von Mises distribution and angular regression methods to investigate the seasonality  
341 of disease onset. *Stat Med*, 25, 1593-1618.doi:10.1002/sim.2463

342 Gelfand, A., & Smith, A. F. M. 1990. Sampling based approaches to calculating marginal  
343 densities. *J Am Stat Assoc.*, 85, 398-409.

344 Gustafson, C. L., & Partch, C. L. (2014). Emerging models for the molecular basis of  
345 mammalian circadian timing. *Biochemistry*, 54, 134-149.doi:10.1021/bi500731f

346 Hulet, C. V., Ercanbrack, S. K., & Knight, A. D. 1984. Development of the Polypay  
347 breed of sheep. *J Anim Sci.*, 58, 15-24.doi:10.2527/jas1984.58115x

348 Ibáñez-Escriche, N., Sorensen, D., Waagepetersen, R., & Blasco, A. 2008. Selection for  
349 environmental variation: a statistical analysis and power calculations to detect response.  
350 *Genetics*, 180, 2209-2226.doi:10.1534/genetics.108.091678

351 Jeffcoate, I. A., Rawlings, N. C., & Howell, W. E. (1984). Duration of the breeding  
352 season and response to reproductive manipulation in five breeds of sheep under northern  
353 prairie conditions. *Theriogenology*, 22, 279-290.doi:10.1016/0093-691x(84)90485-0

354 Landler, L., Ruxton, G. D., & Malkemper, E. P. (2018). Circular data in biology: advice  
355 for effectively implementing statistical procedures. *Behav Ecol Sociobiol*, 72,  
356 128.doi:10.1007/s00265-018-2538-y

357 Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953).  
358 Equations of state calculations by fast computing machines. *J Chem Phys.* 21: 1087-  
359 1092.doi:10.1063/1.1699114

360 Mooney, J. A., Helms P. J., & Joliffe, I. T. (2003). Fitting mixtures of von Mises  
361 distributions: a case study involving sudden infant death syndrome. *Comput Stat Data*  
362 *Anal.*, 41, 505–513.

363 Mulder, K., & Klugkist, I. (2017). Bayesian estimation and hypothesis tests for a circular  
364 Generalized Linear Model. *J Mathematical Psychology*, 80, 4-  
365 14.doi:10.1016/j.jmp.2017.07.001

366 Notter, D. R., & Cockett, N. E. (2005). Opportunities for detection and use of QTL  
367 influencing seasonal reproduction in sheep: a review. *Genet Sel Evol.*, 37(1), S39-  
368 S53.doi:10.1051/gse:2004031

369 Pewsey, A. (2013). *Circular statistics in R (First edition.)*. Oxford University Press.

370 SanCristobal-Gaudy, M., Elsen, J. M., Bodin L., & Chevalet, C. (1998). Prediction of the  
371 response to a selection for canalisation of a continuous trait in animal breeding. *Genet*  
372 *Sel Evol.*, 30, 423-451.

373 Shimatani, I. K., Yoda, K., Katsumata, N., & Sato, K. (2012). Toward the quantification  
374 of a conceptual framework for movement ecology using circular statistical modeling.  
375 PLoS One. 7:e50309.doi:10.1371/journal.pone.0050309

376 Siegmund-Schultze, R. (2006). Probability in 1919/20: the von Mises-Pólya-  
377 Controversy. Arch Hist Exact Sci., 60, 431–515.doi:10.1007/s00407-006-0112-x

378 Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linder, A. 2002. Bayesian  
379 measures of model complexity and fit. J Royal Stat Soc B, 64, 583-  
380 639.doi:10.1111/1467-9868.00353

381 von Mises, R. (1918). Über die ‘Ganzzahligkeit’ der Atomgewichte und verwandte  
382 Fragen. Physikalische Zeitschrift, 19, 490-500.

383 Yenikoye, A. (1984). Annual variations in estrual behavior, rate and possibilities for  
384 ovulation in Peulh ewes from Niger. Reprod Nutr Dev., 24, 11-19.

385

386

387

388 **Table 1.** The number of lambing records and individuals for each sheep breed.

389

390 **Table 2.** Simulated and estimated variance components using linear (Gaussian  
391 distribution) and circular (von Mises distribution) models (averages across 10 replicates).

392 <sup>†</sup>Parameter of concentration of von Mises distribution; <sup>‡</sup>±Standard deviation; <sup>§</sup>Minimum  
393 and maximum estimates across 10 replicates; DIC: Deviance information criterion.

394 <sup>¶</sup>Circular variance  $(1-I_1(\kappa)/I_0(\kappa))$ ; where  $I_1$  and  $I_0$  are the modified Bessel function of order

395 1 and 0, respectively); <sup>‡</sup>Based on the following residual effect:  $\arccos\left(\cos\left(y_i -$

396  $(\mathbf{x}_i\mathbf{b}+\mathbf{z}_{a,i}\mathbf{a}+\mathbf{z}_{p,i}\mathbf{p}+\mathbf{z}_{f,i}\mathbf{f})\right)\right)$ .

397

398 **Table 3.** Estimated variance components of lambing distribution for five different  
399 sheep breeds (three chains each breed).

400  $\sigma_a^2$ : genetic variance,  $\sigma_p^2$ : permanent environmental variance;  $\sigma_f^2$ : flock-year variance,  $\kappa$ :

401 parameter of concentration of von Mises distribution; DIC: Deviance information  
402 criterion.

403 <sup>†</sup>Circular variance  $(1-I_1(\kappa)/I_0(\kappa))$ ; where  $I_1$  and  $I_0$  are the modified Bessel function of order

404 1 and 0, respectively); <sup>‡</sup>Based on the following residual effect:  $\arccos\left(\cos\left(y_i -$

405  $(\mathbf{x}_i\mathbf{b}+\mathbf{z}_{a,i}\mathbf{a}+\mathbf{z}_{p,i}\mathbf{p}+\mathbf{z}_{f,i}\mathbf{f})\right)\right)$ .

406

407 **Figure 1.** Lambing distribution across the year for five purebred Canadian sheep  
408 commercial populations.

409