PREDICTION OF STILLBORN PIGLETS FROM MULTIPAROUS SOWS †

[PREDICCIÓN DE LECHONES NACIDOS MUERTOS DE CERDAS MULTÍPARAS]

D.A. Domínguez-Olvera[1], J.G. Herrera-Haro[1*], J.R. Bárcena-Gama[1], M.E. Ortega-Cerrilla[1], F.E. Martínez-Castañeda[2], A. Rouco-Yañez[3], M.A. Ortíz-Heredia[1] and N.A. Rogers-Montoya[4]

[1]Programa de Recursos Genéticos y Productividad-Ganadería, Colegio de Postgraduados. Kilómetro 36.5 carretera México-Texcoco, Montecillo, Texcoco, Edo. Méx, México. C.P. 56230. Email: haro@colpos.mx
[2]Instituto de Ciencias Agropecuarias y Rurales, Universidad Autónoma del Estado de México, Toluca, Edo. Méx, México. C.P. 56295.
[3]Facultad de Veterinaria, Universidad de Murcia. 5 de Avenida Teniente Flomesta, Murcia, España. C.P. 30003.
[4]Facultad de Medicina Veterinaria y Zootecnia, Universidad Nacional Autónoma de México. Av. Universidad #3000, Ciudad Universitaria, Coyoacán, México, C.P. 04510.
**Corresponding author*

SUMMARY

Background: Assisting sows during parturition reduces the number of stillborn piglets caused by anoxia. However, in industrial settings with a large number of animals, the capacity for assistance is limited. The development of predictive models based on existing data can enable farms to anticipate stillbirths in sows. **Objective:** To develop a predictive model to identify factors affecting the presence of stillborn piglets (PSbP), estimate the probability of their occurrence, and establish a classification criterion accordingly. **Methodology:** Data from 2 415 farrowings in 822 sows (Landrace, Yorkshire, and their crossbreeds) were analyzed. Five variables relating to the current farrowing and five variables related to the preceding one were examined. Our study used cross-validation (groups = 5), modeling the response variable (PSbP, 1: presence, 0: absence). **Results:** The only factor shown to have a negative effect ($p<0.01$) on PSbP was litter weight at birth, while litter size at birth and parity (number of farrowings) were seen to have a positive effect ($p<0.01$). PSbP prevalence during training and testing were 0.297 and 0.296 respectively. The model's estimated probability levels were 0.311 during training and 0.303 during testing, indicating an accurate probability estimation. When categorizing using the optimal cutoff point of 0.395, the predictive efficiency as measured by the area under the Receiver Operating Characteristic (ROC) curve was 0.846 for training and 0.813 for testing. **Implications:** Implementing this model of information-management software could make it possible to provide swift, efficient technical assistance to sows in need, with a high level of predictive efficiency. **Conclusions:** The probabilistic model described here based on a Bayesian approach and adjusted based on a categorization criterion showed effective predictive efficiency in the prediction of stillborn piglets.

Key words: Probabilistic model; logistic regression; cross-validation; *Sus scrofa domesticus*.

RESUMEN

Antecedentes: En cerdas, la asistencia al momento del parto ayuda a reducir los lechones nacidos muertos por anoxia; en condiciones industriales, donde el número de animales es grande, esta capacidad de asistencia se ve

† Submitted May 27, 2024 – Accepted January 28, 2025. http://doi.org/10.56369/tsaes.5658

ISSN: 1870-0462.

ORCID = D.A. Domínguez-Olvera: https://orcid.org/0000-0002-7220-0242; **J.G. Herrera-Haro:** https://orcid.org/0000-0002-6177-8647; **M.E. Ortega-Cerrilla:** https://orcid.org/0000-0001-9971-5169; **J.R. Barcena-Gama:** https://orcid.org/0000-0002-0039-9424; **F.E. Martínez-Castañeda:** https://orcid.org/0000-0003-0168-921X; **M.A. Ortíz-Heredia:** https://orcid.org/0000-0003-4170-7996; **N.A. Rogers-Montoya:** https://orcid.org/0000-0002-6570-7571

limitada. Generar modelos predictivos a partir de la información disponible, permite predecir la respuesta de cerdas respecto a la presencia de lechones nacidos muertos. **Objetivo:** Generar un modelo de predicción para determinar los factores que afectan la presencia de lechones nacidos muertos (PNM), estimar la probabilidad que dicho fenómeno se manifieste y, con base en eso, establecer un criterio de clasificación. **Metodología:** Se analizaron los datos de 2415 partos de 822 cerdas provenientes de distintas granjas, con razas Landrace, Yorkshire y sus cruzas; cinco variables relacionadas con el parto actual, cinco con el desempeño del parto anterior; con validación cruzada (grupos = 5) y se modeló la variable respuesta (PNM, 1: presencia y 0: de otra manera). **Resultados:** Únicamente el peso de la camada al nacimiento tuvo un efecto negativo ($p<0.01$) en la presencia de nacidos muertos; mientras el tamaño de la camada al nacimiento y número de parto tuvieron un efecto positivo ($p<0.01$) en la PNM. La prevalencia de PNM en el entrenamiento fue 0.297 y 0.296 en la prueba, mientras que la probabilidad estimada por el modelo en el entrenamiento fue 0.311 y 0.303 en la prueba, una estimación buena de la probabilidad, pero en la categorización con la obtención y uso del punto de corte óptimo de 0.395, la eficiencia predictiva con una área bajo la curva operativa del receptor de 0.846 para entrenamiento y 0.813 para prueba. **Implicaciones:** La implementación del modelo obtenido en los softwares de gestión de información permitirá dirigir puntualmente la asistencia técnica a las cerdas que lo requieran con una buena eficiencia. **Conclusiones:** El modelo probabilístico con aproximación bayesiana ajustado junto con el criterio de categorización presentó una eficiencia predictiva buena para predecir lechones nacidos muertos.
Palabras clave: Modelo probabilístico; regresión logística; validación cruzada; *Sus scrofa domesticus*.

INTRODUCTION

The increase that has been observed in pig-farm efficiency is attributed to technological progress at various levels of the pig-breeding system, including improvements in reproductive capacity (Caicedo *et al*., 2012), progress in the areas of genetics (Ghio and de la Sota, 2018) and nutrition (Gaillard *et al*., 2020), enhanced biosecurity protocols (López and Sánchez, 2019), and the adoption of good vaccination and medication practices (Maes *et al*., 2019). Recent research has focused on adapting farming practices to meet the unique needs of individual animals, guided by the principles of precision pig farming (Tzanidakis *et al*., 2021).

In this context, precision pig farming covers all the stages of the system, being based on the tenet that each animal has unique traits and requirements. This kind of farming aims to satisfy each animal's individual needs so as to significantly improve pig farming practices by integrating cutting-edge information technologies. To achieve the latter goal, precision pig farming has used integrated models that optimize the use of existing information. These advances facilitate customized herd management focused on individual needs (Pomar and López, 2018; Tzanidakis *et al*., 2021).

On sow farms, productive efficiency (i.e. the number of weaned piglets per sow) is directly influenced by piglet mortality during the birth, lactation and weaning phases, having a cumulative effect that significantly influences technical and economic outcomes on farms (Pomar and Pomar, 2005; Martínez-Castañeda and Perea-Peña, 2012; Stalder, 2017; Faccin *et al*., 2020; Maes *et al*., 2020).

Piglet mortality during birth is primarily caused by anoxia due to a disturbance in the natural birthing process that can lead to umbilical-cord rupturing or premature placenta detachment, both of which hinder proper delivery and ultimately lead to the piglet's death (Castillo and Vicente, 2016). Given the complex nature of the birthing process, the effective tackling of mortality due to anoxia at the farm-level poses a considerable challenge. Precision pig farming aims to focus technical support on sows that are at risk of delivering stillborn piglets (Pomar and López, 2018). This targeted approach is essential for reducing the number of stillbirths and requires predictive models to anticipate the latter (Vanderhaeghe *et al*., 2010).

Logistic regression is of importance in biology due to its ability to interpret binary phenomena that may be influenced by other variables, making it possible to predict probable stillbirths (Vanderhaeghe *et al*., 2010) by establishing a single binary response variable based on several explanatory variables. Predictive models have been developed in order to anticipate the occurrence of stillborn piglets and provide sows with the technical assistance that they require based on an understanding of the risk factors inherent in the birthing process (Vanderhaeghe *et al*., 2010).

In certain cases, the optimization of parameters from logistic models using Bayesian statistics can lead to significant improvements. This is so because there have been cases of overfitting in conventional logistic-regression models, leading to convergence failures (Londoño-Parra *et al*., 2018). Since such conventional models rely on frequency approximation for their fit, it is difficult to accurately estimate their regression parameters (Gelman *et al*., 2008). Furthermore, some authors have found that certain biological phenomena are predicted more accurately via logistic regression than via machine-learning methods (Nusinovici *et al*., 2020).

Given the need for models capable of accurately predicting variables that influence the technical performance of swine farms, the aim of the study was to develop a predictive model for identifying the factors influencing the occurrence of stillborn piglets, estimating its likelihood and establishing classification criteria.

MATERIALS AND METHODS

Data source

Data from 2,415 farrowings in 822 sows on farrow-to-wean pig farms were analyzed during the period 2010-2013. The qualitative variables studied were farm (FA) and breed or genetic line (BLN), while the quantitative variables studied were number of farrowings (NF), litter size at birth (LSB: only live-born piglets), number of stillborn piglets (NSbP), number of mummies (NMu: fetuses that died after the occurrence of bone calcification), litter weight at birth (LWB: only live-born piglets), litter weight and size at birth in the preceding farrowing (LWB-PF and LSB-PF respectively), weight and size of litter at weaning in the preceding farrowing (LWW-PF and LSW-PF respectively), and number of stillborn piglets in the preceding farrowing (NSbP-PF). The standard normal distribution was analyzed for all variables. The dichotomous variable, Presence of Stillborn Piglets (PSbP), was generated defined by PSbP = 1 when stillbirths occurred and PSbP = 0 when they did not occur.

Evaluation of information

Paired and independent Tukey means ($p<0.05$) were compared in order to determine: a) differences between the FA and BLN of the LSB, NSbP, NMu, LWB, LWB-PF, LSB-PF, LWW-PF, LSW-PF and NSbP-PF variables, so as to measure their effect and importance in the building of the model; b) differences between LSB, NSbP and LWB of the LSB-PF, NSbP-PF and LWB-PF respectively, in order to establish a paired-means effect, only observations with more than one farrowing were included (n=1,725), and c) differences among the different variables, LSB, NSbP, LWB, LWB-PF, LSB-PF, LSW-PF, LWW-PF and NSbP-PF, were included, based on the classification variable PSbP: 1. farrowings with stillborn piglets and 2. farrowings with no stillborn piglets.

Modelling

Based on the above variables and without taking stock of the FA or BLN groups, the logistic regression was adjusted via a Bayesian approximation method (Equation 1), using cross-validation where the response variable was PSbP and the modelled probability was the occurrence of stillbirths (P[Y=1|X]). The Bayesian strategy used was the one based on conditional probability, using Cauchy's non-informative prior distribution, that is recommended by Gelman et al. (2008), since, when estimated, the parameters have smaller associated variance using that approach, thus increasing the model's predictive power and avoiding overadjustment.

In order to avoid collinearity and guarantee a parsimonious model, the variables retained were determined using the stepwise regression procedure. Only those variables that had a significant effect ($p<0.05$) on the response variable, PSbP, were considered. The cross-validation carried out was for k-groups (k-fold cross-validation) as described by Berrar (2017).

To minimize bias in the generation of the model, the database was split up into five groups (k=5) so as to ensure that the latter kept the proportions of the response variable when randomizing (Kohavi, 1995); upon splitting up the database into the said random groups, without replacing any of the data, a proportion similar to that of the response variable (PSbP) was found among them. The groups underwent 20% of the observations with k iterations being used, k-1 groups being utilized for training (80% of all the observations), and the model's predictive efficiency being tested on the group that was not considered for training. The K-group-repetition method was not used until the arrangement with the least variance was found, since Molinaro et al. (2005) have shown that, instead of increasing predictive power, such repetitions only bring about slight reductions in variance.

$$P(Y = 1|X_i) = \frac{exp^{(\beta_0+\beta_1X_1+\beta_2X_2+\cdots+\beta_PX_P)}}{1 + exp^{(\beta_0+\beta_1X_1+\beta_2X_2+\cdots+\beta_PX_P)}} \quad Equation\ (1)$$

where P (Y=1|Xi) is the conditional probability of stillborn piglets occurring at birth (PSbP), and assumes a value of 1, given Xi, which is the vector of realization of the individual, i, in the p variables; β0 is the intercept, βP is the regression coefficient for the p variable, and Xp is the value of the p variable.

To determine the predictive power of the models, the factors considered were: 1) the value of the Receiver Operating Characteristic (ROC) area below the curve, which ranges from zero to one, where values close to one have better predictive power; 2) the standard error of the mean; 3) Spearman correlation between the observed PSbP and those estimated based on the model, using the optimal cut-off point (OCP), which is defined as the optimal likelihood for establishing a decision-making borderline at which PSbP assumes a value of 1 based on the probability estimate; 4) evaluation of the observed mean using the predicted mean.

RStudio for R software (Posit team, 2023) was used to process and analyze the data and create the graphs. The following packages were used for cross-validation and model-adjustment ("caret" version 6.0-94; "leaps" version 3.1; and "BGLR" version 1.1.1), for graphs ("MASS" version 7.3-53.1 and "rgl" version 0.106.8), and for the model-performance evaluation ("ROCR" version 1.0-11; "leaps" version 3.1; "agricolae" version 1.3-6, and "PRROC" version 1.3.1.).

RESULTS AND DISCUSSION

Evaluation of information

The average NF was 2.94 farrowings, being the only variable that did not follow normality, having a Gamma distribution. The average LWB was 14.52 kg, 8.22% higher than that reported for F1 Large Whitex Landrace sows by Rendón-del Águila *et al*. (2017), but lower than the 20.79 kg reported by Salazar (2018) in a meta-analysis-based study that included various breeds and genetic lines.

Upon comparing paired means, no differences (p>0.05) were found between the current farrowing variables and preceding farrowing (PF) variables. The average NSbP was 4.11% of all LSB, the average weight per animal at birth was 1.29 kg, higher than that reported by Knol *et al*. (2002) for Large White x Pietrain piglets, indicating that weights under 1.125 kg impede the animals from maintaining an adequate body temperature, which explains the high LSW, with only 10% mortality during lactation (Table 1). Given the homogeneous management practices across the analyzed farms, including artificial insemination for reproduction, a 146-day sow cycle (116 days of gestation, 23 days of lactation, and a 7-day weaning-to-service interval), and similar genetics, no differences were observed between FA and BLN in any of the variables studied. Litters averaged 12 live-born piglets with a 16% mortality rate during lactation and a weaning weight of 7.36 kg. The 130-day fattening phase categorized pigs into pre-starting, starting, growth, development, and finishing stages based on their physiological phase, achieving an average live weight of 110 kg.

The two groups defined by the PSbP, showed differences (p<0.05) between their means in most of the variables. The NF and LSB were higher in the group with stillborn piglets (PSbP =1), while the LWB was higher in the group without stillborn piglets (PSbP =0), with no differences (p>0.05) between the NMu in both groups.

In the group with PSbP =1, the average NSbP was 1.53 piglets; Vanderhaeghe *et al*. (2010) found the same ratio when they studied and split up 532 farrowings. The population without stillbirths had an average litter size at birth of 14.7 piglets, with an average mortality level of 2.02 piglets. A similar pattern was observed in the preceding farrowing variables, although the only categorical variable was considered from the variables relating to the current farrowing, there were no differences (p<0.05) between groups in LWW-PF and LSW-PF (Table 1).

Model

The above information suggests the need for a model that is of help in determining the likelihood of sows having PSbP. In the probabilistic modelling with a Bayesian approach, only NF, LSB and LWB were significant (p<0.05) in the PSbP. Baxter *et al*. (2008), reported that under similar management conditions, piglets with an average weight of 1.17 kg were stillborn.

Table 1. Descriptive statistics and comparative means of the variables studied.

	Variable	Mean	SEM	PSbP = 0	PSbP =1
Current farrowing		n=1725		n=1699	n=716
	NF* (n)	2.94	0.04	2.80[b]	3.26[a]
	LSB (n)	11.19[x]	0.06	10.78[b]	12.18[a]
	LWB (kg)	14.52[y]	0.07	14.69[a]	14.13[b]
	NSbP (n)	0.46[z]	0.02	0.00	1.53
	NMu (n)	0.48	0.03	0.46[a]	0.51[a]
Preceding farrowing		n=1725		n=1203	n=522
	LSB-PF (n)	11.18[x]	0.07	10.92[b]	11.76[a]
	LWB-PF (kg)	14.47[y]	0.09	14.34[b]	14.76[a]
	LSW-PF (n)	9.00	0.05	9.00[a]	9.00[a]
	LWW-PF (kg)	53.83	0.27	54.13[a]	53.14[a]
	NSbP-PF (n)	0.45[z]	0.03	0.40[b]	0.56[a]

*Gamma distribution. a,b: Mean values per column with different literal are statistically different ($p \leq 0.05$). x,y,z: Mean values per column with different literal are statistically different ($p \leq 0.05$). SEM: Standard Error of the Mean. NF: Parity number. LSB: Litter size at birth. LWB: Litter weight at birth. NSbP: Stillborn piglets. NMu: Number of mummies. PSbP =0: Population without stillbirths. PMN=1: Population with stillbirths.

The average estimated values of the regression coefficients for the NF, LSB and LWB variables resulting from the cross-validation process were 0.102, 0.259 and -0.162 respectively (Table 2). The βNF and βLSB estimators had positive values, meaning that these variables had a positive effect on the probability of PSbP being equal to 1, and a negative effect on the number of live-born piglets, with the βLWB coefficient being the only one showing the negative effect on the PSbP response variable. Baxter *et al*. (2008) report the same trend with large LSB having the highest number of stillbirths, beginning, in most cases, after the 10th piglet. This is due to hypoxia at birth, a problem arising from the cumulative effect of uterine contractions and most commonly observed in the last piglets in the litter (Alonso-Spilsbury *et al*., 2007).

Moreover, Canario *et al*. (2007) mention that this problem has worsened due to the genetic selection focused in obtaining lean growth and increasing sow prolificity, both of which result in less mature piglets. The variation of all estimators of the model was less than 13%, meaning that, as expected, there was little variation in the data within the k=5 groups. Equation 2 is the definitive equation of the modelling, which retains the average of the estimators obtained from the cross-validation (Pérez-Planells *et al*., 2015).

Table 2. Regression coefficients and their confidence interval (α=0.05) obtained from the cross-validation (k-groups = 5) of significant variables (*p<0.05*) with the presence of stillbirths.

Estimator	Mean
Intercept	-2.103 [-2.204, -2.001]*
NF	0.102 [0.085, 0.118]*
LSB	0.259 [0.237, 0.287]*
LWB	-0.14 [-0.162, -0.117]*

*with α=0.05.

As mentioned in this study, the logistic-regression model for predicting the occurrence of stillborn piglets is defined by only three variables. It's expression is shown in the following equation:

$$P(Y=1|X) = \frac{exp^{(-2.103+0.102X_1+0.259X_2-0.140X_3)}}{1+exp^{(-2.103+0.102X_1+0.259X_2-0.140X_3)}} \quad Equation\ (2)$$

where P (Y=1|X) is the probability of stillbirths (PSbP) and takes on the value of one: X1= parity, X2= litter size at birth and X3= litter weight at birth.

The results of the model obtained by cross-validation and Spearman correlation between estimated probability and the response variable are stable (Table 3). Almost identical average correlations can be observed between the test and training groups (0.547 and 0.495 respectively); the estimated Mean Squared Error did not differ from that of the test, due to the size of the sample, which in all cases was over 400. Bates *et al*. (2023) assert that a sample size of 100 significantly improves estimation of the Standard Error of the Mean. Basically, upon comparing the training and test results for each group, the only thing that varied was the confidence interval, which was bigger for the test populations, but still with minimal differences. Pérez-Planells *et al*. (2015) established that the Mean Squared Error gives us the error associated with the method, confirming the model's strong predictive power.

The area below the receiver operating characteristic curve of the receptor (ABC-ROC) constitutes a robust option for describing and comparing classification models (Polo and Miot, 2020). The results of the cross-validation were 0.846 (Table 1) and 0.813, values deemed to be good (Nahm, 2022). Upon comparing our results with those pertaining to the area below the curve of different probabilistic models, Pinto and Sanchez Bayle (2017) obtained values of between 0.65 and 0.88 when they created probabilistic models for diagnosing bacterial infections in lactating piglets with fever, managing to achieve good predictions from the said models. While Lynam et al. (2020) obtained ABC-ROC values of over 95% when using logistic regression to model Type-1 and Type-2 diabetes in human adults.

The optimal cut-off point (OCP) shows the limit of the optimal probability for establishing the borderline where the result of the said probability stemming from the model is defined. Presence or absence of stillborn piglets (0, 1) is a value obtained from the ROC graph, being the point at which the sensitivity and specificity are suitable because of the use of laxer criteria to increase sensitivity results in a compensation whereby specificity decreases (Nahm, 2022).

The average value of the model, different from the borderline of 0.5 used by Londoño *et al*. (2018), was 0.395 in the training and 0.407 in testing, indicating the probability at which stillbirths are expected. The test's ABC-ROC was similar to that of the training when used at the OCP, but lower when not taken into account. The average PSbP classification was accurate in 78.55% of the cases, when predicting both stillbirths and the lack thereof (Table 3).

Figure 2 shows the distribution of the 2,415 farrowings among the model's three significant variables (NF, LSB and LWB). On the left is the observed occurrence of PSbP and on the right the predicted occurrence of PSbP based on the model with a OCP of 0.395. Both the predicted and observed occurrences can be seen to follow the same trend whereby PSbP increases in step with increases in LSB. Moreover, while it can be seen that PSbP also increases in step with the NF, the graph does not show any clear trend with regard to LWB, although the negative impact on the response variable had already been defined in the model. The differences between the events observed and the ones estimated based on the model may be due to other variables not contemplated in the latter, such as farrowing interval, age of the gilt at insemination and gilt birth phenotype (Patterson *et al.* 2020; Carrión-López *et al*. 2022; van den Bosch *et al.* 2022). Baxter *et al*. (2008) found important differences in crown-to-rump length when studying the comparative numbers of stillbirths. While Oliviero *et al*. (2010) concluded that the relationship between farrowing length and the number of stillbirths is clear, explaining that the decisive factor in farrowing length is back-fat thickness. Although neither the breed nor its crossbreeds were significant ($p>0.05$) in the model, Leenhouwers *et al*. (1999) found that purebred sows have more stillborn piglets per litter than crossbred ones, although this could depend on litter size and farrowing length.

Vanderhaeghe *et al*. (2013) found 17 non-infectious factors, ranging from genetics and characteristics of the sow to the features and handling of the piglet itself that might have some sort of impact on the occurrence of stillbirths, making it clear that this model needed to include more environmental variables in order to be more precise. Upon graphing the probability of PSbP based on the model (Equation 2), with the combination of the variables (Figure 3), one can see the trend of the latter on the response variable. The observations in red are the events that it is expected will have a value of one as a result of the categorization of the established OCP. While the combination of LWB and NF shows an accumulation of observations with stable LWB that is not shown on the graph. On the other hand, the combination of LSB and NF shows a clear tendency for a higher PSbP in step with increases in both variables. When one charts the distribution of the observations stemming from the combination of LSB and LWB, a clear upward trend in PSbP can be seen, defined by increase in the former and low values for the latter (Figure 3).

Table 3. Model-performance criteria in the cross-validation and its confidence intervals (α=0.05), with k-groups = 5.

Estimator	Mean
Training (80%)	n=1,932
Observed mean	0.296 [0.289, 0.303]
Predicted mean	0.311 [0.304, 0.318]
Spearmen correlation*	0.547 [0.537, 0.557]
SME	0.015 [0.008, 0.021]
ABC-ROC	0.846 [0.781, 0.851]
OCP	0.395 [0.383, 0.407]
Test (20%)	n=483
Observed mean	0.297 [0.267,0.327]
Predicted mean	0.303 [0.296, 0.309]
Spearman correlation*	0.495 [0.471, 0.519]
MSE	0.019 [0.001,0.037]
ABC-ROC	0.813 [0.802, 0.824]
OCP	0.407 [0.359, 0.455]
Effectiveness, %	78.55 [0.767, 0.803]

*Correlation between the stillbirths observed and the estimated probability (*p<0.0001*). MSE: Mean squared error. ABC-ROC: Area below the Receiver Operating Characteristic of the Receptor. OCP: Optimal Cut-off Point. Effectiveness = proportion of true cases (for 0 or 1) in accordance with the classification criteria established based on the OCP and their probability.

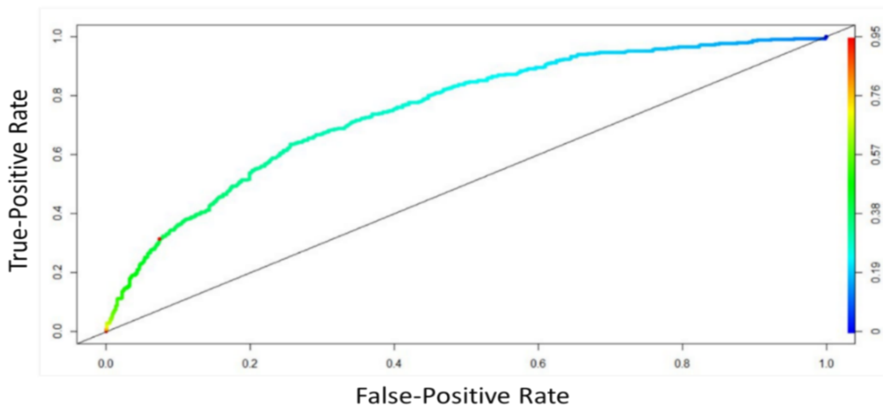

Figure 1. Third-iteration receiver operating characteristic curve (ROC) showing the best features for predicting stillbirths (Area Below the Curve = 0.846). In the right-hand scale, which uses different colors to show probability levels, red indicates the optimal cut-off point (0.395), for establishing the borderline that specifies based on the probability, the probable presence or absence of stillborn piglets.

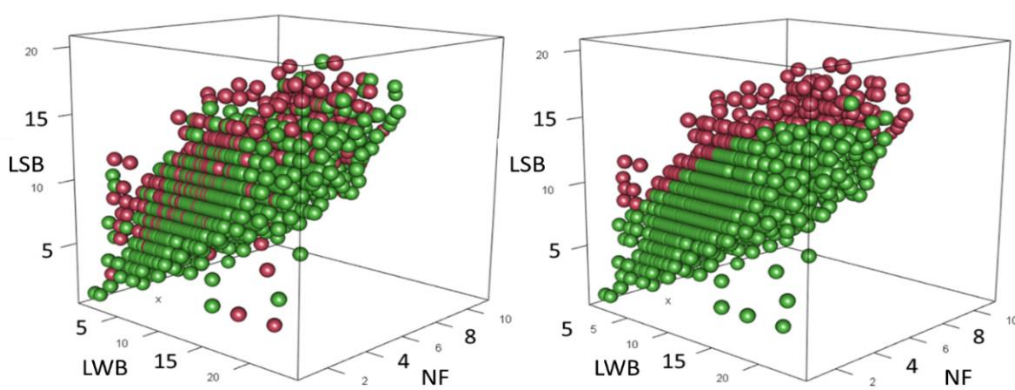

Figure 2. Observed distribution of births (left-hand graph) and predicted distribution of births (right-hand graph) as a function of the variables litter weight at birth (LWB), litter size at birth (LSB), and number of farrowings (NF), including the occurrence of stillborn piglets (red dots).

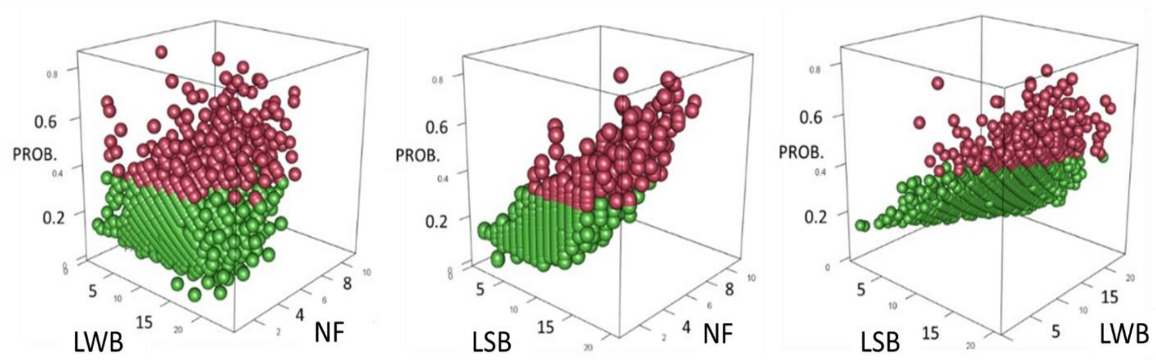

Figure 3. Probability distribution (predicted graphs) of presence of stillborn piglets (PSbP) as a function of the variables: litter weight at birth (LWB), litter size at birth (LSB) and farrowing number (NF) of the studied population; the predicted occurrence of PSbP (red dots) was determined based on the optimal cut-off point (0.395).

CONCLUSION

It is concluded that the adjusted probabilistic model with Bayesian approximation has good predictive efficiency, indicating that sows with heavy litters but few piglets are less likely to suffer stillbirths. While parity (number of farrowings) had an impact on the likelihood of stillbirths, this impact was not as obvious as that of litter size and weight at birth. The performance-in-previous-farrowings (preceding farrowing variables) does not affect the probability of stillbirths. The optimal-cut-off-point variable was suitable for purposes of creating the classification.

Acknowledgments
The authors wish to thank Pigknows©, Roman Martinez and Carlos Cedenio who kindly provided the information for this study.

Funding. This study was funded by the research project ESYCA-144971 of the Consejo Mexiquense de Ciencia y Tecnología and CONAHCyT México granted to Daniel Alonso Domínguez Olvera.

Conflicts of interest/competing interests. The authors declare no conflict of interest.

Compliance with ethical standards. The Ethics Committee of the UNAM “Facultad de Medicina Veterinaria y Zootecnia” approved the research protocol and information gathering tools. Due to the nature of the study, no animals were used to obtain data.

Data availability. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author contribution statement (CRediT). **D.A. Domínguez-Olvera**: Conceptualization, Methodology, Research, Data analysis, Data curation, Writing – Original Draft, Project administration - funding acquisition; **J.G. Herrera-Haro**: Conceptualization, Methodology, Research, Writing – Original Draft, Supervision, Project administration - funding acquisition; **J.R. Bárcena-Gama:** Methodology, Research, Writing –Review & Editing; **M.E. Ortega-Cerrilla**: Research, Writing –Review; **F.E. Martínez-Castañeda**: Conceptualization, Research, Writing – Original Draft, Supervision; **A. Rouco-Yañez**: Research, Writing –Review & Editing. **M.A. Ortíz-Heredia**: Research, Writing –Review & Editing; **N.A. Rogers-Montoya**: Writing –Review & Editing.

REFERENCES

Alonso-Spilsbury, M., Ramirez-Necoechea, R., Gonzalez-Lozano, M., Mota-Rojas, D. and Trujillo-Ortega, M.E., 2007. Piglet survival in early lactation: a review. *Journal of Animal and Veterinary Advances*, 6(1), pp. 76-86. http://docsdrive.com/pdfs/medwelljournals/javaa/2007/76-86.pdf [Accessed 21 March, 2024].

Bates, S., Hastie, T. and Tibshirani, R., 2023. Cross-validation: what does it estimate and how well does it do it? *Journal of the*

American Statistical Association, pp. 1-12. https://doi.org/10.1080/01621459.2023.2197686

Baxter, E.M., Jarvis, S., D'eath, R.B., Ross, D.W., Robson, S.K., Farish, M., Nevison, I.M., Lawrence A.B. and Edwards, S.A., 2008. Investigating the behavioural and physiological indicators of neonatal survival in pigs. *Theriogenology*, 69 (6), pp. 773-783. https://doi.org/10.1016/j.theriogenology.2007.12.007

Berrar, D., 2017. Confidence curves: an alternative to null hypothesis significance testing for the comparison of classifiers. *Mach Learn*. 106, pp. 911–949. https://doi.org/10.1007/s10994-016-5612-6

Caicedo, Q.W., Valle, R.S. and Velázquez, R.F., 2012. Diagnóstico participativo para la producción porcina en el medio periurbano y rural del cantón Pastaza Ecuador. *Revista Electrónica de Veterinaria*, 13(8), pp. 1-9. Available at: https://www.redalyc.org/pdf/636/63624429006.pdf [Accessed 21 March, 2024].

Canario, L., Pere, M.C., Tribout, T., Thomas, F., David, C., Gogué, J. and Le Dividich, J., 2007. Estimation of genetic trends from 1977 to 1998 of body composition and physiological state of Large White pigs at birth. *Animal*, 1(10), pp. 1409-1413. https://doi.org/10.1017/S1751731107000766

Carrión-López, M.J., Orengo, J., Madrid, J., Vargas, A., and Martínez-Miró, S., 2022. Effect of sow body weight at first service on body status and performance during first parity and lifetime. *Animals*, 12(23), p. 3399. https://doi.org/10.3390/ani12233399

Castillo, M.T. and Vicente, S.C., 2016. Puntos clave en el manejo en maternidad. *Anaporc: Revista de la Asociación de Porcinocultura Científica*, 13(134), pp. 22-28. Available at: https://www.archivo-anaporc.com/app/download/6467914011/Puntos+clave+en+el+manejo+de+madres.pdf?t=1494408309 [Accessed 21 march 2024].

Faccin, J.E., Laskoski, F., Hernig, L.F., Kummer, R., Lima, G.F., Orlando, U.A. and Bortolozzo, F.P., 2020. Impact of increasing weaning age on pig performance and belly nosing prevalence in a commercial multisite production system. *Journal of Animal Science*, 98(4), skaa031. https://doi.org/10.1093/jas/skaa031

FAO, 2022. FAOSTAT-Cultivos y productos de ganadería. Available at: https://www.fao.org/faostat/es/#data/QCL [Accessed 02 march 2023].

Gaillard, C., Brossard, L. and Dourmad, J.Y., 2020. Improvement of feed and nutrient efficiency in pig production through precision feeding. *Animal Feed Science and Technology*, pp. 268. https://doi.org/10.1016/j.anifeedsci.2020.114611

Gelman, A., Jakulin, A., Pittau, M.G. and Su, Y.S., 2008. A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 2(4), pp. 1360-1383. https://doi.org/10.1214/08-AOAS191

Ghio, M. and De la Sota, M.N.L., 2015. Actualización sobre mejoramiento genético porcino en el mundo y en la República Argentina. *Semiárida*, 25(2), pp. 72-73. Available at: https://ojs24.unlpam.edu.ar/index.php/semiarida/article/view/2543 [Accessed 21 march 2024].

Knol, E.F., Leenhouwers, J.I. and Van der Lende, T., 2002. Genetic aspects of piglet survival. *Livestock Production Science*, 78(1), pp. 47-55. https://doi.org/10.1016/S0301-6226(02)00184-7

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence. *American Association for Artificial Intelligence*, 2(12), pp. 1137-1143. Available at:

https://www.ijcai.org/Proceedings/95-2/Papers/016.pdf [Accessed 21 march 2024]

Leenhouwers, J.I., Van Der Lende, T. and Knol, E.F., 1999. Analysis of stillbirth in different lines of pig. *Livestock Production Science*, 57(3), pp. 243-253. https://doi.org/10.1016/S0301-6226(98)00171-7

Londoño-Parra, J.S., Cabrera-Torres, K.R. and González-Hurtado, M.I., 2018. Modelo de predicción probabilística de deterioro en jamón de cerdo cocido. *Vitae*, 25(2), pp. 64-74. https://doi.org/10.17533/udea.vitae.v25n2a02

Lynam, A.L., Dennis, J.M., Owen, K.R., Oram, R.A., Jones, A.G., Shields, B.M. and Ferrat, L.A., 2020. Logistic regression has similar performance to optimized machine learning algorithms in a clinical setting: application to the discrimination between type 1 and type 2 diabetes in young adults. *Diagnostic and Prognostic Research*, 4(1), pp. 1-10. https://doi.org/10.1186/s41512-020-00075-2

Maes, D.G., Dewulf, J., Piñeiro, C., Edwards, S. and Kyriazakis, I., 2020. A critical reflection on intensive pork production with an emphasis on animal health and welfare. *Journal of Animal Science*, Volume 98, Issue Supplement 1, pp. S15–S26.

Martínez-Castañeda, F.E. and Perea-Peña, M., 2012. Estrategias locales y de gestión para la porcicultura doméstica en localidades periurbanas del Valle de México. *Agricultura, Sociedad y Desarrollo*, 9(4), pp. 411-425. Available at: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1870-54722012000400003&lng=es&tlng=es [Accessed 21 march, 2022].

Molinaro, A.M., Simon, R. and Pfeiffer, R.M., 2005. Prediction error estimation: a comparison of resampling methods. *Bioinformatics*, 21(15), pp. 3301-3307. https://doi.org/10.1093/bioinformatics/bti499

Nahm, F.S., 2022. Receiver operating characteristic curve: overview and practical use for clinicians. *Korean Journal of Anesthesiology*, 75(1), pp. 25-36. https://doi.org/10.4097/kja.21209

Nusinovici, S., Tham, Y.C., Yan, M.Y.C., Ting, D.S.W., Li, J., Sabanayagam, C., Wong T.Y. and Cheng, C.Y., 2020. Logistic regression was as good as machine learning for predicting major chronic diseases. *Journal of Clinical Epidemiology*, 122, pp. 56-69. https://doi.org/10.1016/j.jclinepi.2020.03.002

Oliviero, C., Heinonen, M., Valros, A. and Peltoniemi, O., 2010. Environmental and sow-related factors affecting the duration of farrowing. *Animal Reproduction Science*, 119(1-2), pp. 85-91. https://doi.org/10.1016/j.anireprosci.2009.12.009

Patterson, J., Bernardi, M.L., Allerson, M., Hanson, A., Holden, N., Bruner, L., Pinilla, J.C., and Foxcroft, G., 2020. Associations among individual gilt birth weight, litter birth weight phenotype, and the efficiency of replacement gilt production. *Journal of Animal Science*, 98(11), p. skaa331. https://doi.org/10.1093/jas/skaa331

Pérez Planells, L., Delegido Gómez, J., Rivera-Caicedo, J.P. and Verrelst, J., 2015. Análisis de métodos de validación cruzada para la obtención robusta de parámetros biofísicos. *Revista Española de Teledetección*, 44, pp. 55-65. http://dx.doi.org/10.4995/raet.2015.4153

Pinto, E.V. and Sánchez-Bayle, M., 2017. Creación de un modelo probabilístico de diagnóstico de infección bacteriana grave en lactantes febriles de 0 a 3 meses de vida. *Anales de Pediatría*, 87(6), pp. 330-336. https://doi.org/10.1016/j.anpedi.2017.02.003

Polo, T.C.F. and Miot, H.A., 2020. Aplicações da curva ROC em estudos clínicos e experimentais. *Jornal Vascular Brasileiro*, pp. 19.

https://doi.org/10.1590/1677-5449.200186

Pomar, J. and López, V., 2018. La porcicultura de precisión es una perspectiva innovadora para el futuro de la producción porcina (I). Albéitar: *Publicación Veterinaria Independiente*, (217), pp. 18-19. Available at: https://core.ac.uk/download/pdf/161804819.pdf [Accessed 21 march, 2024].

Pomar, J. and Pomar, C., 2005. A knowledge-based decision support system to improve sow farm productivity. *Expert Systems with Applications*, 29(1), pp. 33-40. https://doi.org/10.1016/j.eswa.2005.01.002

R Core Team., 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/

Rendón del Águila, J.U., Martínez-Gamba, R.G., Herradora Lozano, M.A. and Alonso-Spilsbury, M., 2017. Efecto del peso al nacer, tamaño de camada y posición en la ubre sobre el crecimiento de cerdos durante la lactancia y engorda. *Revista Mexicana de Ciencias Pecuarias*, 8(1), pp. 75-81. https://doi.org/10.22319/rmcp.v8i1.4317

Salazar, S., 2018. Factores que afectan la vitalidad de los lechones al momento del nacimiento. *Nutrición Animal Tropical*, 12(1), pp. 40-58. https://doi.org/10.15517/nat.v12i1.33670

Schober, P. and Vetter, T.R., 2021. Logistic regression in medical research. *Anesthesia and Analgesia*, 132(2), pp. 365. https://doi.org/10.1213/ane.0000000000005247

Stalder, K.J., 2017. Pork industry productivity analysis. National Pork Board Report. Available at: https://jygatech.com/wp-content/uploads/2021/05/PigmortalityStalderIowa.pdf [Accessed 21 march 2024].

Tzanidakis, C., Simitzis, P., Arvanitis, K. and Panagakis, P., 2021. An overview of the current trends in precision pig farming technologies. *Livestock Science*, pp. 249, 104530. https://doi.org/10.1016/j.livsci.2021.104530

van den Bosch, M., van de Linde, I.B., Kemp, B., and van den Brand, H., 2022. Disentangling litter size and farrowing duration effects on piglet stillbirth, acid–base blood parameters, and pre-weaning mortality. *Frontiers in Veterinary Science*, 9, p. 836202. https://doi.org/10.3389/fvets.2022.836202.

Vanderhaeghe, C., Dewulf, J., de Kruif, A. and Maes, D., 2013. Non-infectious factors associated with stillbirth in pigs: a review. *Animal Reproduction Science*, 139(1-4), pp. 76-88. https://doi.org/10.1016/j.anireprosci.2013.03.007

Vanderhaeghe, C., Dewulf, J., De Vliegher, S., Papadopoulos, G.A., de Kruif, A. and Maes, D., 2010. Longitudinal field study to assess sow level risk factors associated with stillborn piglets. *Animal Reproduction Science*, 120(1-4), pp. 78-83. https://doi.org/10.1016/j.anireprosci.2010.02.010