



## OPEN ACCESS

## EDITED BY

Flaviana Gottardo,  
University of Padua, Italy

## REVIEWED BY

Matthew R. Beck,  
United States Department of Agriculture  
(USDA), United States  
Temple Grandin,  
Colorado State University, United States

## \*CORRESPONDENCE

Dinu Gavojdian  
✉ gavojdian\_dinu@animalsci-tm.ro

RECEIVED 17 December 2023

ACCEPTED 19 January 2024

PUBLISHED 01 February 2024

## CITATION

Gavojdian D, Mincu M, Lazebnik T, Oren A,  
Nicolae I and Zamansky A (2024) BovineTalk:  
machine learning for vocalization analysis of  
dairy cattle under the negative affective state  
of isolation.

*Front. Vet. Sci.* 11:1357109.

doi: 10.3389/fvets.2024.1357109

## COPYRIGHT

© 2024 Gavojdian, Mincu, Lazebnik, Oren,  
Nicolae and Zamansky. This is an open-  
access article distributed under the terms of  
the [Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction  
in other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# BovineTalk: machine learning for vocalization analysis of dairy cattle under the negative affective state of isolation

Dinu Gavojdian<sup>1\*</sup>, Madalina Mincu<sup>1</sup>, Teddy Lazebnik<sup>2,3</sup>,  
Ariel Oren<sup>4</sup>, Ioana Nicolae<sup>1</sup> and Anna Zamansky<sup>4</sup>

<sup>1</sup>Cattle Production Systems Laboratory, Research and Development Institute for Bovine, Balotesti, Romania, <sup>2</sup>Department of Mathematics, Ariel University, Ariel, Israel, <sup>3</sup>Department of Cancer Biology, University College London, London, United Kingdom, <sup>4</sup>Tech4Animals Laboratory, Information Systems Department, University of Haifa, Haifa, Israel

There is a critical need to develop and validate non-invasive animal-based indicators of affective states in livestock species, in order to integrate them into on-farm assessment protocols, potentially via the use of precision livestock farming (PLF) tools. One such promising approach is the use of vocal indicators. The acoustic structure of vocalizations and their functions were extensively studied in important livestock species, such as pigs, horses, poultry, and goats, yet cattle remain understudied in this context to date. Cows were shown to produce two types of vocalizations: low-frequency calls (LF), produced with the mouth closed, or partially closed, for close distance contacts, and open mouth emitted high-frequency calls (HF), produced for long-distance communication, with the latter considered to be largely associated with negative affective states. Moreover, cattle vocalizations were shown to contain information on individuality across a wide range of contexts, both negative and positive. Nowadays, dairy cows are facing a series of negative challenges and stressors in a typical production cycle, making vocalizations during negative affective states of special interest for research. One contribution of this study is providing the largest to date pre-processed (clean from noises) dataset of lactating adult multiparous dairy cows during negative affective states induced by visual isolation challenges. Here, we present two computational frameworks—deep learning based and explainable machine learning based, to classify high and low-frequency cattle calls and individual cow voice recognition. Our models in these two frameworks reached 87.2 and 89.4% accuracy for LF and HF classification, with 68.9 and 72.5% accuracy rates for the cow individual identification, respectively.

## KEYWORDS

cattle, animal communication, affective states, vocal parameters, welfare indicators

## Introduction

Farm animal welfare is commonly defined as the balance between positive and negative emotions, where positive emotions are considered as the main indicators of a moral animal life (“a life worth living” concept), with most of the recent research body of literature outlining the importance of affective states in farmed animals’ health and wellbeing (1, 2). Non-human mammals’ affective states might vary in valence (positive to negative) and arousal levels (high

to low), having functional adaptations linked to behavioral decisions that facilitate individual survival and reproduction while promoting approaches toward rewards and avoidance (3, 4).

There is an evident need to develop valid non-invasive animal-based indicators of emotions in domestic animals in order to integrate them into future on-farm assessment protocols, potentially via the use of precision livestock farming (PLF) tools, such as novel sensors (5, 6). To this end, the use of bioacoustics to evaluate health, emotional states, and stress responses has been validated for some of the most important livestock species such as pigs (*Sus scrofa domestica*), goats (*Capra hircus*), horses (*Equus caballus*), and poultry (*Gallus gallus domesticus*). The research findings consistently show that vocal parameters differ substantially during positive and negative experiences (7–15). Consequently, these developments started to be implemented and used in commercial settings in order to automatically classify animal vocalizations and identify health issues. For instance, the AI-based solution SoundTalks® was introduced in pig farms to detect respiratory diseases. However, compared to the aforementioned species, there is a significant knowledge gap regarding cattle communication behavior (16, 17). A potential explanation of this might be that cattle have a lower incidence of emitting vocalizations (16, 18), especially alarm and pain-specific vocalizations, developed as an adaptive response of the species as prey animals in order to avoid the risk of alarming potential predators.

Domestic cattle vocalizations were shown to contain information on individuality, given the high levels of inter-cow variability in the acoustic characteristic of the vocalizations emitted under various contexts, as well as allowing facilitation of short- and long-distance interactions with herd-mates. This variability found in vocalizations produced by cattle allows for each animal to be identified by the “fingerprint” of their call (19–22). Cattle are highly gregarious and form complex social relationships, having a strong innate motivation for continuous social contact (23, 24), with isolation from conspecifics resulting in physiological changes such as increased heart rates, cortisol levels, ocular and nasal temperature, and an increase in vocalization production (22, 25). Furthermore, it was suggested that individual cattle vary in susceptibility to emotional stressors and challenges (26–28), with limited research being undertaken to evaluate the effects that isolation over prolonged periods of time has on vocalization response in adult cattle. Throughout a typical production cycle, dairy cows face a series of negative emotional challenges and stressors, such as separation from a calf immediately after calving, frequent regrouping based on production levels and lactation phase, re-establishing social hierarchy and dominance, frequent milking, isolation from herd-mates for insemination, pregnancy check-ups, being at high risks of developing metabolic disorders, isolation in sickness pens, etc.

Cattle are known to produce two types of vocalizations, which are modulated by the configuration of the supra-laryngeal vocal tract (21). The first type is low-frequency calls (LF), produced by the animal with the mouth closed or partially closed, used for close distance contact, and regarded as indicative of lower distress or positive emotions. The second type is open-mouth emitted high-frequency calls (HF), produced for long-distance communication, and indicating higher arousal emotional states, generally associated with negative affective states (22, 29).

In domestic ungulates, individuality was proven to be encoded in a wide range of vocal parameters, most evidently in the F0-contour (15, 30), amplitude contour and call duration (31), as well as in filter-related vocal parameters including formant frequencies (21). Individuality expression was shown to be distinct for each call type (22, 30), with individual differences in cattle high-frequency calls being attributed mainly to sound formants (21, 22), while vocalization formants are being modulated in turn by the caller vocal tract morphology (32). Given that cattle can produce vocalizations with fundamental frequencies of over 1,000 Hz (30), which are more likely to occur during times of higher arousal affective states, it was hypothesized by the authors that high-frequency calls encode a larger amount of individuality information, than their low-frequency equivalents, due to their propagation over longer distances where vision and/or olfactory signaling are not possible.

Methods of studying animal vocal communication are becoming increasingly automated, with a growing body of research validating the use of both hardware and software that are capable of automatically collecting and processing bioacoustics data [reviewed by McLoughlin et al. (18)]. In this vein, Shorten and Hunter (33) found significant variability in cattle vocalization parameters, and suggested that such traits can be monitored using animal-attached acoustic sensors in order to provide information on the welfare and emotional state of the animal. Therefore, automated vocalization monitoring could prove to be a useful tool in precision livestock farming (18, 34, 35), especially as dairy farming systems become increasingly automated with wide-scale use of milking and feeding robots, all this having the potential to dynamically adjust the management practices while the number of animals per farm unit tends to increase.

Machine learning techniques are therefore increasingly applied in the study of cattle vocalizations. Some tasks that have been addressed to date include the classification of high vs. low frequency calls (33), ingestive behavior (35), and categorization of calls such as oestrus and coughs (34).

This study makes the following contributions to the investigation of cattle vocalizations using machine learning techniques. First of all, we present the largest dataset to date of ( $n = 20$ ) cows' vocalizations collected under a controlled “station” setting, exclusively for negative affective states. Furthermore, we develop two types of AI models: deep-learning-based and explainable machine-learning-based for two tasks: (1) classification of high and low-frequency calls, and (2) individual cow identification. Finally, we investigated the feature importance of the explainable models.

## Materials and methods

### Ethical statement

All experiments were performed in accordance with relevant guidelines and regulations. The experimental procedures and protocols were reviewed and approved by the Ethical Committee from the Research and Development Institute for Bovine, Balotesti, Romania (approval no. 0027, issued on July 11, 2022), with the isolation challenge producing exclusively temporary distress to cows.

## Subjects and experimental approach

The study was carried out at the experimental farm of the Research and Development Institute for Bovine in Balotesti, Romania. At the experimental facilities, cattle were managed indoors year-round (zero-grazing system), being housed under tie-stall conditions (stanchion barn) in two identical animal barns with a housing capacity of 100 heads/barn, having access to outdoor paddocks (14–16 m<sup>2</sup>/head) 10 h/day, between milkings (7:00–17:00). Cows had *ad libitum* access to water and mineral blocks, receiving a daily feed ration of 30 kg corn silage (37% dry matter, DM), 6 kg of alfalfa hay (dehydrated whole plant, 90.5% DM), and 6 kg of concentrates (88.3% DM). Concentrates were given immediately after morning and evening milkings (3 kg/feeding session); while silage and hay were offered at the feeding rails in the outside paddocks, assuring a feeding space of 0.7 m/cow. In total, 20 lactating adult multiparous cows of the Romanian Holstein breed were tested between August and September 2022. The Romanian Holstein breed (RH, national name Băltațu cu Negru Românească) belongs to the dairy Holstein-Friesian strain, with a current census of 264,000 cows, representing 22% of the breed structure in Romania (36). The RH originates in the 19th century, being the result of systematic crossbreeding between Friesian bulls imported from Denmark and Germany and local cattle, such as Romanian Spotted and Dobruja Red. The average milk yield for the RH breed ranges between 6,000 and 8,500 kg of milk/lactation, with adult body weights of cows ranging between 550 and 650 kg. The selection index for the RH breed is focused on milk yield (90%) and fertility related traits (10%) (37). In order to avoid bias and to have a homogeneous study group, cattle included in our research were of similar age (lactations II & III), were habituated previously to the housing system (min. 40 days in milk), and were comparable for body weight (619.5 ± 17.40 kg, mean ± SEM). Cows were individually isolated inside the two identical barns, remaining tethered at the neck to their stall (stall dimensions of 1.7/0.85 m), starting at 7:00 AM for 240 consecutive minutes post-milking when the rest of the herd members were moved to the outside paddocks. The two paddocks were in the immediate vicinity of the barns, having one lateral shared concrete compact wall, thus the animal that remained inside the barn was visually isolated, while being able to hear and communicate vocally with their herd-mates. The isolated cows had *ad-libitum* access to water throughout individual drinkers and fresh wheat straws bedding was provided for animal comfort. After the commencement of recordings, animal caretakers were restricted from access to barns, and human traffic and machinery noise production outside the two barns were limited as much as possible.

## Vocalization recordings

The vocalizations for this study were obtained using two identical directional microphones (Sennheiser MKH416-P48U3, frequency response 40–20,000 Hz, max. sound pressure level 130 dB at 1 kHz, producer Sennheiser Electronic®, Wede-mark, Germany) attached to Marantz PMD661 MKIII digital solid-state recorders (with file encryption, WAV recording at 44.1/48/96 kHz, 16/24-bit, recording bit rates 32–320 kbps, producer Marantz Professional®, United Kingdom). The microphones were directed toward the animal using tripods placed on the central feeding alleys at a

distance of 5–6 m from the cows. For shock and noise reduction, Sennheiser MZW 415 ANT microphone windshields were used. After the end of each experimental day, vocal recordings were saved as separate files in the WAV uncompressed format, at 44.1 kHz sampling rate and a 16-bit amplitude resolution. Despite the fact that all 20 multiparous lactating cows were isolated and recorded for 240 min post-milking, under identical conditions, not all cows vocalized with a similar frequency during the trials, resulting in the analysis of 1,144 vocalizations (57.2 vocalizations per cow, ranging between 33 and 90 vocalizations per cow), out of which 952 were high-frequency vocalizations (HF) and 192 low-frequency vocalizations (LF). All sounds included in our investigation had undergone quality control check, while looking for clear, under- and un-saturated vocalizations, without combined environmental noises such as rattling equipment, chains clanging, or wind. Vocalizations were visualized on spectrograms using the fast Fourier transform method, at window lengths of 0.03 s, time steps of 1,000, frequency steps of 250, dynamic range of 60 dB, and a view range between 0 and 5,000 Hz (Figure 1).

Vocalization recordings were then analyzed using Praat DSP package v.6.0.31 (38), as well as previously developed custom-built scripts (10, 15, 39–41), for the automatic extraction of the 23 acoustic features of each vocalization, with the vocal parameters studied and their definitions being presented in Table 1, the output data being exported to Microsoft Excel for further analysis.

## Classification models

We developed two different computational frameworks of the following types:

- i Explainable model—a pipeline that uses as features the 23 vocal parameters described in Table 1, which have been studied in the context of cattle vocalizations. By using features that are highly relevant to our domain, we increase the explainability of our pipeline, allowing for the study of the feature importance of our model.
- ii Deep learning model, which uses learned features and operates as a “black box” that is not explainable. This model was expected to be more flexible and to have increased performance.

The explainable framework was based on the TPOT (42), AutoSklearn (43), and H2O (44) automatic machine learning libraries. Namely, we assumed a dataset represented by a matrix  $x \in R^{n \times m}$  and a vector  $y \in R^n$ , where  $n$  is the number of rows and  $m$  is the number of features in the dataset. Notably, we used the features described in Table 1, which made the model more explainable, as the contribution of each feature to the model's prediction could be computed. This dataset was divided into training and testing sets, such that the first has 80% of the data and the latter the remaining 20%, divided randomly. The training cohort was used to train the model and the testing cohort was used to evaluate its performance. Moreover, we randomly picked 90% of the training dataset each time for  $r = 50$  times, making sure each value was picked at least half of the times. For each of these cohorts, we first obtained a machine learning pipeline from TPOT, AutoSklearn, and H2O aiming to optimize the following loss function:  $\sum_{i=1}^k \left( \frac{a^i + f^i}{2k} \right)$  where  $a^i$  and  $f^i$  are the  $i_{th}$  model's

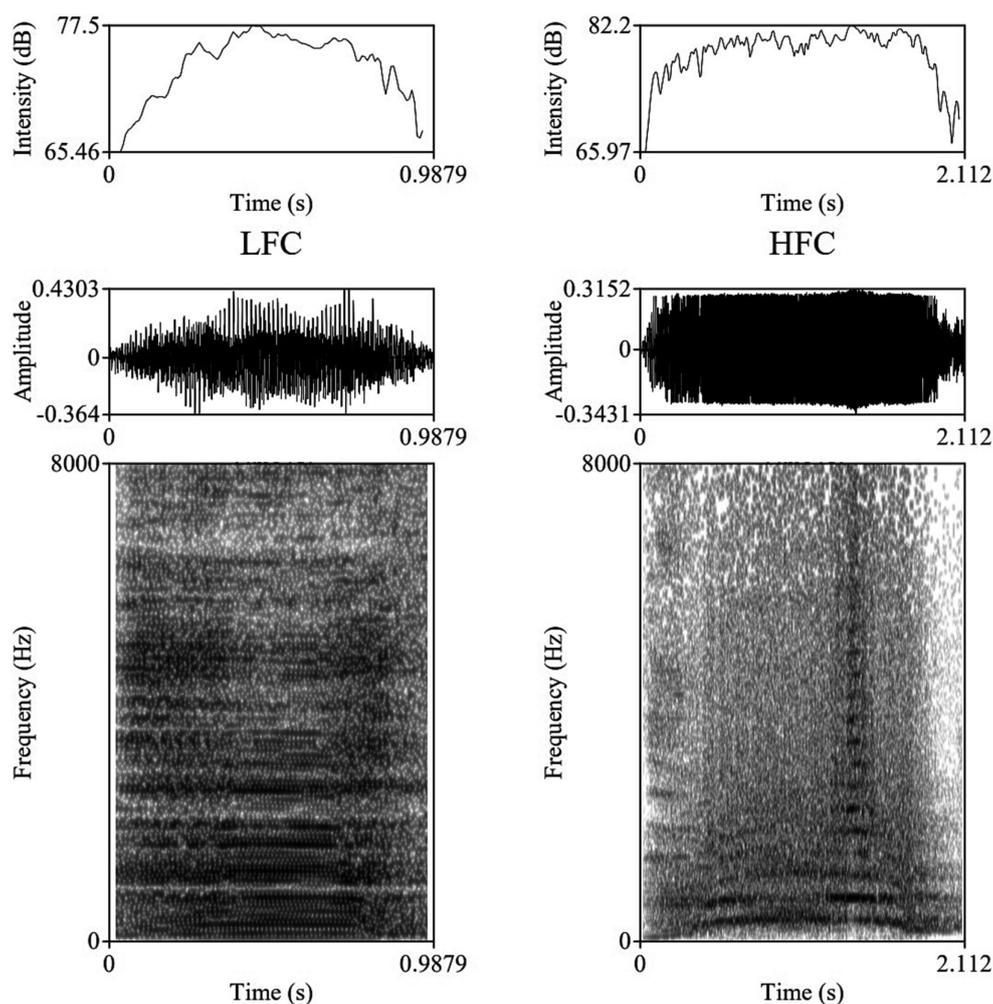


FIGURE 1

Example of a Low Frequency Call (LFC) left-side and a High Frequency Call (LFC) right-side with intensity contours (above), oscillograms (middle), and spectrograms [below; fast Fourier transform (FFT) method, at window lengths of 0.03 s, time steps of 1,000, frequency steps of 250, dynamic range of 60 dB, and a view range between 0 and 5,000 Hz] of typical vocalizations produced by cows during the isolation challenge.

instance accuracy and  $F_1$  scores, respectively, where  $k$  was the number of  $k$ -folds in the cross-validation analysis (45). Once all three models were obtained, we used all three of them to generate another cohort containing their predictions and the corresponding  $y$  value. These were then used to train an XGboost (46) model for the final prediction. For the hyperparameter tuning of the XGboost, we took advantage of the grid-search method. Finally, a majority vote between the  $r$  instances was used to determine the final model's prediction. We reported the results for  $k=5$  fold cross-validation over the entire dataset.

The DL framework was adopted from Ye and Yang (47) which proposed a deep-gated recurrent unit (GRU) neural network (NN) model, combining a two-dimensional convolution NN and recurrent NN based on the GRU cell unit that gets as input the spectrogram of the audio signal. Generally, the two-dimensional convolution NN is used as a feature extraction component, finding spatio-temporal connections in the signal which than is being fed into the recurrent NN that operates as a temporal model able to detect short-and long-term connections in this feature space over time, these being effectively the rules for the voice identifications. For the

hyperparameters of the model such as batch size, learning rate, optimization, etc., we adopted the values from Ye and Yang (47).

## Results

In this section, we examined the data obtained and outlined the performance of the proposed explainable and DL models for the two different tasks based on the collected dataset. First, we provided a descriptive statistical analysis of the obtained dataset and its properties. Secondly, we present the performance of the models in classification between high and low-frequency calls. Finally, we present the models' ability to identify each cow according to its vocalizations, divided into low, high, and all low + high vocalizations. Table 2 summarizes the results of the explainable and DL models' performance in separating between the high and low-frequency calls. The results are shown as mean  $\pm$  standard deviation for  $k=5$  fold. Importantly, we made sure that the train and test cohorts had vocalizations from both classes at each fold and that the ratio between the classes was kept between folds. Both

TABLE 1 Abbreviations and definitions of the 23 vocal parameters studied (21, 29).

Abbreviation/Unit of measure	Definition
F0Mean (Hz)	Mean F0* frequency value across the call
F0Max (Hz)	Maximum F0 frequency value across the call
F0Min (Hz)	Minimum F0 frequency value across the call
F0Range (Hz)	Difference between minimum and maximum F0
Q25 (Hz)	Frequency value at the upper limit of the first quartiles of energy (below 25%)
Q50 (Hz)	Frequency value at the upper limit of the second quartiles of energy (below 50%)
Q75 (Hz)	Frequency value at the upper limit of the third quartiles of energy (below 75%)
Fpeak (Hz)	Frequency of peak amplitude
Sound duration (s)	Duration of the call from start to end, measured on the oscillogram
AMVar (dB/s)	Cumulative variation in amplitude divided by the total call duration
AMRate (s-1)	Number of amplitude modulations in a certain time frame
AMExtent (dB)	Mean-to-mean peak variation of each amplitude modulation
Harmonicity (dB)	Degree of acoustic periodicity, also called harmonic-to-noise ratio—higher values indicate more tonal voice
F1Mean (Hz)	Mean frequency value of the first formant
F2Mean (Hz)	Mean frequency value of the 2nd formant
F3Mean (Hz)	Mean frequency value of the 3rd formant
F4Mean (Hz)	Mean frequency value of the 4th formant
F5Mean (Hz)	Mean frequency value of the 5th formant
F6Mean (Hz)	Mean frequency value of the 6th formant
F7Mean (Hz)	Mean frequency value of the 7th formant
F8Mean (Hz)	Mean frequency value of the 8th formant
Formant dispersal (Hz)	Minimum spacing of the formants
Wiener entropy mean	Spectral flatness of a sound, calculated as the ratio of a power spectrum's geometric mean to its arithmetic mean measured on a logarithmic scale

TABLE 2 The high- and low-frequency calls (LF and HF) classifier models' performances.

Model	Train set accuracy	Test set accuracy
Explainable	89.9 ± 2.2%	87.2 ± 4.1%
Deep learning	91.5 ± 2.6%	89.4 ± 3.8%

The results are shown as mean standard deviation for  $k = 5$  fold.

models achieved good results with almost nine out of 10 correct detections. One can notice that the DL model outperforms the explainable model. One explanation for this is that the DL is more expressive and therefore captures more complex dynamics, which are not necessarily expressed by the features provided to the explainable model (see Table 2). Figure 2 presents the features' importance of the explainable model for the high and low-frequency calls calculated by reducing one feature from the input and calculating its influence on the model's performance. One can notice that AMvar, AMrate, AMExtent, Formant dispersal, and the Wiener entropy mean are the most important features, with a joint importance of 55.36%.

Table 3 summarizes the results of the explainable and DL models' individual cow identification accuracy. The results are shown as mean ± standard deviation for  $k = 5$  fold. For this case, we made sure

that the train and test cohorts had vocalizations from all cows such that the proportion of the vocalizations of each individual cow was present both in the train and test cohorts at each fold. The models obtained around 70% accuracy, with a relatively low standard deviation. Like the previous experiment, the DL model outperformed the explainable model.

Figure 3 presents the features' importance of the explainable model for the cow identification task, calculated by reducing one feature from the input and calculating its influence on the model's performance. The sound duration played a critical role with 14.27% importance, indicating that different cows have a significant pattern in their vocal duration, or at least a non-linear connection between the vocal duration to other features that allows for capturing unique identification patterns. The Wiener Entropy mean is the second best, with 11.65% importance.

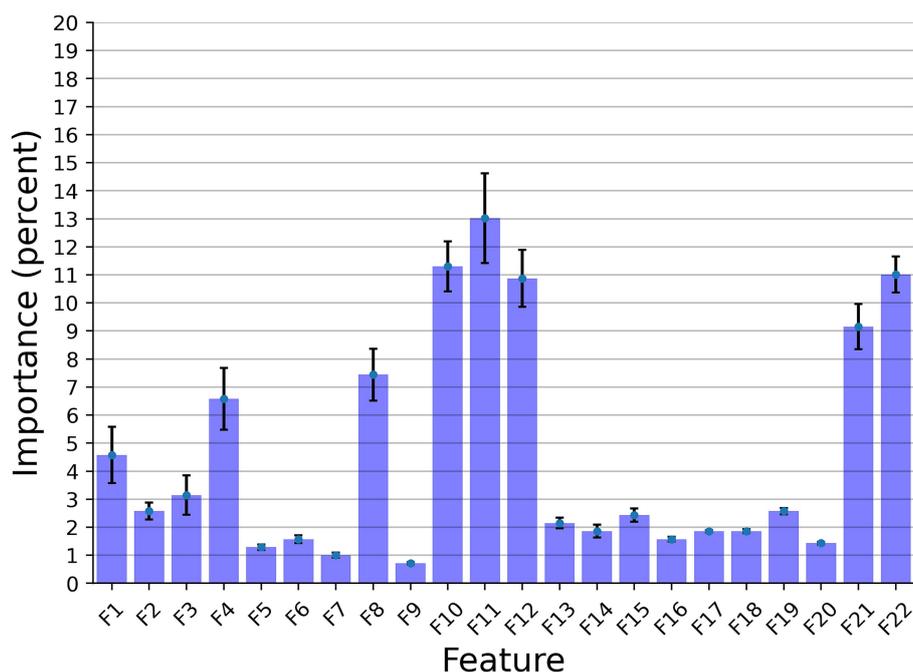


FIGURE 2

The distribution of the features' importance for the high and low frequency calls (LF and HF) explainable classifier model. The results are shown as the average of a  $k = 5$  fold cross-validation where the error bars indicate one standard deviation.

TABLE 3 The individual identification classifier models' performances.

Dataset	Model	Train set accuracy	Test set accuracy
Low Frequency + High Frequency	Explainable	73.0 ± 3.3%	68.9 ± 5.1%
	Deep learning	76.3 ± 4.2%	72.5 ± 4.7%
Low Frequency	Explainable	58.2 ± 1.3%	50.9 ± 2.8%
	Deep learning	65.5 ± 1.8%	46.8 ± 3.3%
High Frequency	Explainable	79.6 ± 2.6%	68.4 ± 3.2%
	Deep learning	74.9 ± 3.0%	70.8 ± 3.4%

The results are shown as mean standard deviation for  $k = 5$  fold.

## Discussion

In this study, we present a dataset of cattle vocal recordings during negative affective states caused by isolation of the cows, which is, to the best of our knowledge, the largest dataset collected to date. The data from  $n = 20$  cows has been manually cleaned from background noises and trimmed to contain only the low-frequency (LF) and high-frequency (HF) calls, to ensure as high quality of data as possible. The resulting dataset comprises 1,144 records in total. Based on this data, we conducted two sets of tasks. Firstly, we provided a classifier for separating between low and high-frequency calls. Secondly, we provided a classifier for identifying individual cows based on their high-, low-, or high + low-frequency vocalizations produced.

As shown in Table 2, both the explainable and DL models were able to accurately classify between the low- and high-frequency calls, with 87.2 and 89.4% accuracy, respectively. This outcome slightly outperforms (2%, 4.4%) the current state-of-the-art model (33), which used a smaller dataset of  $n = 10$  individuals. Notably,

the differences between the models' performances between the training and testing cohorts was around 2.5% toward the training cohort, compared to the state-of-the-art, which reports a 14.2% difference. As such, our model resulted in less over-fitting, if any at all, than the previous model. In addition, as the standard deviations of both models were 4.1 and 3.8%, this indicates that both models are robust.

For the individual cow identification task for both the LF and HF data, the explainable and DL models obtained 68.9 and 72.5% accuracy, respectively. When focusing only on the HF calls, the results were similar, with only 0.5 and 1.7% decrease in performance. On the other hand, when using only the LF samples, the accuracy sharply dropped to 50.9 and 46.8%, respectively, while also revealing overfit over the training dataset. This may be an indication that high-frequency calls contain more individuality information than low-frequency calls in cattle. These results are in accordance with previous findings across non-human mammals (29, 48), where an increase in the arousal states was shown to lead to higher frequency vocalizations for both F0 and formant-related

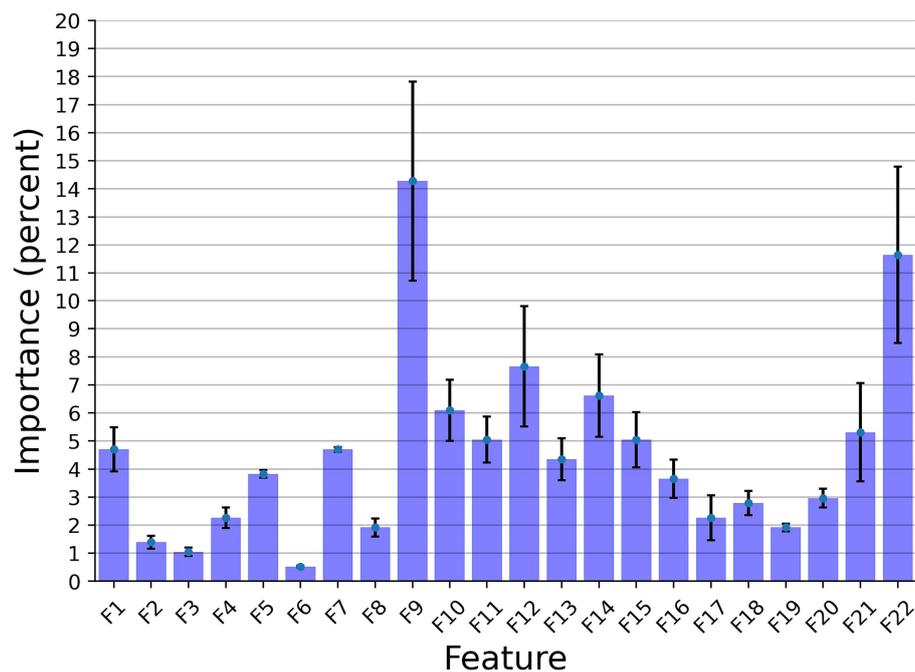


FIGURE 3

The distribution of the features' importance for identification explainable classifier model for the Low Frequency + High Frequency dataset. The results are shown as the average of a  $k = 5$  fold cross-validation, where the error bars indicate one standard deviation.

features, with vocalization parameters being more variable in negative-high arousal states. An alternative explanation for this might be attributed to the reduced amount of LF data, which contained 192 samples (i.e., 16.8% of the entire dataset). While the performance of the model of Shorten and Hunter (33) was better, this study worked with a reduced dataset for LF calls. In addition, their results may be indicative of overfitting, while explainable frameworks were not considered.

Considering homologies in the physiology of vocalization production and the commonalities found across species (48), the current findings could be extrapolated to other European cattle (*Bos taurus*), both dairy and beef specialized breeds. In a comparative study conducted on the two cattle sub-species (*B. taurus* and *B. indicus*), Moreira et al. (49) found *B. indicus* animals to be more reactive to both low and high frequencies sounds, which the authors attributed to the smaller auricle and greater interaural distance found in *B. taurus*, when compared to the indicine cattle. Although the hearing range might differ among closely related species, Maigrot et al. (14) found the functions of vocalizations to exceed intraspecies exchanges of information in domestic horses and Przewalski's horses, wild boars, and domestic pigs, these species being able to discriminate among positive and negative vocalizations produced by heterospecifics, including humans. Moreover, another potential contribution of the current research becomes apparent based on the experimental design and data collection. Whereas the studies conducted on cattle communication behavior to-date analyzed predominantly vocalizations emitted by cows either in an un-controlled setting (e.g., mob on pasture or inside the barn), or assessed and compared calls among a wider set of contexts (e.g., positive and/or negative, with different putative valences and arousal levels) (21, 22, 33, 35).

Conversely, our experimental setting was exclusively focused on a single negative context, while changes in affective states of the same animal being proven previously to result in modulations of the vocal parameters and behavior, e.g., during dam-calf separation and reunion in beef cattle (50). However, it is worth pointing out that our isolation challenge replicates an event that is occurring frequently under production conditions, with cows being individually isolated for health (e.g., sickness and veterinary visits) and reproduction (e.g., artificial insemination, fertility treatments, and pregnancy diagnosis) reasons.

Moreover, the use of vocalization scoring as an indicator of welfare during cattle handling at slaughter was shown to be a feasible approach, with vocalizations frequency and cortisol levels being influenced by the use of electric prods, deficient stunning and aversive handling (51, 52).

Our results are in alignment with previous research which showed that isolation from herd-mates induces a wide range of behavioral and physiological responses in cattle (22, 25), given the much higher incidence of HF calls observed during the isolation challenge, and taking into account the previous research results which suggested that the production and broadcasting of a repetitive single call type is indicative of persistent negative affective states (53), while reflecting a high urgency for the animal itself (54).

This research is, however, not without its limitations. Factors such as emotional contagion among herd-mates, and thus the potential biological role of the distress vocalizations emitted by cows during the isolation challenge were not studied in the current trial. For instance, Rhim (55) found that vocalization and behavior of Holstein cows and calves during partial separation (with vocal and olfactory communication) has led to significantly higher vocalization rates in

both cows and calves, when compared to complete separation. To address this, in our future research, we plan to include the use of additional sensors such as heart rate monitors, infrared thermography, and stress-related biomarkers, to have a more generalized approach when evaluating emotional responses to negative contexts. Moreover, considering the psychology and behavioral patterns of the species, mental processes such as learned helplessness could have contributed to the time-modulation of the vocal parameters following herd isolation, with animals abandoning their attempts to signal the negative event due to a perceived lack of control, which, however, does not mean that the negative event is being perceived as neutral by the animals. Additionally, the study herd consisted of multiparous adult cows, with various degrees of existing habituation to social isolation being presumed.

To summarize, cattle vocalizations can be seen as commentaries emitted by an individual on their own internal affective state, with the challenges lying in understanding and deciphering those commentaries. Looking forward, significantly more work needs to be done, taking into account a wider range of contexts and potential influencing factors on the vocal cues, in order to be able to draw strong conclusions regarding arousal or valence in cattle bioacoustics. Our study highlights the promising applications of machine learning approaches in cattle vocalization behavior. In order for such approaches to be validated for commercial use and adopted at farm-level, further important developments need to occur, such as designing of hardware and software capable to filter and limit external farm noises, and to process vocalizations automatically. Furthermore, in order for vocalizations to be reliable welfare and health predictors, the training and testing vocalization sets need to include a significant wider variety of negative affective states (e.g., sickness, cow-calf separation, weaning, estrous, and pain inducing vocalizations), all while keeping a special focus on changes in individual animal vocalizations.

## Conclusion

In this study, we compiled a data set of dairy cattle vocal recordings during the negative affective state of isolation, which is one of the largest and cleanest datasets of its kind. Through our experiments using explainable and DL models, we have demonstrated the effectiveness of these models in classifying high- and low-frequency calls, as well as for identifying individual cows based on their vocalization productions. These results highlight the future potential of vocalization analysis as a valuable tool for assessing the emotional valence of cows and for providing new insights into promoting precision livestock farming practices. By monitoring cattle vocalizations, animal scientists could gain crucial insights into the emotional states of the animals, empowering them to make informed decisions to improve the overall farm animal welfare. Future work endeavors can take these results a step forward, gathering cattle vocalizations at different critical affective states to identify possible health risks or early real-time disease diagnostics.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession

number(s) can be found here: <https://gitlab.com/is-annazam/bovinetalk>.

## Ethics statement

The experimental procedures and protocols were reviewed and approved by the Ethical Committee from the Research and Development Institute for Bovine, Balotesti, Romania (approval no. 0027, issued on July 11, 2022). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent was obtained from the owners for the participation of their animals in this study.

## Author contributions

DG: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. MM: Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. TL: Data curation, Formal Analysis, Investigation, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. AO: Data curation, Formal Analysis, Investigation, Software, Validation, Writing – original draft, Writing – review & editing. IN: Formal Analysis, Investigation, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. AZ: Data curation, Formal Analysis, Investigation, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS—UEFISCDI, project number PN-III-P1-1.1-TE-2021-0027, within PNCDI III.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Webster J. Animal welfare: freedoms, dominions and “a life worth living”. *Animals*. (2016) 6:35. doi: 10.3390/ani6060035
- Webb LE, Veenhoven R, Harfeld JL, Jensen MB. What is animal happiness? *Ann N Y Acad Sci*. (2019) 1438:62–76. doi: 10.1111/nyas.13983
- Kremer L, Holkenborg SK, Reimert I, Bolhuis J, Webb L. The nuts and bolts of animal emotion. *Neurosci Biobehav Rev*. (2020) 113:273–86. doi: 10.1016/j.neubiorev.2020.01.028
- Laurijs KA, Briefer EF, Reimert I, Webb LE. Vocalisations in farm animals: a step towards positive welfare assessment. *Appl Anim Behav Sci*. (2021) 236:105264. doi: 10.1016/j.applanim.2021.105264
- Ben Sassi N, Averós X, Estevez I. Technology and poultry welfare. *Animals*. (2016) 6:62. doi: 10.3390/ani6100062
- Matthews SG, Miller AL, Clapp J, Plötz T, Kyriazakis I. Early detection of health and welfare compromises through automated detection of behavioural changes in pigs. *Vet J*. (2016) 217:43–51. doi: 10.1016/j.tvjl.2016.09.005
- Silva M, Ferraris S, Costa A, Aerts JM, Guarino M, Berckmans D. Cough localization for the detection of respiratory diseases in pig houses. *Comput Electron Agric*. (2008) 64:286–92. doi: 10.1016/j.compag.2008.05.024
- Tallet C, Linhart P, Policht R, Hammerschmidt K, Šimeček P, Kratinova P, et al. Encoding of situations in the vocal repertoire of piglets (*sus scrofa*): a comparison of discrete and graded classifications. *PLoS One*. (2013) 8:e71841. doi: 10.1371/journal.pone.0071841
- Whitaker B. M., Carroll B. T., Daley W., Anderson D. V. (2014). “Sparse decomposition of audio spectrograms for automated disease detection in chickens” in IEEE Global Conference on Signal and Information Processing (GlobalSIP), IEEE, 2014. 1122–1126.
- Briefer EF, Tettamanti F, McElligott AG. Emotions in goats: Mapping physiological, behavioural and vocal profiles. *Anim Behav*. (2015) 99:131–43. doi: 10.1016/j.anbehav.2014.11.002
- Linhart P, Ratcliffe VF, Reby D, Špinková M. Expression of emotional arousal in two different piglet call types. *PLoS One*. (2015) 10:e0135414. doi: 10.1371/journal.pone.0135414
- Briefer EF, Mandel R, Maigrot AL, Briefer Freymond S, Bachmann I, Hillmann E. Perception of emotional valence in horse whinnies. *Front Zool*. (2017) 14:1–12. doi: 10.1186/s12983-017-0193-1
- McGrath N, Dunlop R, Dwyer C, Burman O, Phillips CJ. Hens vary their vocal repertoire and structure when anticipating different types of reward. *Anim Behav*. (2017) 130:79–96. doi: 10.1016/j.anbehav.2017.05.025
- Maigrot A-L, Hillmann E, Briefer EF. Encoding of emotional valence in wild boar (*sus scrofa*) calls. *Animals*. (2018) 8:85. doi: 10.3390/ani8060085
- Briefer EF, Sypherd CCR, Linhart P, Leliveld LMC, Padilla de la Torre M, Read ER, et al. Classification of pig calls produced from birth to slaughter according to their emotional valence and context of production. *Sci Rep*. (2022) 12:3409. doi: 10.1038/s41598-022-07174-8
- Ede T, Lecorps B, von Keyserlingk MA, Weary DM. Symposium review: scientific assessment of affective states in dairy cattle. *J Dairy Sci*. (2019) 102:10677–94. doi: 10.3168/jds.2019-16325
- Green AC, Lidfors LM, Lomax S, Favaro L, Clark CE. Vocal production in postpartum dairy cows: temporal organization and association with maternal and stress behaviors. *J Dairy Sci*. (2021) 104:826–38. doi: 10.3168/jds.2020-18891
- McLoughlin MP, Stewart R, McElligott AG. Automated bioacoustics: methods in ecology and conservation and their potential for animal welfare monitoring. *J Royal Soc Interface*. (2019) 16:20190225. doi: 10.1098/rsif.2019.0225
- Watts JM, Stookey JM. The propensity of cattle to vocalise during handling and isolation is affected by phenotype. *Appl Anim Behav Sci*. (2001) 74:81–95. doi: 10.1016/S0168-1591(01)00163-0
- Yajuvendra S, Lathwal SS, Rajput N, Raja TV, Gupta AK, Mohanty TK, et al. Effective and accurate discrimination of individual dairy cattle through acoustic sensing. *Appl Anim Behav Sci*. (2013) 146:11–8. doi: 10.1016/j.applanim.2013.03.008
- de la Torre MP, Briefer EF, Reader T, McElligott AG. Acoustic analysis of cattle (*Bos taurus*) mother–offspring contact calls from a source–filter theory perspective. *Appl Anim Behav Sci*. (2015) 163:58–68. doi: 10.1016/j.applanim.2014.11.017
- Green A, Clark C, Favaro L, Lomax S, Reby D. Vocal individuality of Holstein-Friesian cattle is maintained across putatively positive and negative farming contexts. *Sci Rep*. (2019) 9:18468. doi: 10.1038/s41598-019-54968-4
- Boissy A, Le Neindre P. Behavioral, cardiac and cortisol responses to brief peer separation and Reunion in cattle. *Physiol Behav*. (1997) 61:693–9. doi: 10.1016/S0031-9384(96)00521-5
- Holm L, Jensen MB, Jeppesen LL. Calves’ motivation for access to two different types of social contact measured by operant conditioning. *Appl Anim Behav Sci*. (2002) 79:175–94. doi: 10.1016/S0168-1591(02)00137-5
- Müller R, Schrader L. Behavioural consistency during social separation and personality in dairy cows. *Behaviour*. (2005) 142:1289–306. doi: 10.1163/156853905774539346
- van Reenen CG, O’Connell NE, van der Werf JTN, Korte SM, Hopster H, Jones RB, et al. Responses of calves to acute stress: individual consistency and relations between behavioral and physiological measures. *Physiol Behav*. (2005) 85:557–70. doi: 10.1016/j.physbeh.2005.06.015
- Lecorps B, Kappel S, Weary DM, von Keyserlingk MA. Dairy calves’ personality traits predict social proximity and response to an emotional challenge. *Sci Rep*. (2018) 8:16350. doi: 10.1038/s41598-018-34281-2
- Nogues E, Lecorps B, Weary DM, von Keyserlingk MA. Individual variability in response to social stress in dairy heifers. *Animals*. (2020) 10:1440. doi: 10.3390/ani10081440
- Briefer EF. Vocal expression of emotions in mammals: mechanisms of production and evidence. *J Zool*. (2012) 288:1–20. doi: 10.1111/j.1469-7998.2012.00920.x
- Volodin IA, Lapshina EN, Volodina EV, Frey R, Soldatova NV. Nasal and oral calls in juvenile goitred gazelles (*Gazella subgutturosa*) and their potential to encode sex and identity. *Ethology*. (2011) 117:294–308. doi: 10.1111/j.1439-0310.2011.01874.x
- Sèbe F, Poindron P, Ligout S, Sèbe O, Aubin T. Amplitude modulation is a major marker of individual signature in lamb bleats. *Bioacoustics*. (2018) 27:359–75. doi: 10.1080/09524622.2017.1357146
- Taylor AM, Charlton BD, Reby D. Vocal production by terrestrial mammals: source, filter, and function. *Vertebrate Sound Product Acoustic Commun*. (2016) 53:229–59. doi: 10.1007/978-3-319-27721-9\_8
- Shorten P, Hunter L. Acoustic sensors for automated detection of cow vocalization duration and type. *Comput Electron Agric*. (2023) 208:107760. doi: 10.1016/j.compag.2023.107760
- Jung DH, Kim NY, Moon SH, Jhin C, Kim HJ, Yang JS, et al. Deep learning-based cattle vocal classification model and real-time livestock monitoring system with noise filtering. *Animals*. (2021) 11. doi: 10.3390/ani11020357
- Li G, Xiong Y, Du Q, Shi Z, Gates RS. Classifying ingestive behavior of dairy cows via automatic sound recognition. *Sensors*. (2021) 21:5231. doi: 10.3390/s21155231
- Ilie DE, Gao Y, Nicolae I, Sun D, Li J, Han B, et al. Evaluation of single nucleotide polymorphisms identified through the use of SNP assay in Romanian and Chinese Holstein and Simmental cattle breeds. *Acta Biochim Pol*. (2020) 67:341–6. doi: 10.18388/abp.2020\_5080
- Mincu M, Gavojdian D, Nicolae I, Olteanu AC, Vlagioiu C. Effects of milking temperament of dairy cows on production and reproduction efficiency under tied stall housing. *J Vet Behav*. (2021) 44:12–7. doi: 10.1016/j.jvbe.2021.05.010
- Boersma P., Praat D.W. (2022). Doing phonetics by computer [computer program]. Available at: <http://www.praat.org/>
- Reby D, McComb K. Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Anim Behav*. (2003) 65:519–30. doi: 10.1006/anbe.2003.2078
- Gabriël J.L.B. (2004). Wiener entropy, script developed in praat v. 4.2.06. Available at: <https://gbeckers.nl/pages/phonetics.html>
- Briefer EF, Vizier E, Gyga L, Hillmann E. Expression of emotional valence in pig closed-mouth grunts: involvement of both source- and filter-related parameters. *J Acoust Soc Am*. (2019) 145:2895–908. doi: 10.1121/1.5100612
- Le TT, Fu W, Moore JH. Scaling tree-based automated machine learning to biomedical big data with a feature set selector. *Bioinformatics*. (2020) 36:250–6. doi: 10.1093/bioinformatics/btz470
- Feurer M, Eggenberger K, Falkner S, Lindauer M, Hutter F. Auto-sklearn 2.0: hands-free automl via meta-learning. *J Mach Learn Res*. (2022) 23:11936–96. doi: 10.48550/arXiv.2007.04074
- Daugela K., Vaiciukynas E. (2022). “Real-time anomaly detection for distributed systems logs using apache kafka and h2o AI” in *International Conference on Information and Software Technologies*. Springer, 33–42.
- Fushiki T. Estimation of prediction error by using k-fold cross-validation. *Stat Comput*. (2011) 21:137–46. doi: 10.1007/s11222-009-9153-8
- Chen T., Guestrin C. (2016). “Xgboost: a scalable tree boosting system” in *Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining*, 785–794.
- Ye F, Yang J. A deep neural network model for speaker identification. *Appl Sci*. (2021) 11:3603. doi: 10.3390/app11083603
- Briefer EF. Coding for ‘dynamic’ information: vocal expression of emotional arousal and valence in non-human animals In: T Aubin and N Mathevon, editors. *Animal signals and communication: coding strategies in vertebrate acoustic communication*. Springer, Cham. (2020) 7:137–62. doi: 10.1007/978-3-030-39200-0\_6

49. Moreira SM, Barbosa Silveira ID, da Cruz LAX, Minello LF, Pinheiro CL, Schwengber EB, et al. Auditory sensitivity in beef cattle of different genetic origins. *J Vet Behav.* (2023) 59:67–72. doi: 10.1016/j.jveb.2022.10.004
50. Schnaider MA, Heidemann MS, Silva AHP, Taconeli CA, Molento CFM. Vocalization and other behaviors as indicators of emotional valence: the case of cow-calf separation and Reunion in beef cattle. *J Vet Behav.* (2022) 49:28–35. doi: 10.1016/j.jveb.2021.11.011
51. Grandin T. The feasibility of using vocalization scoring as an indicator of poor welfare during cattle slaughter. *Appl Anim Behav Sci.* (1998) 56:121–8. doi: 10.1016/S0168-1591(97)00102-0
52. Hemsworth PH, Rice M, Karlen MG, Calleja L, Barnett JL, Nash J, et al. Human–animal interactions at abattoirs: relationships between handling and animal stress in sheep and cattle. *Appl Anim Behav Sci.* (2011) 135:24–33. doi: 10.1016/j.applanim.2011.09.007
53. Collier K, Townsend SW, Manser MB. Call concatenation in wild meerkats. *Anim Behav.* (2017) 134:257–69. doi: 10.1016/j.anbehav.2016.12.014
54. Kershenbaum A, Blumstein DT, Roch MA, Akçay Ç, Backus G, Bee MA, et al. Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol Rev.* (2016) 91:13–52. doi: 10.1111/brv.12160
55. Rhim SJ. Vocalization and behavior of Holstein cows and calves after partial and complete separation. *Rev Colomb Cienc Pecuarías.* (2013) 26:24–9.