Open Dataset of Acoustic Recordings of Foraging Behavior in Dairy Cows

Luciano S. Martinez-Rau[1,2,*], José O. Chelotti[1,3], Mariano Ferrero[1], Santiago A. Utsumi[4,5], Alejandra M. Planisich[6], Leandro D. Vignolo[1], Leonardo L. Giovanini[1], H. Leonardo Rufiner[1,7], and Julio R. Galli[6,8]

[1]Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i), FICH-UNL/CONICET, Argentina
[2]Department of Electronics Design, Mid Sweden University, 85170 Sundsvall, Sweden
[3]TERRA Teaching and Research Center, University of Liège, Gembloux Agro-Bio Tech (ULiège-GxABT), 5030 Gembloux, Belgium
[4]W.K. Kellogg Biological Station and Department of Animal Science, Michigan State University, United States
[5]Department of Animal and Range Science, New Mexico State University, United States
[6]Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Argentina
[7]Laboratorio de Cibernética, Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Argentina
[8]Instituto de Investigaciones en Ciencias Agropecuarias de Rosario, IICAR, UNR-CONICET, Argentina
[*]corresponding author(s): Luciano S. Martinez-Rau (luciano.martinezrau@miun.se)

ABSTRACT

Monitoring of livestock feeding behavior is essential to assess animal wellness, and nutritional status and optimize pasture management. Nevertheless, manual monitoring can be challenging due to the need for continuous monitoring over long periods of time. To overcome this challenge, the use of automatic techniques based on the acquisition and analysis of data from sensors is adopted. This work describes an extensive open dataset of acoustic recordings of the foraging behavior of dairy cows. The dataset includes 662 hours of daily records obtained using unobtrusive and non-invasive instrumentation mounted on five lactating multiparous Holstein cows continuously monitored for six non-consecutive days in pasture and barn. Labeled recordings precisely delimiting grazing and rumination bouts are provided for a total of 400 hours and for over 6,200 masticatory jaw movements. Companion information on the audio recording quality and expert-generated labels is also provided to facilitate data interpretation and analysis. This comprehensive dataset is a useful resource for studies that are aimed to explore new tools and solutions for precision livestock farming.

Background & Summary

Advances in information and communication technologies is allowing the implementation of precision livestock farming (PLF) systems and solutions with promising application to enhance farm operational efficiencies and animal welfare[1,2]. Over the last three decades, PLF has grown substantially, attracting farmers, operators and industries around the globe[3,4]. New PLF developments include methodologies to enable the individual monitoring of livestock feeding behavior, which in most production systems could be used to trace changes of animal welfare with direct insights into animal nutrition, health or performance[5–7]. For example, minutes to hourly changes in the pattern, duration and periodicity of meals, could be useful to decide pasture allowance management[8] and diets[9,10], indicate a state of stress[11,12] or anxiety[13], serve as an early indicator of diseases[14–16], rumen health[17,18], initiation of the birthing process[19–21] or heat detection aid[22–24].

Wearable sensors are the most common data acquisition method to monitor feeding behavior[25,26]. Accelerometers and inertial measurement units determine the head and neck movements and have been used mainly in confined environments[27,28]. Acoustic sensors are typically preferred over motion sensors in free-ranging conditions[29] to recognize animal jaw movements (JMs)[30–33] and inform changes of foraging behavior[34]. Furthermore, distinguishing differences between different types of masticatory JMs is useful to estimate changes in grazing and rumination bouts[35], estimate differences in dry matter intake, or discriminate different feedstuffs and plants[36,37].

The acoustic monitoring of foraging behavior is a complex engineering task that requires confident solutions tolerable to noise, interference and disturbance [29]. The opportunities of using acoustic methods for practical farm-level management and animal research are ample[38], but the limited availability of public/open acoustic datasets usually hinders new and relevant research[39]. In one case, an open dataset of 52 audio recordings of JMs of dairy cows grazing on two contrasting forage species at two sward heights was published in Vanrell et al.[40]. In other related cases, few samples of cattle vocalizations are provided but not the entire dataset[41,42].

This work presents a full dataset of audio recordings of masticatory sounds of dairy cows along with their corresponding event identification labels. The dataset is organized in three subsets. The first subset includes 24 h audio recordings of foraging sounds of dairy cows continuously monitored while grazing at pasture or visiting the dairy milking barn. A total of 662.5 h of sounds were recorded, from which 400.4 h correspond to foraging sounds registered in a free-range pasture environment. Annotations of grazing and rumination bouts for each of the cows are provided. Periods during which dairy cows remained indoors inside the dairy barn are also indicated. The second subset contains two audio files of 54.6 min of grazing and 30.0 min of rumination, with corresponding labels for masticatory JMs. Experts identified and labeled 4,221 and 2,006 masticatory JM produced during grazing and rumination, respectively. The third subset provides a comprehensive description of the different types of masticatory JMs and recorded animal behaviors, and specific information of the audio recordings. The dataset presented here could be implemented to further test improved PLF applications and algorithms for automatic detection and quantification of JMs and foraging behavior of cattle [40,43–46].

Methods

The field study took place from July 31 to August 19, 2014, and was conducted at the W.K. Kellogg Biological Station's Pasture Dairy Research Center of Michigan State University, located in Hickory Corners, Michigan, US (GPS coordinate 42° 24′ 21.8″ N 85° 24′ 08.4″ W). The procedures for animal handling, care and use in this experiment were revised, approved and conducted by the Institutional Animal Care and Use Committee of Michigan State University (#02/17 − 020 − 00). Animals were managed on a pasture-based robotic milking system with free cow traffic in Watt et al.[47]. Voluntary milking (3.0 ± 1.0 daily milkings) was conducted using two Lely A3-Robotic milking units (Lely Industries NV, Maassluis, The Netherlands). Permissions for milking were set by a minimum expected milk yield of 9.1 kg or a 6 h milking interval. Thus, milking frequency varied across cows according to milk yield. Dairy cows were fed a grain-based concentrate at 1 to 6 kg per kg of extracted milk (daily maximum 12 kg/cow) during milking and through automatic feeders located inside the dairy milking barn. The neutral detergent fiber (NDF), net energy for lactation (NEL), and average crude protein (CP) of the grain-based concentrate pellet supplied (Cargill Inc, Big Lake, MN) were 2.05 Mcal/kg dry matter (DM), 99.4 g/kg DM, and, 193.0 g/kg DM respectively. Cows were allowed 24 h access to grazing paddocks with predominance of orchardgrass (Dactylis glomerata), tall fescue (Lolium arundinacea) and white clover (Trifolium repens), or perennial ryegrass (Lolium perenne) and white clover. Two allocations of ~ 15 kg/cow of fresh pasture were offered from 10:00 to 22:00 and from 22:00 to 10:00 (GMT-5), each day. Allocations of fresh ungrazed pasture were made available at opposite sides of the farm (south and north) to entice cow traffic through the milking shed. An average of ~ 30 kg of DM/cow was offered daily. Thirty readings of sward height (SH,±x) along each paddock were conducted by a plate meter to estimate pre-grazing and post-grazing herbage biomass to a ground level ($Y, Y = 125x; r^2 = 0.96$). Across the 16 paddocks used in this study, the average pre-grazing herbage biomass was 2387 ± 302 kg DM/ha (19.2 ± 2.5 cm SH) and the average post-grazing herbage biomass was 1396 ± 281 kg DM/ha (11.2 ± 2.2 cm SH). Composite hand-plucked samples from the 16 paddocks were used to determine the 48 h in vitro digestibility of DM (IVDMD) (Daisy II, Ankom Technology Corp.), the acid (ADF) and neutral detergent fiber (NDF) (Fiber Analyzer, Ankom Technology Corp., Fairport, NY), the crude protein (CP) (4010 CN Combustion, Costech Analytical Technologies Inc., Valencia, CA), and the acid detergent lignin (ADL) content of consumed forages. The values of DM expressed in terms of g/kg for IVDMD, CP, NDF, ADF and ADL were 781 ± 30, 257 ± 20, 493 ± 45, 187 ± 25, 33 ± 8, respectively.

For this study, 5 lactating high-producing multiparous Holstein cows were selected from a herd of 146 Holstein cows and used to collect acoustic signals and to continuously monitor their foraging behavior using non-invasive techniques. Specific characteristics of individual cows are provided in Table 1. Individualized 24 h audio recordings were conducted on July 31, and August 4, 6, 11, 13 and 18, 2014, respectively. Records were completed following a 5 x 5 Latin-square design (Table 2) using 5 independent monitoring systems (halters, microphones and recorders) that were rotated daily across the 5 cows and throughout 6 non-consecutive recording days. This design was decided to control differences of sound data associated with a particular cow, recording systems or experiment day. On the first day, each recording system was randomly assigned to each cow. On the sixth day, the recording systems were reassigned to cows using the same order that was used on the first day. No training to the use of the recording systems was deemed necessary before study onset. Recording problems were registered with the recording system number 2. On the first day, the recording trial had to be stopped a few hours before completion because the recording system was unfastened from the cow. On the sixth day, the recording system failed to register any sound because the microphone connector was disconnected from the recorder. This trial was repeated on the next day (August 19) to complete the Latin-square design. Changes in the order and completion of recording trials should be considered in case of considering trial days as a random variable in the experimental design. The weather conditions during the study were registered by the National Weather Service Station located at the Kellogg Biological Station (Table 3).

Each recording system consists of two directional electret microphones connected to the stereo-input channels of a digital recorder (Sony Digital ICD-PX312, Sony, San Diego, CA, USA). This instrumentation was enclosed in a weather proof protective case (1015 Micron Case Series, Pelican Products, Torrance, CA, USA) mounted to the top side of a halter neck strap (Fig. 1). One microphone was positioned facing inwards in a non-invasive way and pressed against the forehead of the animal. The other microphone was placed facing outwards to capture the bone-transmitted vibrations and sounds produced by the animal, respectively. To achieve better microphone contact, hair of the central forehead area was removed using a sharp clipper. The microphones were held in the desired position by using a rubber foam and elastic headband attached to the halter. This design prevented microphone movements and allowed the insulation of microphones from environmental noises caused by wind, friction and scratches[48,49].

The dataset includes audio recordings registered in two settings: indoors while cows visited the dairy milking barn and outdoors while cows had free access to grazing pasture. After the first milking in the morning, cows were automatically separated to a holding pen and restrained by using head lockers for installation of the recording systems. The date and relevant information of recording systems and cows was kept in a logbook. A similar process was repeated daily and until completion of trials according to the Latin-square design. In each recorder system, the two microphones were connected randomly to the stereo-input channels of the recorder at the beginning of trials. This information was not logged. Experienced animal handlers, who had extensive experience in animal behaviors, data collection and analysis, directly observed the focal animals for blocks of ~ 5 minutes each hour. Observation of foraging behavior and other relevant parameters were documented and registered in the logbook. Handlers also checked the correct placement and location of recording systems on the cows. Observations were conducted at a distance from the animals to minimize disruptions of behavior.

Data Records

The audio recordings were saved in MPEG-1 Audio Layer III (MP3) format[50] with a sampling rate of 44.1 kHz, providing a nominal recording bandwidth of 22 kHz and a dynamic range of 96 dB. The recordings were made in stereo, using one microphone per channel with a resolution of 16 bits at 192 kbps. The digital recorder automatically crops and generates a new MP3 file if the current audio recording is longer than 6 h. Thus, 24 h audio recordings are partitioned into 4 parts of approximately 6 h each. The data included in the dataset are accessible through repositories available in Figshare (*data upload with the manuscript and cited after acceptance*). The dataset are organized in three subsets (Fig. 2) as follows:

1. Subset of continuous monitoring: contains 30 ZIP files that correspond to the different recording trials of this study. Each ZIP file comprises ~ 24 h of audio recordings and the corresponding label files generated by two experts (Fig. 2). A total of 662.5 h of audio recordings are included in this subset, consisting of 262.1 h registered indoors while cows visited the dairy milking barn, and 400.4 h registered outdoors while cows remained at pasture. The label files are a list of timestamps indicating the start and end of identified animal behaviors and other annotation remarks. Animal behavior categories labels include the following activities: grazing, rumination, walking, and idle or resting. Other annotation labels include the animal location into the barn and the time for the installation and retrieval of the recording systems.

2. Subset of masticatory JMs: consist of a ZIP file containing 2 audio files and 2 corresponding label files of masticatory JMs. Audio files correspond to a grazing and rumination bout respectively extracted from the channel 1 of the 'D3RS4ID2909P3.mp3' file. The label files indicate the timestamps (start and end) of distinct masticatory JMs associated with grazing and rumination. The labels were generated by inspection of audio files by two experts in sound signal processing. In addition to these files, 2 more files of masticatory JM labels were generated using a Python script that automatically adjusted the timestamps provided by the experts. This subset also includes a file of the masticatory JM labels ('D3RS4ID2909P3_JM.txt/csv'), with their timestamps aligned for use with the 'D3RS4ID2909P3.mp3' file.

3. Subset of additional information:

    - The 'BehaviorLabelsDescription.pdf' file provides a comprehensive description of animal behavior categories, including the registered annotations and the criteria used to determine the start and end of each behavior.

    - The 'MasticatoryJMDescription.pdf' file provides details on the different types of masticatory JMs produced during grazing and rumination activities.

    - The 'AudioInformation.xlsx' file provides detailed information of the audio recordings. Information consists of the corresponding trials of the Latin-square design (day, cow and recording system), audio duration, sound quality, registered animal behaviors, audio channels, and companion comments.

Technical Validation

The subset of continuous monitoring of animal behaviors and corresponding annotations comprises 133 audio recordings and 133 label files, respectively. The label files were generated by two experts with extensive experience in animal behaviors scouting and digital analysis of audio signals[40,46,48,49,51,52]. The label-making process was overseen by an expert and the results were double-inspected and checked by another expert. The experts were guided by the logbook and used Audacity software for observing the sound waveforms and for listening to sounds to identify, classify, and label data into animal behavior categories and annotations of interest. The criteria used by the experts to delimit the start and end of each label are specified in the 'BehaviorLabelsDescription.pdf' file. Although the experts matched all label assignments, there were some small differences in the start and/or end times (timestamp) of some labels. In those cases, both experts revised the labels together until they reached a mutual agreement. Additionally, as previously mentioned, the two microphones of each recording system were randomly connected to the stereo-input channels of the recorder throughout the trials. As a consequence, the stereo-input channels are swapped across the audio recordings. To address this, the experts marked the one-to-one correspondence between the stereo-input channels and the two microphones (facing inwards and outwards of the forehead of the animal) for all audio recordings. The experts made their decision based on the individualized observation and listening to audio recordings, and on a final mutual agreement.

During the continuous monitoring of cows, rumination and feeding activities inside the milking barn were annotated in the logbook by the animal handler and recorded by the recording systems. However, the experts did not label these activities because the presence of acoustic noise in the audio recordings made it difficult to ensure their proper delimitation. The main focus of the experiment was to collect acoustic signals of foraging behavior while cows grazed in free-range condition. The foraging behavior can be identified by masticatory JMs on a short timescale and defined grazing and rumination activities on a long timescale.

Cattle engage in two primary foraging activities, grazing and rumination, that can last from several minutes to hours[53,54]. The duration of grazing and rumination bouts collected in this study are shown in Fig. 3. Grazing entails the search, apprehension, chew, and swallow of herbage. A grazing bout consists of a non-predefined sequence of masticatory JMs performed rhythmically every ~ 1 s[29], with occasional interruptions produced during the search and displacement of the animal to a new feeding station. In this dataset, interruptions of the regular JMs greater than 90 s were considered to delimit a grazing bout. These interruptions could be associated with an animal distraction or animal displacement to a distant feeding patch. The great sensitivity to interruptions of regular JMs generates multiple short grazing bouts that can be aggregated into longer grazing meals and making it useful to estimate minute to hourly grazing time budgets. Thus, about 40% of the grazing bouts last less than 25 min (see Fig. 3), while a typical grazing meal lasts more than 1 hour[29]. The waveform and spectrogram of audio signals during grazing are shown in Fig. 4a. Rumination is characterized by repetitive cycles of 40-60 s of chew performed rhythmically every ~ 1 s that are followed by a 3-7 s pause required for swallowing and regurgitating the feed cud (bolus)[6,52]. The waveform and spectrogram of audio signals during rumination are shown in Fig. 4b. The bottom panel of Fig. 4b shows a zoom-in of the waveform region produced during the pause between two consecutive chewing periods. About 85% of the rumination bouts lasts less than 75 min (Fig. 3). A more detailed explanation of grazing and rumination activities is provided in the 'BehaviorLabelsDescription.pdf' file.

The subset of masticatory JMs produced in a grazing and rumination bout comprises 2 audio files, along with the corresponding label files. The masticatory JM labels were generated by the same experts, following the same approach and label criteria used for the subset of continuous monitoring of animal behaviors. The 'MasticatoryJMDescription.pdf' file explains the marks and characteristics used to distinguish the different masticatory JMs. A total of 6,227 masticatory JMs were individualized, delimited and classified. This is a complex task that requires significant processing time and expertise in audio signal processing and inspection. Therefore, in the label files generated by experts ('JM_grazing.txt/csv' and 'JM_rumination.txt/csv'), the start and end (timestamp) of the masticatory JM labels could be subjective and may vary from the true bounds of the JMs in the audio files. To address this potential bias, an algorithm was used to automatically adjust the timestamps of the JM labels generated by experts.

Masticatory JMs have a duration of approximately 0.3-0.6 s and can be classified into three types: bite, chew, and chew-bite. A bite occurs when herbage is apprehended and severed, while a chew comminutes the herbage. A chew-bite is a combination of chewing and biting in a single JM[34,37,55]. The three types of JMs are present during grazing, whereas chews are present exclusively during rumination to comminute the feeding bolus[6,56] (bottom panels in Fig. 4). Waveforms and spectral characteristics of the JMs are shown in Fig. 5. Variation in chews produced during grazing and rumination occurs due to differences in moisture content of the ingested matter in each case[45,52]. In the subset, the 6,227 masticatory JMs correspond to 578 bites (9.3%), 1,136 chews (18.2%), and 2,507 chew-bites (40.3%) during grazing, as well as 2,006 chews (32.2%) during rumination. Only three possible non-labeled jaw movements (<0.1%) were observed. A more detailed explanation of the three types of JMs is provided in the 'MasticatoryJMDescription.pdf'.

To evaluate the sound quality of the audio recordings obtained from the continuous monitoring of dairy cows, only the relevant sections related to the foraging activities of interest were examined. Initially, the experts conducted a subjective analysis by listening to random segments of each of these sections and confirmed that all foraging activities were aurally discriminated from the background noise. This statement was further confirmed through a quantitative analysis. Quality assessment of these sections involved examining the masticatory JMs executed during grazing and rumination. For each audio recording, two quality indicators of masticatory JMs were computed individually for grazing and for rumination using previously established parameters[45].

The first parameter, the JM modulation index (MI) is useful to locate the masticatory JMs. The MI is a measure based on the difference of the audio signal intensity produced during masticatory JMs and the background noise. Give that the masticatory JMs are performed rhythmically every ~ 1 s during grazing and rumination, the MI was computed as:

$$MI_{JM} = \left(\overline{JM_{intra}} - \overline{JM_{inter}}\right) / \left(\overline{JM_{intra}} + \overline{JM_{inter}}\right) \in [0;1]$$

where $\overline{JM_{intra}}$ and $\overline{JM_{inter}}$ are the mean audio signal intensity produced during masticatory JMs and mean audio signal intensity produced in the short-pauses between consecutive masticatory JMs respectively, and defined as:

$$\overline{JM_{intra}} = \frac{1}{l_{intra}} \sum_{k=1}^{l} x^2[k]w[k]$$

$$\overline{JM_{inter}} = \frac{1}{l_{inter}} \sum_{k=1}^{l} x^2[k](1-w[k])$$

where $x[k]$ is the audio signal, l is the length in samples of the audio signal, l_{intra} and l_{inter} are the total number of samples with and without masticatory JMs, respectively, and $w[k]$ is a logical function indicating the presence of an masticatory JM in the k-th sample.

The second parameter is the signal-to-noise ratio (SNR). This parameter indicates the extent to which the background noise affects the sound produced during masticatory JMs, thus helping to differentiate between masticatory JMs associated with chews, bites and chew-bites. To compute the SNR, the sound produced during masticatory JMs must be isolated from the background noise. A multiband spectral subtraction algorithm assuming uncorrelated additive noise in the audio recordings was used to estimate a noise-free signal $\hat{s}[k]$ and a noisy signal $\hat{n}[k]$[57]. The SNR is computed as follow:

$$SNR(dB) = 10log\left(\sum_{k=1}^{l} \hat{s}^2[k]\right) - 10log\left(\sum_{k=1}^{l} \hat{n}^2[k]\right) \in \mathbb{R}$$

Examples of audio recordings with high- and low-quality sound are available in an open repository (Data Citation 2[58]). Their waveforms are presented in Fig. 6a and 6b. The higher the MI and SNR values, the better the audio recording quality. The frequency distribution of the estimated values of MI and SNR for both rumination and grazing computed over the audio recordings of the subset of continuous monitoring are shown in Fig. 7. Fig 7a. shows a considerable variation in the MI values of rumination and grazing. The MI values of rumination tend to be smaller than the MI values of grazing. This indicates that the masticatory JMs produced in rumination (exclusively chews) are more difficult to distinguish from the background noise. This is partly due to the lower intensity of the masticatory JMs produced during rumination compared to grazing, as shown in Fig. 4 and Fig. 5a. Fig 7b shows that the masticatory JMs produced during grazing are less affected by background noise than those produced during rumination. This could be due to the difference in the energy spectral density of the masticatory JMs produced in both grazing and rumination compared to that of the background noise[59].

Usage Notes

The 'AudioInformation.xlsx' is a spreadsheet file that provides specific information of the audio recordings that were obtained from the continuous monitoring of dairy cows. The sheet called "Audiofile properties" describes the Latin-square design for this experiment, which could be useful to analyze variations related to animals, experimental days or recording systems. Additionally, the correspondence between the direction of the microphones (inwards/outwards) and the channels in the audio recordings elaborated by the experts is also indicated. It should be noted that some errors may have occurred in the channel assignment due to the diverse sound quality detected across audio recordings. Any observations or particularities presented in the audio recordings are also mentioned. The sheet named "Cattle activities" specifies the kind of animal behavior categories and annotations presented in the audio recordings. This enables users to filter activities of interest.

Audio recordings qualities can vary greatly due to differences in animals, microphones and recording channels. We hypothesize that these variations were caused by differences in microphone response, microphone setup at the onset of recordings, and microphone movement during recordings. The sheet named "Audio quality" shows the values of the quality parameters of the audio recordings, using a background color scale from green to red to indicate high- and low-quality sound, respectively. This enables users to choose the optimal audio recordings or apply signal enhancement techniques, among other options. We recommend listening to the audio recordings in stereo or mono, depending on their preferred comfort and result, as this can vary from user to user due to differences in hearing capacity and audio signal intensity. We suggest listening in stereo for audio recordings with high-quality sound and listening only to the channel corresponding to the microphone facing inward for those with low-quality sound, as indicated in the 'AudioDescription.xlsx' file.

The subset enclosing information of JMs produced during grazing and rumination can be used as a standalone dataset for JMs analysis, or as an audiovisual guide to generate new JM labels using other audio recordings from foraging or rumination. We encourage users to utilize this subset as a reference for generating new JM labels from other grazing or rumination audio recordings.

The data described in this article is released under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license, indicating that it may be used for non-commercial purposes. We encourage users to cite this paper when using the data for proper attribution.

Code availability

The code for automatically adjusting the timesteps of masticatory JM labels and for technical validation is available (Data Citation 3[60]). All code was written in Python 3.8.10 and distributed under the MIT license. Small changes should be conducted in the scripts by specifying the audio file path of the execution environment.

References

1. Slob, N., Catal, C. & Kassahun, A. Application of machine learning to improve dairy farm management: A systematic literature review. *Prev. Vet. Medicine* **187**, 105237, https://doi.org/10.1016/j.prevetmed.2020.105237 (2021).
2. Lovarelli, D., Bacenetti, J. & Guarino, M. A review on dairy cattle farming: Is precision livestock farming the compromise for an environmental, economic and social sustainable production? *J. Clean. Prod.* **262**, 121409, https://doi.org/10.1016/j.jclepro.2020.121409 (2020).
3. Michie, C. *et al.* The internet of things enhancing animal welfare and farm operational efficiency. *J. Dairy Res.* **87**, 20–27, https://doi.org/10.1017/S0022029920000680 (2020).
4. Tzanidakis, C., Tzamaloukas, O., Simitzis, P. & Panagakis, P. Precision livestock farming applications (plf) for grazing animals. *Agriculture* **13**, https://doi.org/10.3390/agriculture13020288 (2023).
5. Banhazi, T. M. *et al.* Precision livestock farming: an international review of scientific and commercial aspects. *Int. J. Agric. Biol. Eng.* **5**, 1–9, https://doi.org/10.3965/j.ijabe.20120503.001 (2012).
6. Hodgson, J. & Illius, A. W. *The ecology and management of grazing systems* (Wallingford (United Kingdom) CAB International, 1998).
7. Garcia, R., Aguilar, J., Toro, M., Pinto, A. & Rodriguez, P. A systematic literature review on the use of machine learning in precision livestock farming. *Comput. Electron. Agric.* **179**, 105826, https://doi.org/10.1016/j.compag.2020.105826 (2020).
8. Connor, E. Invited review: Improving feed efficiency in dairy production: challenges and possibilities. *Animal* **9**, 395–408, https://doi.org/10.1017/S1751731114002997 (2015).
9. White, R. R., Hall, M. B., Firkins, J. L. & Kononoff, P. J. Physically adjusted neutral detergent fiber system for lactating dairy cow rations. i: Deriving equations that identify factors that influence effectiveness of fiber. *J. dairy science* **100**, 9551–9568, https://doi.org/10.3168/jds.2017-12765 (2017).
10. Nørgaard, P., Nadeau, E. & Randby, Å. T. *A new Nordic structure evaluation system for diets fed to dairy cows: a meta analysis*, chap. 9, 112–120 (Wageningen Academic Publishers, 2011).
11. Abeni, F. & Galli, A. Monitoring cow activity and rumination time for an early detection of heat stress in dairy cow. *Int. journal biometeorology* **61**, 417–425, https://doi.org/10.1007/s00484-016-1222-z (2017).
12. Schirmann, K., von Keyserlingk, M., Weary, D., Veira, D. & Heuwieser, W. Technical note: Validation of a system for monitoring rumination in dairy cows. *J. Dairy Sci.* **92**, 6052–6055, https://doi.org/10.3168/jds.2009-2361 (2009).
13. Bristow, D. J. & Holmes, D. S. Cortisol levels and anxiety-related behaviors in cattle. *Physiol. & Behav.* **90**, 626–628, https://doi.org/10.1016/j.physbeh.2006.11.015 (2007).
14. DeVries, T., Beauchemin, K., Dohme, F. & Schwartzkopf-Genswein, K. Repeated ruminal acidosis challenges in lactating dairy cows at high and low risk for developing acidosis: Feeding, ruminating, and lying behavior. *J. Dairy Sci.* **92**, 5067–5078, https://doi.org/10.3168/jds.2009-2102 (2009).

15. Paudyal, S. *et al.* Rumination time and monitoring of health disorders during early lactation. *Animal* **12**, 1484–1492, https://doi.org/10.1017/S1751731117002932 (2018).

16. Osei-Amponsah, R. *et al.* Heat stress impacts on lactating cows grazing australian summer pastures on an automatic robotic dairy. *Animals* **10**, 869, https://doi.org/10.3390/ani10050869 (2020).

17. Beauchemin, K. A. Ingestion and mastication of feed by dairy cattle. *Vet. Clin. North Am. Food Animal Pract.* **7**, 439–463, https://doi.org/10.1016/S0749-0720(15)30794-5 (1991).

18. Beauchemin, K. Invited review: Current perspectives on eating and rumination activity in dairy cows. *J. Dairy Sci.* **101**, 4762–4784, https://doi.org/10.3168/jds.2017-13706 (2018).

19. Pahl, C., Hartung, E., Grothmann, A., Mahlkow-Nerge, K. & Haeussermann, A. Rumination activity of dairy cows in the 24 hours before and after calving. *J. Dairy Sci.* **97**, 6935–6941, https://doi.org/10.3168/jds.2014-8194 (2014).

20. Kovács, L., Kézér, F., Ruff, F. & Szenci, O. Rumination time and reticuloruminal temperature as possible predictors of dystocia in dairy cows. *J. Dairy Sci.* **100**, 1568–1579, https://doi.org/10.3168/jds.2016-11884 (2017).

21. Schirmann, K., Chapinal, N., Weary, D., Vickers, L. & Von Keyserlingk, M. Rumination and feeding behavior before and after calving in dairy cows. *J. dairy science* **96**, 7088–7092, https://doi.org/10.3168/jds.2013-7023 (2013).

22. Pahl, C., Hartung, E., Mahlkow-Nerge, K. & Haeussermann, A. Feeding characteristics and rumination time of dairy cows around estrus. *J. Dairy Sci.* **98**, 148–154, https://doi.org/10.3168/jds.2014-8025 (2015).

23. Reith, S. & Hoy, S. Relationship between daily rumination time and estrus of dairy cows. *J. Dairy Sci.* **95**, 6416–6420, https://doi.org/10.3168/jds.2012-5316 (2012).

24. Dolecheck, K. A. *et al.* Behavioral and physiological changes around estrus events identified using multiple automated monitoring technologies. *J. dairy science* **98**, 8723–8731, https://doi.org/10.3168/jds.2015-9645 (2015).

25. Aquilani, C., Confessore, A., Bozzi, R., Sirtori, F. & Pugliese, C. Review: Precision livestock farming technologies in pasture-based livestock systems. *Animal* **16**, 100429, https://doi.org/10.1016/j.animal.2021.100429 (2022).

26. Mahmud, M. S., Zahid, A., Das, A. K., Muzammil, M. & Khan, M. U. A systematic literature review on deep learning applications for precision cattle farming. *Comput. Electron. Agric.* **187**, 106313, https://doi.org/10.1016/j.compag.2021.106313 (2021).

27. Riaboff, L. *et al.* Predicting livestock behaviour using accelerometers: A systematic review of processing techniques for ruminant behaviour prediction from raw accelerometer data. *Comput. Electron. Agric.* **192**, 106610, https://doi.org/10.1016/j.compag.2021.106610 (2022).

28. Lovarelli, D. *et al.* Development of a new wearable 3d sensor node and innovative open classification system for dairy cows' behavior. *Animals* **12**, 1447, https://doi.org/10.3390/ani12111447 (2022).

29. Andriamandroso, A., Bindelle, J., Mercatoris, B. & Lebeau, F. A review on the use of sensors to monitor cattle jaw movements and behavior when grazing. *Biotechnol. Agron. Société et Environnement* **20**, https://doi.org/10.25518/1780-4507.13058 (2016).

30. Ferrero, M. *et al.* A full end-to-end deep approach for detecting and classifying jaw movements from acoustic signals in grazing cattle. *Eng. Appl. Artif. Intell.* **121**, 106016, https://doi.org/10.1016/j.engappai.2023.106016 (2023).

31. Li, G., Xiong, Y., Du, Q., Shi, Z. & Gates, R. S. Classifying ingestive behavior of dairy cows via automatic sound recognition. *Sensors* **21**, 5231, https://doi.org/10.3390/s21155231 (2021).

32. Duan, G. *et al.* Short-term feeding behaviour sound classification method for sheep using lstm networks. *Int. J. Agric. Biol. Eng.* **14**, 43–54, https://doi.org/10.25165/j.ijabe.20211402.6081 (2021).

33. Wang, K., Wu, P., Cui, H., Xuan, C. & Su, H. Identification and classification for sheep foraging behavior based on acoustic signal and deep learning. *Comput. Electron. Agric.* **187**, 106275, https://doi.org/10.1016/j.compag.2021.106275 (2021).

34. Ungar, E. D. & Rutter, S. M. Classifying cattle jaw movements: comparing iger behaviour recorder and acoustic techniques. *Appl. animal behaviour science* **98**, 11–27, https://doi.org/10.1016/j.applanim.2005.08.011 (2006).

35. Martínez Rau, L., Chelotti, J. O., Vanrell, S. R. & Giovanini, L. L. Developments on real-time monitoring of grazing cattle feeding behavior using sound. In *2020 IEEE International Conference on Industrial Technology (ICIT)*, 771–776, 10.1109/ICIT45562.2020.9067192 (2020).

36. Laca, E. A., WallisDeVries, M. F. *et al.* Acoustic measurement of intake and grazing behaviour of cattle. *Grass Forage Sci.* **55**, 97–104, https://doi.org/10.1046/j.1365-2494.2000.00203.x (2000).

37. Galli, J. *et al.* Monitoring and assessment of ingestive chewing sounds for prediction of herbage intake rate in grazing cattle. *Animal* **12**, 973–982, https://doi.org/10.1017/S1751731117002415 (2018).

38. Ritter, C., Mills, K. E., Weary, D. M. & von Keyserlingk, M. A. Perspectives of western canadian dairy farmers on the future of farming. *J. Dairy Sci.* **103**, 10273–10282, https://doi.org/10.3168/jds.2020-18430 (2020).

39. Cockburn, M. Review: Application and prospective discussion of machine learning for the management of dairy farms. *Animals* **10**, https://doi.org/10.3390/ani10091690 (2020).

40. Vanrell, S. R. *et al.* Audio recordings dataset of grazing jaw movements in dairy cattle. *Data Brief* **30**, 105623, https://doi.org/10.1016/j.dib.2020.105623 (2020).

41. Jung, D.-H. *et al.* Deep learning-based cattle vocal classification model and real-time livestock monitoring system with noise filtering. *Animals* **11**, 10.3390/ani11020357 (2021).

42. Pandeya, Y. R., Bhattarai, B. & Lee, J. Visual object detector for cow sound event detection. *IEEE Access* **8**, 162625–162633, https://doi.org/10.1109/ACCESS.2020.3022058 (2020).

43. Chelotti, J. O. *et al.* An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle. *Comput. Electron. Agric.* **173**, 105443, https://doi.org/10.1016/j.compag.2020.105443 (2020).

44. Martinez-Rau, L. S., Weißbrich, M. & Payá-Vayá, G. A 4 μ w low-power audio processor system for real-time jaw movements recognition in grazing cattle. *J. Signal Process. Syst.* 1–18, https://doi.org/10.1007/s11265-022-01822-y (2022).

45. Martinez-Rau, L. S. *et al.* A robust computational approach for jaw movement detection and classification in grazing cattle using acoustic signals. *Comput. Electron. Agric.* **192**, 106569, https://doi.org/10.1016/j.compag.2021.106569 (2022).

46. Deniz, N. N. *et al.* Embedded system for real-time monitoring of foraging behavior of grazing cattle using acoustic signals. *Comput. electronics agriculture* **138**, 167–174, https://doi.org/10.1016/j.compag.2017.04.024 (2017).

47. Watt, L. *et al.* Differential rumination, intake, and enteric methane production of dairy cows in a pasture-based automatic milking system. *J. Dairy Sci.* **98**, 7248–7263, https://doi.org/10.3168/jds.2015-9463 (2015).

48. Milone, D. H., Galli, J. R., Cangiano, C. A., Rufiner, H. L. & Laca, E. A. Automatic recognition of ingestive sounds of cattle based on hidden markov models. *Comput. Electron. Agric.* **87**, 51–55, https://doi.org/10.1016/j.compag.2012.05.004 (2012).

49. Chelotti, J. O. *et al.* A real-time algorithm for acoustic monitoring of ingestive behavior of grazing cattle. *Comput. Electron. Agric.* **127**, 64–75, https://doi.org/10.1016/j.compag.2016.05.015 (2016).

50. Bosi, M. & Goldberg, R. E. *MPEG-1 Audio*, 265–313 (Springer US, Boston, MA, 2003).

51. Chelotti, J. O., Vanrell, S. R., Galli, J. R., Giovanini, L. L. & Rufiner, H. L. A pattern recognition approach for detecting and classifying jaw movements in grazing cattle. *Comput. Electron. Agric.* **145**, 83–91, https://doi.org/10.1016/j.compag.2017.12.013 (2018).

52. Galli, J. R. *et al.* Discriminative power of acoustic features for jaw movement classification in cattle and sheep. *Bioacoustics* **29**, 602–616, https://doi.org/10.1080/09524622.2019.1633959 (2020).

53. Kilgour, R. J. In pursuit of "normal": A review of the behaviour of cattle at pasture. *Appl. Animal Behav. Sci.* **138**, 1–11, https://doi.org/10.1016/j.applanim.2011.12.002 (2012).

54. Phillips, C. *Cattle behaviour and welfare* (John Wiley & Sons, 2008).

55. Laca, E. A., Ungar, E. D., Seligman, N. G., Ramey, M. R. & Demment, M. W. An integrated methodology for studying short-term grazing behaviour of cattle. *Grass Forage Sci.* **47**, 81–90, https://doi.org/10.1111/j.1365-2494.1992.tb02250.x (1992).

56. Benvenutti, M. A., Pavetti, D. R., Poppi, D. P., Gordon, I. J. & Cangiano, C. A. Defoliation patterns and their implications for the management of vegetative tropical pastures to control intake and diet quality by cattle. *Grass Forage Sci.* **71**, 424–436, https://doi.org/10.1111/gfs.12186 (2016).

57. Loizou, P. C. *Speech enhancement: theory and practice* (CRC press, 2013).

58. Martinez-Rau, L. S. *et al.* Open dataset of acoustic recordings of foraging behavior in dairy cows. *GitLab* https://gitlab.com/luciano.mrau/acoustic_dairy_cow_dataset/-/tree/master/data/sound_quality (2023).

59. Oppenheim, A. V., Willsky, A. S., Nawab, S. H. & Ding, J.-J. *Signals and systems*, vol. 2 (Prentice hall Upper Saddle River, NJ, 1997).

60. Martinez-Rau, L. S. *et al.* Open dataset of acoustic recordings of foraging behavior in dairy cows. *GitLab* https://gitlab.com/luciano.mrau/acoustic_dairy_cow_dataset (2023).

61. Brand, A., Allen, L., Altman, M., Hlava, M. & Scott, J. Beyond authorship: attribution, contribution, collaboration, and credit. *Learn. Publ.* **28**, 151–155, https://doi.org/10.1087/20150211 (2015).

Acknowledgements

The authors would like to express their gratitude to the staff at the KBS Robotic Dairy Farm for their invaluable support and assistance during the completion of this study. The operation of this farm and research was possible through funding provided by USDA-NIFA MICL0222 and MICL0406 projects, and support from Michigan State University AgBioResearch. This work has been funded by various organizations including Universidad Nacional del Litoral, Argentina, with projects CAID 50620190100080LI and 50620190100151LI, Universidad Nacional de Rosario, Argentina, with projects 80020180300053UR, 2016-AGR266 and 2013-AGR216, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina, with project 2017-PUE-sinc(i). Support for weather data was provided by the KBS LTER Program (NSF Award DEB 2224712).

Author contributions statement

Individual authorships and contributionships are described using the terms described by the Contributor Roles Taxonomy (CRediT) author statement[61]. L.S.M.R: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. J.O.C: Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization. M.F: Validation, Data Curation, Writing - Review & Editing. S.A.U: Methodology, Design, Formal analysis, Investigation, Data Curation, Writing - Review & Editing. A.M.P: Data Curation, Writing - Review & Editing. L.D.V: Methodology, Writing - Review & Editing. L.L.G: Funding acquisition, Investigation, Methodology, Project administration, Supervision. H.L.R: Funding acquisition, Investigation, Methodology, Project administration, Writing – review & editing, Supervision. J.R.G: Funding acquisition, Investigation, Methodology, Design, Data Curation, Writing - Review & Editing.

Competing interests

The authors declare no competing financial interests.

Figures & Tables

	Cow 1 (ID: 2936)	Cow 2 (ID: 2909)	Cow 3 (ID: 2948)	Cow 4 (ID: 21036)	Cow 5 (ID: 2976)	Mean ± SD
Weight [kg]	653	651	674	657	663	659.7 ± 9.4
Lactation number	3	3	2	2	3	2.6 ± 0.5
Days in milk [d]	130	125	68	141	62	105.2 ± 37.3
Milk yield [kg/d]	35.4	37.8	44.1	40.3	44.0	40.3 ± 3.8

Table 1. Specific traits and description of the dairy cows used to acquire the audio recordings. The measurements were carried out on the first day of the experiment.

Cow	Days (Date)					
	1 (Jul 31)	2 (Aug 4)	3 (Aug 6)	4 (Aug 11)	5 (Aug 13)	6 (Aug 18)
1 (ID: 2936)	1	2	3	4	5	1
2 (ID: 2909)	2	3	4	5	1	2*
3 (ID: 2948)	3	4	5	1	2	3
4 (ID: 21036)	4	5	1	2	3	4
5 (ID: 2976)	5	1	2	3	4	5

* Audio recording repeated on August 19 due to problems associated with the microphone connector and recorder.

Table 2. Latin-square design for recording systems, cows and days.

Cow	Days (Date)						
	1 (Jul 31)	2 (Aug 4)	3 (Aug 6)	4 (Aug 11)	5 (Aug 13)	6 (Aug 18)	6 (Aug 19)*
Total Rain [mm]	0	3.048	4.064	0	1.016	1.778	0
Average Wind speed [m/s]	1.976	1.079	1.102	1.651	2.895	1.419	1.034
Wind Vector cells: Direction [m/s]	88.6	79.7	73.38	249.3	91.3	253.4	266.2
Average Radiation [kW/m²]	0.238	0.279	0.193	0.251	0.068	0.078	0.166
Total Radiation [MJ/m²]	0.02851	0.03348	0.02318	0.03011	0.008138	0.009370	0.01988
Average Air Temperature [ºC]	17.93	20.69	20.9	21.46	17.02	20.21	20.66
Maximum Air Temperature [ºC]	23.32	28.2	25.13	27.86	19.04	21.82	26.84
Minimum Air Temperature [ºC]	12.73	13.14	17.69	13.98	13.31	19.11	15.86
Relative Humidity [%]	85.9	91.6	96.8	80.2	92.5	91.2	96.4

* Extra day to complete the six experimental days.

Table 3. Weather conditions during audio recording trials.

Figure 1. Recording system used to record the acoustic signals composed of inward and outward facing microphones (a). Wired microphones were covered by an elastic headband (b) and plugged (c) to a recorder housed inside a weather proof case attached to the top side of a halter neck strap (d).

```
Dataset
├─ Daily_recordings_subset
│  ├─ Day1
│  │  ├─ D1RS1ID2936.zip
│  │  │  ├─ D1RS1ID2936P1.mp3
│  │  │  ├─ D1RS1ID2936P1.csv
│  │  │  ├─ D1RS1ID2936P1.txt
│  │  │  └─ ...
│  │  ├─ D1RS2ID2909.zip
│  │  ├─ D1RS3ID2948.zip
│  │  ├─ D1RS4ID21036.zip
│  │  └─ D1RS5ID2976.zip
│  ├─ Day2
│  │  └─ ...
│  └─ ...
├─ Jaw_movements_subset
│  ├─ JM_grazing.wav
│  ├─ JM_grazing.txt
│  ├─ JM_grazing.csv
│  ├─ JM_grazing_adjusted.txt
│  ├─ JM_grazing_adjusted.csv
│  ├─ JM_rumination.wav
│  ├─ JM_rumination.txt
│  ├─ JM_rumination.csv
│  ├─ JM_rumination_adjusted.txt
│  ├─ JM_rumination_adjusted.csv
│  ├─ D3RS4ID2909P3_JM.txt
│  ├─ D3RS4ID2909P3_JM.csv
│  ├─ D3RS4ID2909P3_JM_adjusted.txt
│  └─ D3RS4ID2909P3_JM_adjusted.csv
├─ Additional_files_subset
│  ├─ BehaviorLabelsDescription.pdf
│  ├─ MasticatoryJMDescription.pdf
│  └─ MP3AudioInformation.xlsx
└─ Index.xlsx
```

Daily_recordings_subset: Subset of audio recordings (MP3 format) and label (CSV and TXT format) files. Files are sorted by trial day and recording device, and package into files (ZIP format).
The name of files provides coded information of trial day, recording system, cow ID and recording partition. For example file 'D4RS3ID2976P3.mp3' refers to data collected on trial day 4, recording system 3, cow 2976 and recording period 3 (corresponding to the 12-18 h of that trial).

Jaw_movements_subset: Subset of audio (WAV format) and label (CSV and TXT format) files of jaw movements associated with grazing and rumination activities. Labels for individualized jaw movements were generated by close inspection of audio files by two experts in sound signal processing and by use of a script coded in phyton.

Additional_files_subset: Subset of additional description (PDF format) and spreadsheet (XLSX format) files with detailed explanations and information.

Figure 2. Internal dataset organization in bundled files and naming.

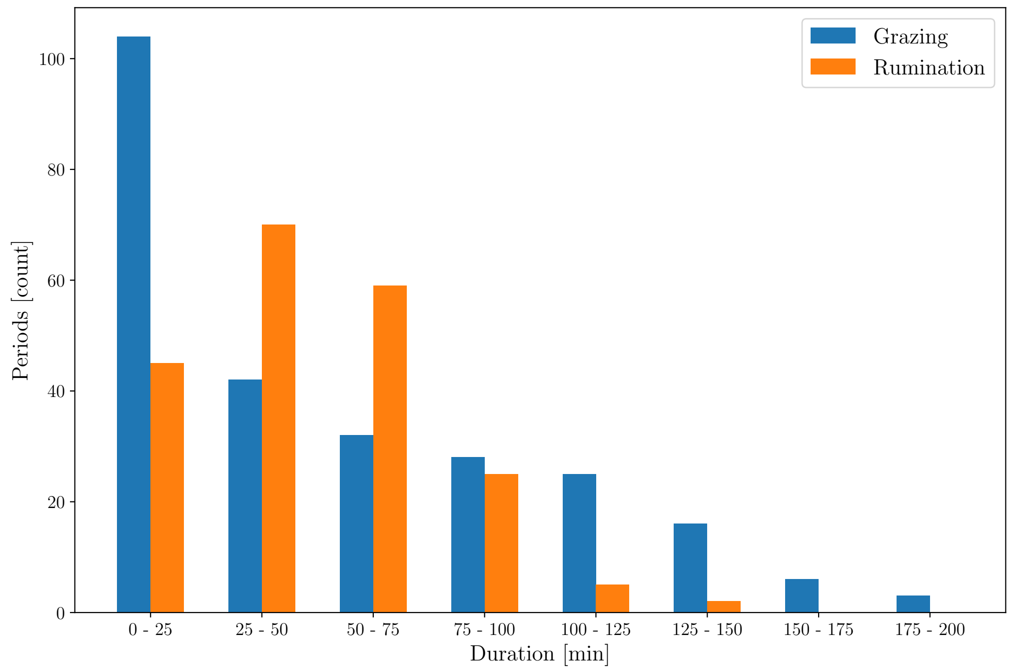


Figure 3. Histogram showing the frequency distribution of the duration of grazing and rumination bouts grouped in 25 min intervals. A total of 257 grazing bouts and 206 rumination bouts are present in the dataset.

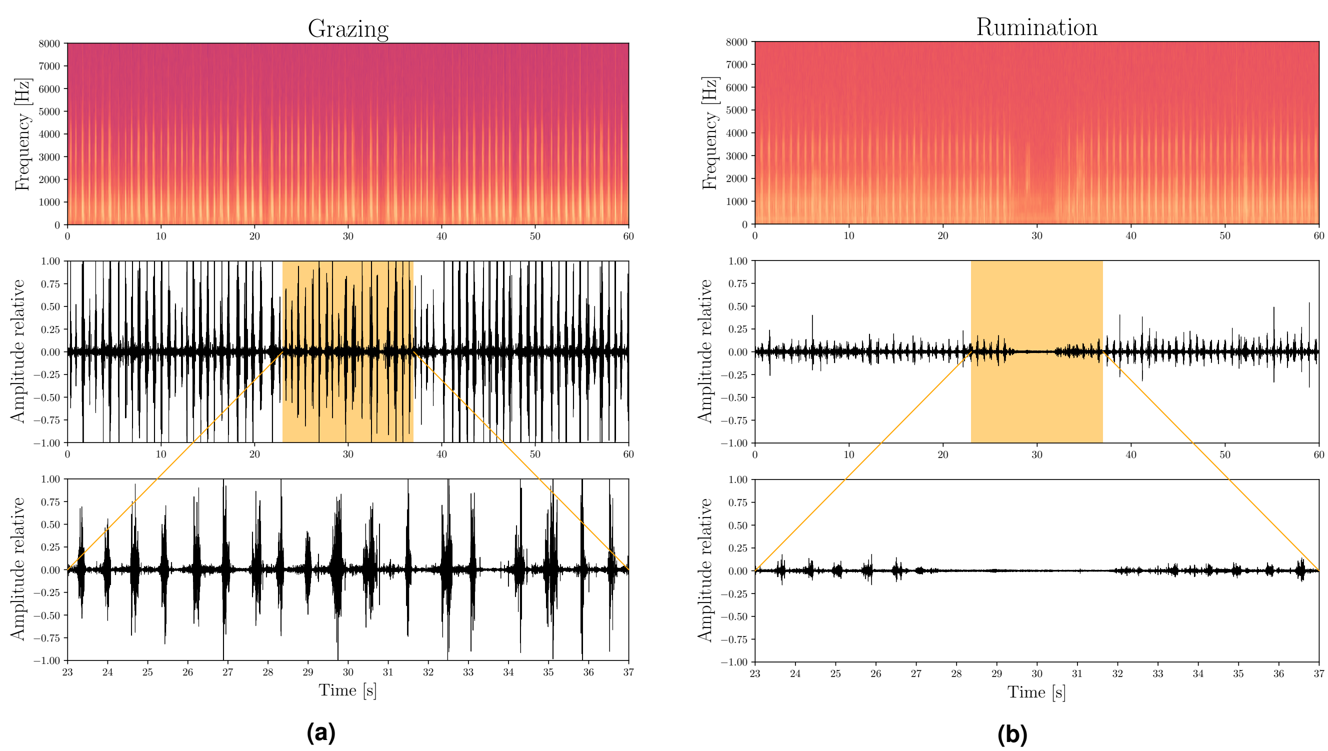
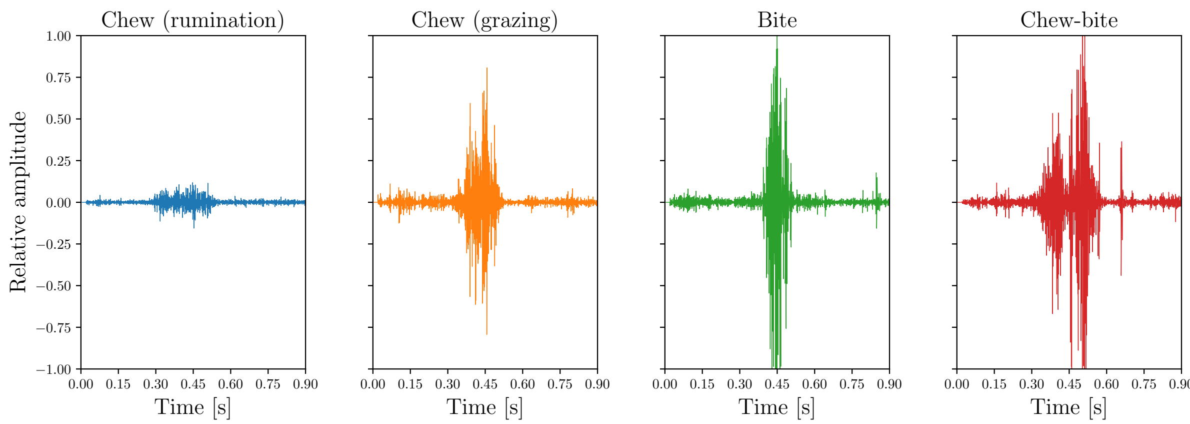
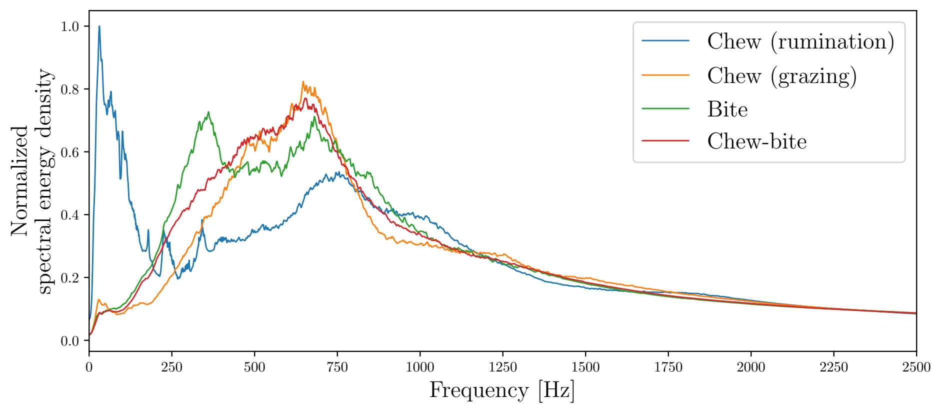


Figure 4. Spectrogram and waveform (with zoom) of foraging audio signals associated with (a) grazing and (b) rumination activities.

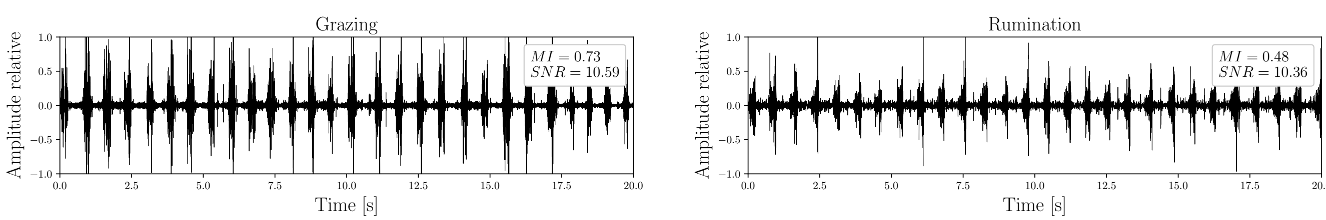


(a)

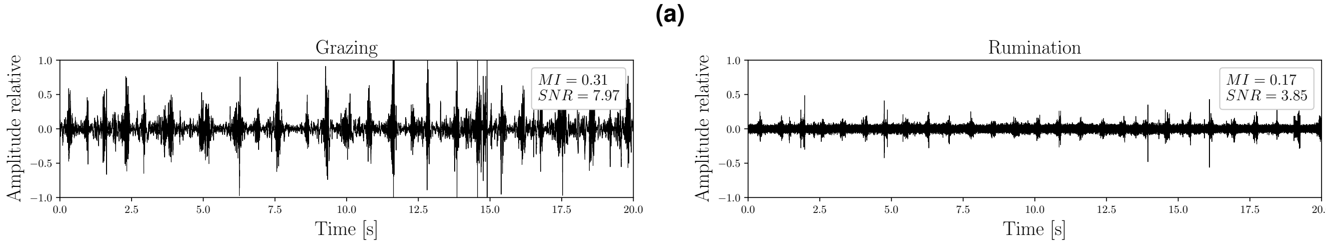


(b)

Figure 5. Typical waveform (a) and average spectrum (b) for the different types of masticatory JMs: chew produced during rumination and chew, bite and chew-bite produced during grazing. Energy spectra were averaged across all masticatory JMs and normalized to the maximum value.

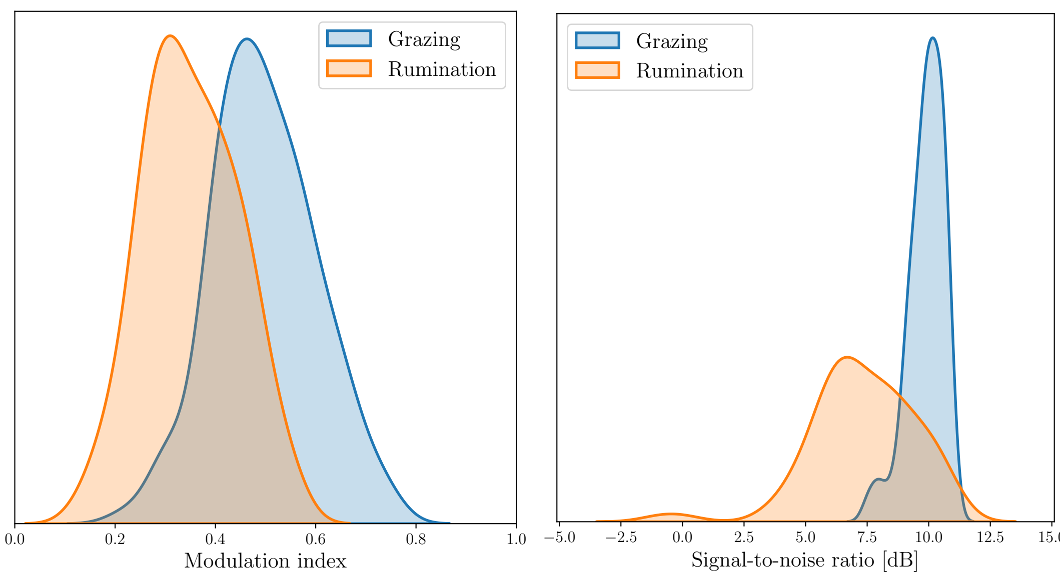


(a)



(b)

Figure 6. Waveforms of segments of audio recordings with (a) high- and (b) low-quality sound.



(a) **(b)**

Figure 7. Frequency distribution of the audio recording quality in terms of (a) the modulation index and (b) the signal-to-noise ratio.