



OPEN

DATA DESCRIPTOR

MSPB: a longitudinal multi-sensor dataset with phenotypic trait measurements from honey bees

Yi Zhu¹✉, Mahsa Abdollahi¹, Ségolène Maucourt², Nico Coallier³, Heitor R. Guimarães¹, Pierre Giovenazzo² & Tiago H. Falk¹

We present a one-year-long multi-sensor dataset collected from honey bee colonies (*Apis mellifera*) with rich phenotypic measurements. Data were collected non-stop from April 2020 to April 2021 from 53 hives located at two apiaries in Québec, Canada. The sensor data included audio features, temperature, and relative humidity. The phenotypic measurements contained beehive population, number of brood cells (eggs, larva and pupa), *Varroa destructor* infestation levels, defensive and hygienic behaviors, honey yield, and winter mortality. Our study is amongst the first to combine a wide variety of phenotypic trait measurements annotated by apicultural science experts with multi-sensor data, which facilitate a broader scope of analysis. We first summarize the data collection procedure, sensor data pre-processing steps, and data composition. We then provide an overview of the phenotypic data distribution as well as a visualization of the sensor data patterns. Lastly, we showcase several hive monitoring applications based on sensor data analysis and machine learning, such as winter mortality prediction, hive population estimation, and the presence of an active and laying queen.

Background & Summary

Apis mellifera, commonly referred to as honey bees, hold significant commercial value not only for the hive products but also due to their fundamental role as pollinators. Their pollination activities are beneficial for both agricultural crops and biodiversity^{1–3}. Historically, beekeepers have relied on manual and visual inspections to monitor their beehives^{4,5}, but these are time-consuming and disruptive to the colonies. While beekeepers typically inspect their hives on a regular basis (e.g., weekly to monthly) during pollination or honey production, important changes in colonies can occur within that time frame, making continuous monitoring essential⁴. Meanwhile, large-scale colony losses have been observed worldwide in recent years, caused by multiple stressors acting independently or synergistically, such as pesticides, pathogens, parasites, climate changes, and many other factors^{6–9}. Human inspections of hives can offer valuable insights, such as early detection of foulbrood and identification of *Varroa* mite infestations, but they require significant time and cannot provide continuous monitoring.

Recently, computer-aided automated beehive monitoring systems have been developed to address the limitations of human management⁴. Existing systems typically place sensors inside the hive to record the environmental changes as well as colony status. High-level features are then extracted from sensor data then fed into machine learning (ML) models for downstream tasks, such as the early prediction of colony winter survivability¹⁰, estimation of colony strength¹¹, and discrimination of different bee activities¹², just to name a few. Among different sensor modalities, temperature and relative humidity are known as the two most widely used ones that are closely related to hive status. For example, studies have shown an increased amount of honey production and lower mortality rates under well-controlled temperature levels^{13–15}. The optimum relative humidity of a beehive varies between 50% to 60%, while higher or lower levels are shown to have an impact on brood development and mite infestation levels^{16,17}. Abrupt changes in internal hive temperature and relative humidity have also been used to detect the presence of an active queen bee, as well as to predict swarming^{18,19}.

More recent studies have explored the use of acoustic sensors to infer the present state of the colony⁴. Honey bees contract their thoracic wing muscles, creating a vibration that generates complex acoustic signals^{20,21}. Compared to conventional modalities, audio signals provide a more direct measurement of the hive status and

¹INRS-EMT, Université du Québec, Montréal, Canada. ²Département de biologie, Université Laval, Laval, Canada.

³Nectar Technologies Inc., Montréal, Canada. ✉e-mail: Yi.Zhu@inrs.ca

Dataset	#Hives	Modality	Location	#Days	Annotations	Availability
NU-Hive ^{†28}	2	Audio Temp Humid CO ₂	IT	5	Queen presence	Subset
BUZZ ^{‡51,52}	6	Audio	UT, USA	109	Sound type (e.g., bee, noise, etc.)	Upon requests
OSBH ^{‡53}	6	Audio	EU; USA; AU	2	Queen presence	Subset
MSPB (Ours)	53	Audio Temp Humid	QC, CA	365	Population Honey yield Queen conditions Defensive behavior Hygienic behavior Winter mortality	Fully available

Table 1. Comparison of MSPB with other publicly available beehive sensor datasets. [†]NU-Hive and OSBH described here are publicly available subsets of their parental databases. [‡]BUZZ comprises four subsets (where BUZZ 3 is a subset of BUZZ 4), we report the metadata after aggregating all subsets.

reflects the instant response of bees to outer changes. Given the advantages of beehive acoustics, a substantial body of work has applied ML algorithms to audio data to detect the presence of a queen bee^{22–24}, swarming^{22,25,26}, as well as other activities^{27,28}.

While automated beehive monitoring systems are advantageous in multiple aspects, massive amounts of data are needed to enable accurate ML model training and decision-making^{29,30}. To this end, we curated the Multi-modal Sensor dataset with Phenotypic trait measurements from honey Bees (MSPB), which is composed of audio, temperature, and relative humidity data recorded from 53 hives located in Québec, Canada during a one-year period. This paper presents a detailed description of the data collection procedure, sensor data pre-processing, data records, and our preliminary findings based on statistical analysis and ML-driven hive monitoring tools.

To highlight the novelty of our database, a comparison with existing publicly available databases is summarized in Table 1. It should be noted that with NU-Hive and OSBH, only subsets were found available at <https://zenodo.org/records/1321278> while the full versions are not publicly available. Therefore, the descriptions in Table 1 correspond only to these subsets. Despite the multiple efforts made, it can be seen from Table 1 that existing datasets are limited in terms of hive sample size, time range, number of sensor modalities, and variety of phenotypic trait measurements, thus making the development of ML tools challenging. As such, we introduce the MSPB dataset to tackle these limitations and to provide the research community with a richer dataset to help advance beehive monitoring.

The MSPB database was collected non-stop for one year from April 15th, 2020 to April 14th, 2021, using 53 hives located at two apiaries in Québec, Canada. The audio, temperature, and relative humidity data were recorded synchronously throughout the whole year, resulting in a total of 365 days of data. Compared to existing publicly available databases, the MSPB dataset covers the longest time range and provides much richer phenotypic traits annotated by apicultural science experts, including the colony honey bee population, honey yield, queen-related conditions (e.g., swarming, supersedure, egg laying), health status (e.g., *Varroa* mite infestation, winter survivability), and multiple behavioral evaluation results. These phenotypic trait measurements, together with the large number of hives, would allow for a more systematic analysis of sensor data to understand honey bee activities.

Methods

Honey bee colonies. Our study was conducted with 53 honey bee colonies selected amongst the livestock of the Centre de recherche en sciences animales de Deschambault (CRSAD) Québec, Canada (N46°40.270, W10°71.500). Selected colonies had sister queens, were of equivalent strength (6–7 frames of bees/brood) and housed in 10-frame Langstroth hives mounted with a Plastic *Varroa* Stainless Steel Screen Bottom Board (Propolis-etc..., Saint-Pie, QC, Canada; SD-1500). Colonies were managed for honey production with a single brood chamber and placed in two farmland sites, i.e., the Dubuc (N46°42.27, W71°34.33) and Côté (N46°44.302, W71°28.284) apiaries. The Dubuc apiary is on a small hill and is exposed to more wind compared to the Côté Apiary. Figure 1 shows the photos of hives taken at the two apiaries, as well as a closer view of one hive with one brood chamber at the start of experiment.

Hive management. During the summer, extra honey supers were added over the brood chamber and separated by a queen excluder. The positioning of the sensor on the top of the frames of the brood chamber and the overall structure of a hive are illustrated in Fig. 2. At the beginning of September, honey supers were removed, and colonies had one brood chamber. Fall feeding started on September 15th, 2020 and all colonies were given 24 liters of a sucrose 2:1 solution using a top box feeder (Wooden Miller feeder # FE-1100 at Propolis-etc..., Beloeil, Québec). Colonies received a Thymovar anti-*Varroa* treatment (Propolis-etc..., Saint-Pie, QC, Canada; TH-1110), applied as per label, starting on September 17th, 2020, followed by an oxalic acid treatment (Propolis-etc..., Saint-Pie, QC, Canada; AO-1201) on October 28th, 2020 (drip method: 35 gL⁻¹ in a sucrose 1:1 solution, 5 ml between frames of the hive body crowded with honey bees). Colonies were wintered indoors in an environmentally controlled room (4–5°, 50–60% relative humidity) from November 14th, 2020 to April 14th, 2021 and then moved into a spring apiary in Deschambault, Québec near the bee research facility until mid-May.

Beehive monitoring system overview. The beehive monitoring system comprises two fundamental components: (1) a multi-modal sensor system with continuous data recording, and (2) phenotypic traits annotated by apicultural science experts on a bi-weekly basis. An overview of the beehive monitoring system can be seen in Fig. 3. With the sensor data collection, a multi-modal sensor is positioned at the top of the central frame of the brood box of a Langstroth hive housing honey bees (see Fig. 2a). This sensor is capable of concurrent

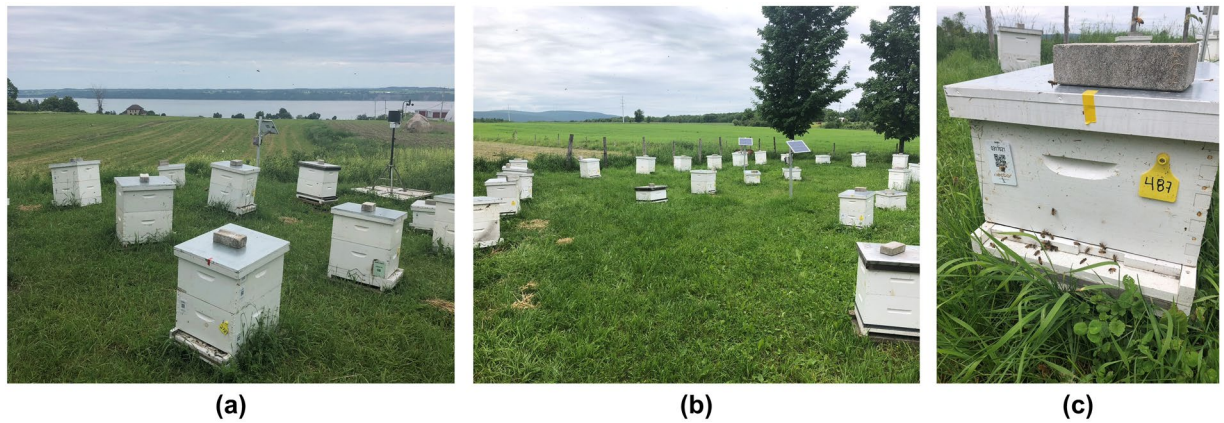


Fig. 1 Photos of (a) Dubuc hives, (b) Côté hives, and (c) a closer view of one hive chamber.

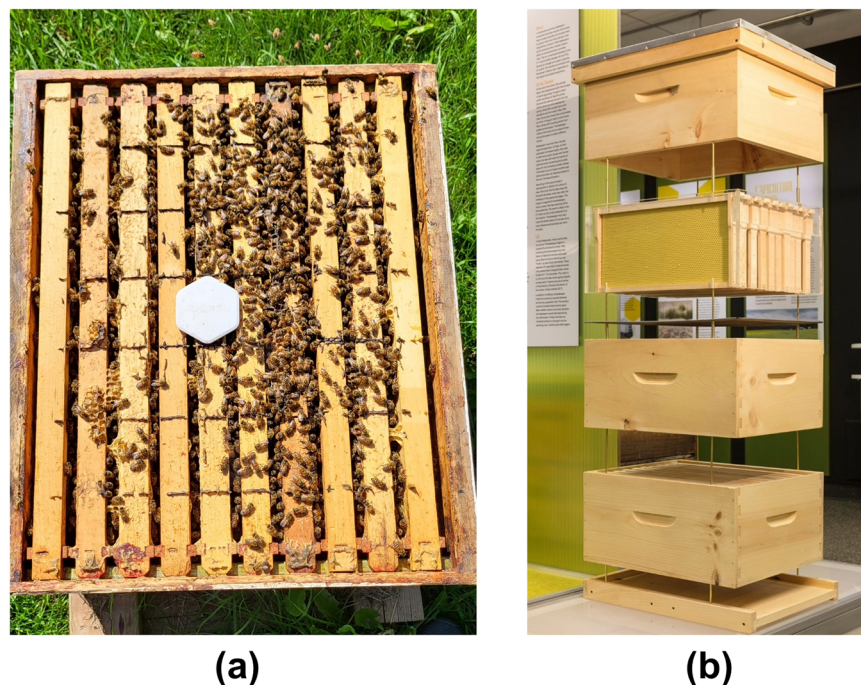


Fig. 2 Illustrations of (a) the position of the sensor over the frames of the bottom brood chamber on the base board, and (b) the decomposed overall structure. From bottom-up: base board, two brood chambers, queen excluder, honey super showing frames and box, top cover roof.

recording of audio, relative humidity, and temperature data at regular intervals of 5 min, 15 min, and 15 min respectively. The recorded sensor data is wirelessly transmitted to a central data aggregator powered by solar energy and securely stored in the cloud. The sensor data was collected 24 hours a day, 7 days per week from April 15th, 2020 to April 14th, 2021. Besides continuous sensor recording, apicultural science experts visited hives bi-weekly to monitor the hive status and conducted evaluations on a regular basis. Colony phenotypic trait measurements, such as honey bee population, honey yield, and health status, were also collected, hence providing valuable context to interpret the sensor data.

Phenotypic trait measurements. Phenotypic trait measurements were collected and are summarized in Table 2. Details about the collection procedure are outlined as follows:

- **Number of brood cells.** The number of brood cells (eggs + larvae + pupae) can be used to indicate the hive strength^{31–33}, which was estimated by measuring both width and length of the brood area on each side of every brood frame. The rectangular area obtained was multiplied by 0.8 to compensate for the elliptic form of the brood pattern³¹. These values were added to calculate the total brood area in each colony. A factor of 25 cells per 6.25 cm² was used to convert the area to obtain a number of brood cells, which was related to the cell size of a standard worker cell (Propolis-etc...; WH-1302).

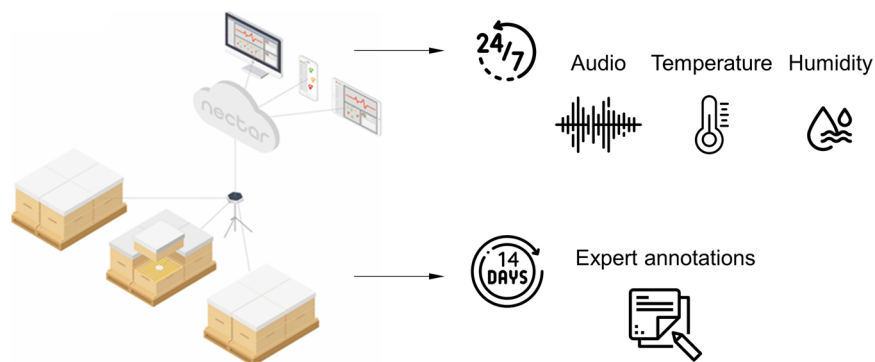


Fig. 3 An overview of the beehive monitoring system with continuous multi-modal data recording and phenotypic traits. Sensor data were collected 24/7 at a fixed interval of 5 min for audio, and 15 min for temperature and relative humidity. Phenotypic traits were annotated by apicultural science experts every two weeks.

Data type	Collection date (YYYY-MM-DD)	Description
Brood surface	2020-06-09	Number of capped, uncapped, and total cells in the 2 brood chambers
Honey production	2020-11-06	Weight of honey produced by each colony
Defensive behavior	2020-07-21 & 2020-08-04	Number of stings on a leather flag waved over the chamber for 2 minutes
Hygienic behavior	2020-07-23 & 2020-08-06	Liquid nitrogen frozen-brood test and evaluation after 24h
<i>Varroa</i> infestation level	2020-05-06 & 2020-08-13	Evaluation of the <i>Varroa</i> infestation level with the <i>Varroa</i> wash technique

Table 2. Types of phenotypic trait measurements annotated by apicultural science experts.

- **Frames of bees.** For each colony, we measured the size of the bee cluster by opening the hive and counting the number of frames occupied by bees (FoB), as seen from top to bottom. Since one hive could comprise multiple honey supers, we repeated this counting process for each super. The number of frames covered by bees per super ranges from 0 to 10.
- **Honey yield.** Colonies are equipped with honey supers, placed above a queen excluder. Each colony has at least two honey supers during honey flow, each with stretched comb frames. Each colony was weighed before and after harvest using a platform scale (CAS-USA, East-Rutherford, NY, USA; CAS CI-2001BS), where the weight difference was considered as the honey yield.
- **Hygienic behavior.** Hygienic behavior is evaluated with a test that measures the cleaning capacity of a colony's bees on a percentage level. During the test, the colony was opened, and a comb was selected containing a solid slab of sealed worker brood in the pupal stage, with pupae having pink or purple eyes. Two PVC tubes (5.08 cm internal diameter) were pressed to the comb's midrib. The number of empty (i.e., missed) cells in each tube was counted. Liquid nitrogen was then applied at a rate of 300 ml per tube to freeze the brood. Frames were marked and returned to the colony. After 24 hours, the number of cells removed was counted and divided by the total number of cells, the resultant percentage value was then used as a measure of the hygienic behavior^{34,35}. This test was carried out twice during the 2020 summer, at the honey flow low period, i.e., during a period when nectar resources are poor for honey bees, early August.
- **Defensive behavior.** The flag stinging test was used to measure a colony's defensive behavior. When the hive was open, a flag was waved rhythmically (amplitude approximately 20 cm), with an oscillation every 2 s, 5 to 10 cm above the colony's brood chamber for 2 minutes. The flag is a piece of black suede leather (measuring 10 × 8 cm) suspended from a piece of light wood (0.7 × 0.5 × 100 cm). After the test, the number of stings on the flag was counted^{36,37}.
- ***Varroa* infestation level.** Two methods were employed to evaluate the *Varroa* infestation level at the end of August. The first was the natural mite-fall method, where the *Varroa destructor* infestation level was assessed by the natural mite-fall method using sticky boards placed on the bottom boards of hives^{38,39}. Mites that had fallen on sticky boards (5-7 days) were counted to obtain a daily mite drop value. The second followed the alcohol washing method³⁹. This method involved sampling around 200 to 300 honey bees from a colony honey frame in 70% alcohol. For each sample, the honey bees were counted and then placed in a *Varroa* EasyCheck or *Varroa* Mite Test Bottle type sampling device and covered with 70% ethanol. The sampling device was then placed on a horizontal shaker at 150 rpm for 5 minutes. The honey bees were then removed and the number of mites in the sampling device was counted. The process was repeated until there were no more mites in the sample (up to 3 washes). The total number of *Varroa* mites counted gave the number of *Varroa* mites per 100 bees (i.e., total number of *Varroa* mites in the sample × 100/number of bees in the sample)^{40,41}.

Modality	Freq.	Fs	Feature	Description
Audio	5 min	15625 Hz	Hive power	Overall power within [122 Hz-519 Hz] in dB
			Audio Band Density Ratio	Ratio of hive power over the overall power
			Audio Density Variation	Amount of changes in audio power within 0.98 s
			Audio Band Coefficients	Power of 16 linearly spaced frequency bands in dB
Temperature	15 min	—	Raw value in degree	Temperature insides the chambers
Relative Humidity	15 min	—	Raw value in percentage	Relative humidity insides the chambers

Table 3. List of stored features computed from multi-modal sensor data. Temperature and relative humidity data were stored as raw values, while audio data were stored as acoustic features. Freq: The frequency (time intervals) of the sensor data collection. Fs: Sampling rate of the sensor signals.

Sensor data collection and pre-processing. Table 3 presents the sensor data modalities, the frequency in which parameters were extracted, and in the case of the audio modality, the microphone sampling frequency. To optimize bandwidth, battery life, and storage requirements, the audio data were not stored in the raw waveform format. Instead, four types of acoustic features were computed to encapsulate relevant information, including hive power, audio band density ratio, audio density variation, and audio band coefficients. Audio signals were originally sampled at 15625 Hz, the fast Fourier transform (FFT) was subsequently computed over 30 non-overlapping frames once every 5 min, each frame with a length of 512 points (i.e., 0.98 s of audio data considered every 5 min). We denote the resultant spectrogram as $X_{j,k}$, where j corresponds to the frame index ($j \in \{0, 1, \dots, 29\}$), and the frequency bin index as k ($k \in \{0, 1, \dots, 256\}$).

The hive power (P_{hive}) reflects the overall power between 122 Hz and 515 Hz, where:

$$ABD_j = \sum_{k=4}^{k=17} \|X_{j,k}\|^2 \quad \text{and} \quad (1)$$

$$P_{hive} = 10 \lg \left(\frac{\sum_{j=0}^{j=29} ABD_j}{30 \times (17 - 4 + 1)} \right) = 10 \lg \left(\frac{\sum_{j=0}^{j=29} ABD_j}{420} \right). \quad (2)$$

Here, ABD_j corresponds to the audio band density at the time frame j , and the extra division by 14 in the denominator normalizes the energy for a per-bin representation, but can be removed by a single shift in dB. The selection of such frequency range was based on our exploratory analysis, where non-bee sounds generally manifest at different frequency ranges (e.g., rainfall) or exhibit abrupt fluctuations in signal power (e.g., human speech and thunder sounds). The audio band density ratio (ABD_{ratio}) is defined as the ratio of hive power with regard to the power of the whole frequency range (AD):

$$AD_j = \sum_{k=4}^{k=256} \|X_{j,k}\|^2 \quad (3)$$

$$ABD_{ratio} = \frac{ABD_j}{AD_j} \quad (4)$$

The audio band density variation ($ABDR$) reflects the amount of changes within the 0.98 s:

$$ABDR = 10 \lg \frac{\max(AD_j)}{\min(AD_j)} \quad (5)$$

For a consistent audio event like a bee sound, this value will be at a lower dB, while for other events such as thunder or human speech, the density variation is expected to be at a higher value. Lastly, we chose 16 linearly spaced frequency bins and computed the power, respectively, as the 16 coefficients:

$$Bin_N = \frac{N \times 15625}{512}, \quad n \in \{4, 5, 6, \dots, 19\}, \quad (6)$$

$$Coef_N = 10 \lg \left(\frac{\sum_{j=0}^{j=29} \|X_{j,N}\|^2}{30} \right). \quad (7)$$

The aforementioned acoustic features were extracted in real-time and stored as the final format on the cloud server. As for relative humidity and temperature, the raw data were preserved in percentage and degrees, respectively, where higher values corresponding to increased relative humidity and higher temperatures within the beehive boxes.

Data Records

The MSPB dataset is made fully available at the Zenodo repository⁴². The sensor data and phenotypic traits were stored separately in .csv format, each of which was further divided into two files based on the time range, resulting in a total of four .csv files. To distinguish summer and winter data, those collected between April 15th, 2020 and November 6th, 2020 received a 'D1' label in the file name, while the data between November 6th, 2020 and April 14th, 2021 were labelled as 'D2'. The detailed file composition is summarized in Table 4. The total size of the shared files is about 500 MB.

D1 and D2 sensor data are both paired with (1) the time stamp (date and time) of the data collection, (2) hive ID, which is a unique number to identify each hive, (3) apiary ID, which indicates the apiary location of the hive, (4) temperature values, (5) relative humidity values, and (6) twenty audio features. The D1 phenotypic traits file has three sub-sheets, which details (1) the visit date and time of the human evaluations, as well as the evaluation tasks, (2) the population size of the colonies measured at each visit, (3) other phenotypic trait measurements, such as *Varroa* infestation levels, defensive and hygienic behavior, honey yield, etc. During the D2 period, hives were maintained in the winter chambers and only evaluated once in the Spring to check their winter survival rate. Hence, the D2 phenotypic traits file contains the survival status, as well as the mortality causes (if any) of all hives.

Technical Validation

Apiary and hive-level population. Figure 4 demonstrates an overview of changes in the number of hives during D1 and D2 for two apiaries. With an initial total of 53 hives in April, five and three hives failed in Côté and Dubuc during D1, respectively. Six out of the eight failed hives went through a queen change, which leads to an overturn of the colonies and affect their behavior, especially in honey production. We therefore flagged these hives as 'failed' along with the rest that died of chalk brood, and did not measure the honey production. Among the remaining 45 hives, 10 did not survive the winter (4 in Côté and 6 in Dubuc). Figure 5 further shows the changes in hive population for the two apiaries based on FoB counted manually at six human evaluations. Hives that died before the date of evaluation were removed from the plot. The average FoB increases from 10 to 20 from June 9th to July 9th, then stabilizes at between 20 and 25 with the peak seen at the end of July. Meanwhile, it was observed that the population size varies markedly across hives. For example, while the majority (1st to 3rd quartile) had 20 to 30 FoB during July and August, as few as 10 FoB were seen in smaller colonies. Compared to the hives in Côté, Dubuc hives demonstrate larger variations in populations across colonies.

Phenotypic trait representation. The phenotypic traits of hives from both apiaries are depicted in Fig. 6. Overall, the phenotypic data distributions of the two apiaries follow similar patterns. Hives are equally distributed across two apiaries, each with 26-27 hives (see Fig. 6a). In terms of total honey production, the majority produced 30 to 60 kg, while 11 hives produced less than 10 kg of honey. Among the low-productivity hives, queen cells were observed in five hives, which led to the division or failure of the entire colony. Regarding bee population, the average of total brood (i.e., eggs, larva and pupa of honey bees) was approximately 25,000, with the majority varying between 20,000 to 40,000 (see Fig. 6c). During the summer, apicultural experts also evaluated the *Varroa* condition, cleaning capacity, and hive defensive behavior on a regular basis. We calculated the average from all evaluations and summarized these data in Fig. 6d–f respectively. At the end of August, *Varroa* infestations were below the economic threshold level of 3% (see Fig. 6d). Majority of the hives had a cleaning capacity of over 80% (see Fig. 6e); 11 hives once exhibited defensive behaviors (i.e., number of stings > 0) (see Fig. 6f).

To investigate the associations between different colony phenotypic traits, the Spearman's rank correlation coefficient r was calculated between honey yield, *Varroa* infection, hygienic behavior, and defensive behavior for each of the two apiaries (see Fig. 7). While different correlation patterns are seen across the two apiaries, the honey yield was consistently shown to be positively associated with the defensive behavior, with statistical significance seen for both apiaries ($p_{\text{côté}} < .01$; $p_{\text{dubuc}} < .05$). The *Varroa* mite infection and cleaning capacity were found with weaker but significant correlation only with the honey yield of hives in Côté. Meanwhile, though being insignificant, negative correlation values are seen with cleaning capacity and *Varroa* mite infection, suggesting that the hives with higher cleaning capacity suffer less from *Varroa* mite infection.

Temporal patterns of multi-modal sensor data. Similar to weather changes, the behaviors of honey bees follow a specific pattern that repeats on a daily and yearly level. Here, we show that such patterns can be captured by the multi-modal sensor data. With the yearly pattern, we first aggregated data points from each day and calculated the daily average for each hive. We then computed the mean and standard deviation across all hives to obtain the general pattern throughout the year. With the daily pattern, data points were aggregated hourly and averaged across the whole year. An average was then computed across all hives to obtain the 24 h changes.

Yearly changes. The changes of multi-modal sensor data from April 15th, 2020 to April 14th, 2021 are depicted in Fig. 8, annotated with the honey bee experts evaluations and interventions. Compared to temperature and humidity, the audio modality exhibits greater variations across time, many of which are conditioned on human activities. For example, the hive power increases and decreases accordingly when the number of supers were being added or reduced to each hive. Peaks in hive power are also seen during evaluations and treatments, such as during the Thymovar® *Varroa* treatment in mid-September and two behavior evaluations before and after August 1st. After being moved to the winter chambers, all hives manifest a similarly steady pattern, which is reflected by the reduced variation across time and the smaller standard deviation across hives. Furthermore, hive power also exhibits higher standard deviations across hives compared to the other modalities, indicating that bee acoustics might encapsulate richer information unique to each hive. In general, the sensor data are found to be a good indicator of the arousal level of honey bees, especially the audio.

File name	Sub-sheets	Variables	Description	
D1_sensor.csv & D2_sensor.csv	—	Published_at	Time stamp (YYYY-MM-DD HH24:MI:SS)	
		Hive ID	ID unique to each hive	
		Apiary ID	ID of the apiary	
		Temperature	Temperature in degree Celsius	
		Relative humidity	Humidity in percentage	
		Audio features	twenty audio feature values	
D1_ant.xlsx	Visit	Yard	Name of the apiary	
		Dates	Date of the visits	
		Arrival time	Time of arrival (HH-MM)	
		Departure time	Time of departure (HH-MM)	
		Manipulations	Details of human evaluations	
		Yard location	Coordinates of two apiaries	
	Evaluation [NUMBER]	Dates	Date of the population evaluation	
		Yard	Name of the apiary	
		Hive ID	ID unique to each hive	
		Number of boxes	Bottom chamber + honey supers	
	Phenotypic measurements	NoF covered by bees	Population measurement	
		Brood surface	Capped, uncapped, and total brood cells	
		Varroa infestation	Severity of <i>Varroa</i>	
		Defensive behavior	Defensiveness measured by flag test	
		Hygienic behavior	Cleaning capacity	
	Honey weight	Total honey produced during summer		
	READ ME	—	An introduction on file content	
	D2_ant.xlsx	—	Apiary	Name of the apiary
			Hive ID	ID unique to each hive
Mortality cause			Causes of the failed hives	
Weight			Weight values (kg) before and after winterization	
Bees frames			Population before and after winterization	
Syrup consumption			Syrup consumption (kg) during winterization	

Table 4. Structure of the multi-modal sensor data and phenotypic trait measurement files.

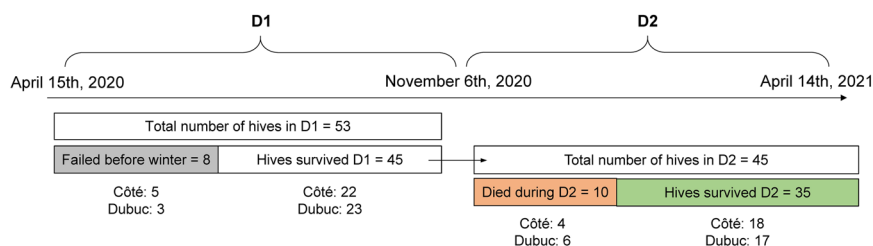


Fig. 4 Changes in number of hives during D1 and D2 for two apiaries.

Daily changes. Honey bees are known to follow a particular daily routine driven by their circadian clock⁴³. Figure 9 shows the summer-averaged 24 h pattern in hive power, relative humidity, and temperature, together with the sunrise and sunset time across all hives. From sunrise, the majority of worker bees leave the hives to forage⁴⁴, resulting in decreasing audio power and humidity. Between sunrise and 3PM, the hive power is maintained constantly low, while small fluctuations can be seen which could be due to the activities of remaining honey bees inside the hives or noise from the outside environment (e.g., rain, wind, human speech). After 3PM, a rapid increase is seen with hive power and temperature, suggesting the return of honey bees. The peak of hive power appears subsequently between 6-9PM, which is close to the sunset time, indicating that the majority of honey bees have returned and may have started cleaning honeycomb cells. After sunset, the hive power and temperature gradually decrease while the humidity slightly increases, these findings are in line with the decreased activities and mobility reported with honey bees during nighttime^{44,45}. Additionally, although outside humidity and temperature can have large variations throughout the day (e.g., environment humidity can vary from 40% to 90% on rainy days, temperature can vary up to $\pm 10^\circ\text{C}$), minor changes are seen with inner-hive temperature and humidity, suggesting the capability of honey bees to control the hive environment^{16,46}. In summary, the sensor data are found to be highly correlated with the daily activities of honey bees, demonstrating the potential to be used to infer and monitor honey bee behaviors.

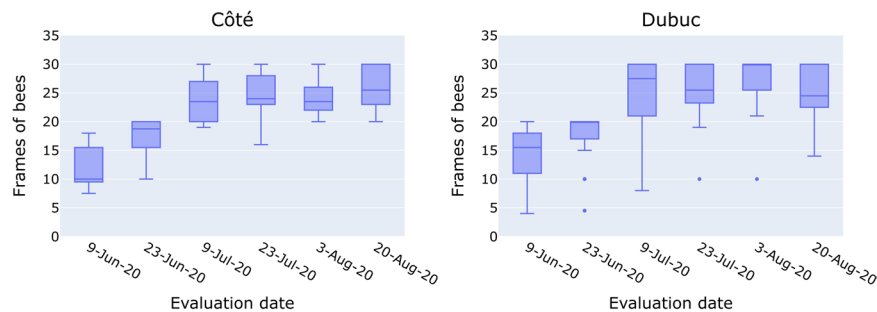


Fig. 5 Changes in hive population from June 9th, 2020 to August 20th, 2020 for Côté (left) and Dubuc (right). The box region covers the interquartile (IQR) range (25% to 75% percentages) of the data, and the whiskers are set to $1.5 \times$ IQR. When a group of hives have the similar frames of bees, the median line would be shown close to the box boundaries.

Early detection of beehive winter survivability. *Hive power and colony population as potential indicators.* For the past fifteen years, winter loss of honeybee colonies have been observed throughout the world. In Canada, about 30% of hives have failed during the winter months since 2007^{9,47}. However, very limited tools can be used by beekeepers to identify the high-risk colonies at an early stage^{9,47}. Our exploratory analysis, in turn, has shown that hives that failed during the winter manifested a different pattern in multi-modal sensor data compared to those that survived. Figure 10 depicts such change in behavior for the three signal modalities. As can be seen, the main difference is observed with the audio modality, where failed hives showed significantly lower average hive power than the survivors based on Welch's t-test (p -value = 0.028), indicating that the failed hives were less active before entering the wintering room. With the relative humidity and temperature, however, no significant difference has been found (p -value = 0.824 and 0.702 respectively).

While hive power is demonstrated with potentials to indicate winter survivability, it may be affected by the colony population, where populated hives might be expected to show larger hive power values. To investigate the acoustics-population relationship as well as to compare their respective predictive power for winter survivability, Fig. 11 provides a side-by-side visualization of hive power and colony population for survived and failed hives from the two apiaries. Since the colony population was only measured between June 9th and August 20th, we limit the hive power within the same time range for alignment with the population curves. Similar to the trend seen in Fig. 10, failed hives in both apiaries show lower average hive power than the survivors. When comparing the populations of survived and failed hives, the two apiary groups show different patterns. For hives in Côté, the two classes in general demonstrate similar trends, where the failed ones are observed with significantly larger populations on August 3rd (p -value = 0.012). While the failed hives in Dubuc have significantly smaller populations observed on June 9th (p -value = 0.05), July 9th (p -value = 0.01), and July 23rd (p -value = 0.01). Results here suggest that hive power could be more suitable than colony population as an indicator for winter survivability.

Additionally, the changes in colony populations do not follow the exact same pattern as that in hive power. Though the population is observed to rise with hive power between June 9th to July 9th, the hive power starts to decrease after August 3rd where the population still maintains at the plateau for hives in both apiaries. We further calculated the Spearman's correlation coefficient between the curves of hive power (sampled at the six evaluation dates) and the colony population per hive, then reported the distribution of the correlation values for each apiary (box plots on the left side of the curves). As indicated in Fig. 11, though positive average correlation values are obtained between the two measures, the individual correlation values vary from -0.4 to 1 for hives in Côté and from -0.8 to 1 for hives in Dubuc. These findings suggest that while hive power is correlated with colony population on a group-level, specifically during hive growing periods, such correlation is not consistent through time and varies across hives. Therefore, the acoustic features, such as hive power, cannot be substituted by colony population measures to predict winter survivability.

Multi-modal ML model for winter survivability prediction. Our recent study¹⁰ explored the use of the aforementioned multi-modal sensor data (audio, relative humidity, temperature) to predict winter survivability using only features computed from the summer months. In the experiments, all hives were divided into two categories: ones that survived the 2020-21 winter and ones that failed (see D2 hive division in Fig. 5). A hand-crafted feature set was then proposed and derived from the multi-modal data, which aggregated the descriptors from different time-levels (Fig. 12a). For classification, parameter-free unsupervised out-of-distribution detection models were used (e.g., isolation forest with predefined parameters) to detect the hives that were less likely to survive the winter. The top-performing model achieved an AUC-ROC score of 0.730 based only on the data from 1st May 2020 to 31st October 2020. A visualization of the decision boundary obtained from the isolation forest classifier can be found in Fig. 12b. Table 5 quantifies the importance of different sensor modalities by comparing the performance obtained by individual modality and their combinations. For the single-modality category, audio is shown with the highest predictive power (AUC-ROC 0.669). Improvement can be achieved when multi-modal features are concatenated, where the fusion of all three modalities lead to the best performance (AUC-ROC 0.730). Overall, these findings suggest that multi-modal features can indeed encompass indicators of the hive winter survivability.

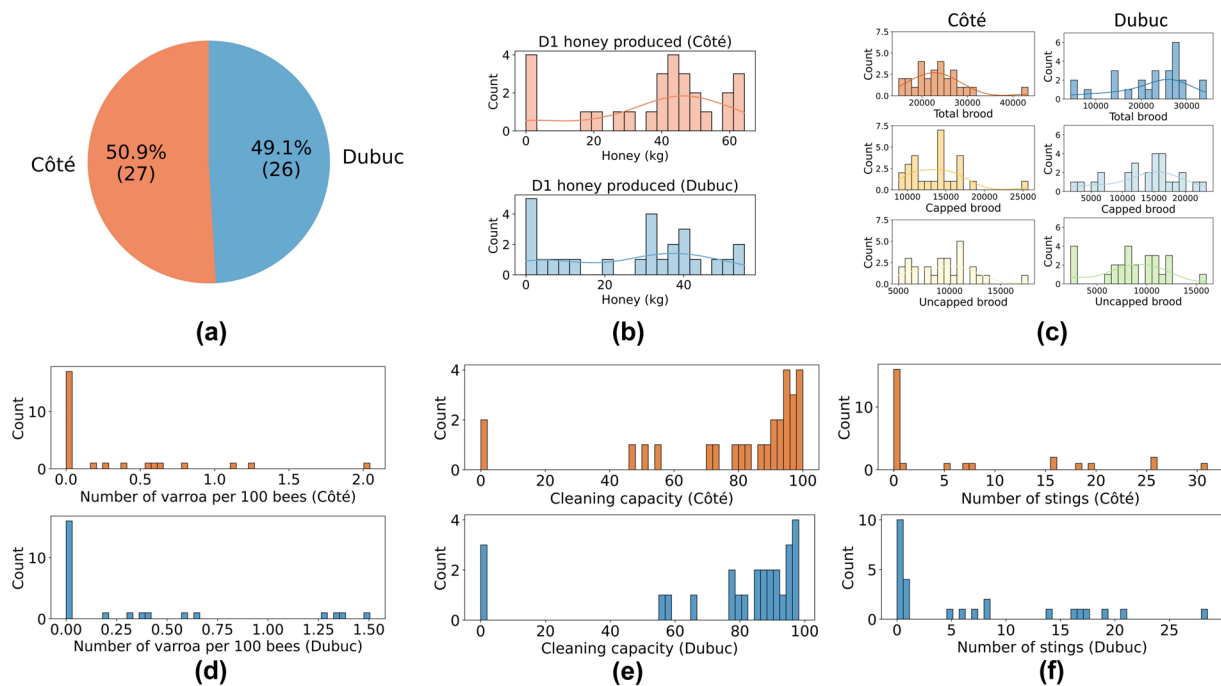


Fig. 6 Apiary-wise phenotypic data distribution: (a) Number of hives, (b) Honey yield, (c) Brood population, eggs, larvae and pupa, during spring colony build up in June, (d) *Varroa* infection levels, (e) Hygienic behavior quantified by cleaning capacity (percentage), and (f) Defensive behavior quantified by number of stings.

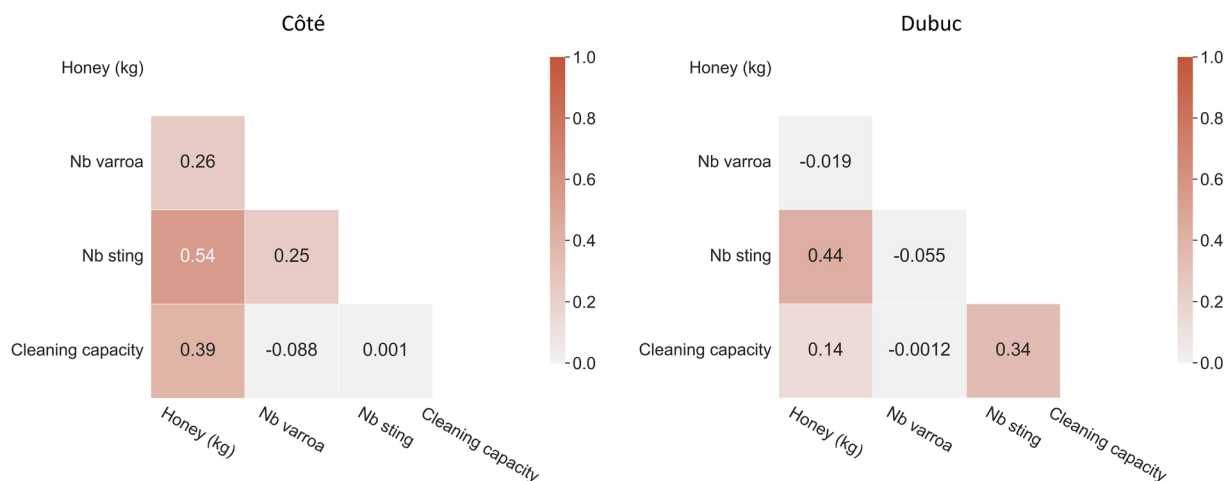


Fig. 7 Spearman's rank correlation coefficients between different phenotypic trait measurements. Darker color suggests stronger positive correlation. The most impactful factors for honey yield is shown to be the defensive behavior (i.e., number of stings), followed by the cleaning capacity. Additionally, no significant correlation is found between cleaning capacity and *Varroa* mite infection.

Population estimation. Colony population is one of the most important measures of hive strength³². In our study, for example, the colony population was estimated by apicultural experts, which was a time-consuming and arduous process. Recent studies have shown the usefulness of raw audio to predict hive population⁴⁸. An acoustic-based population estimation tool can assist with the evaluation of hive strength by providing continuous monitoring. Here, we developed a population estimation model using features extracted from multi-modal sensor data. Colonies in both apiaries from the six evaluation dates (see Fig. 5) were combined and treated as independent samples, which resulted in a total number of 318 data points (53×6). These were then divided into two classes based on the FoB, namely smaller hives comprised of $\text{FoB} < 20$ and larger hives with $\text{FoB} > 20$. We then extracted features (i.e., mean, 1st-order and 2nd-order deltas) from the sensor data three days before and after each evaluation date, and aggregated the daily features by computing low-level descriptors across the seven days (i.e., mean, standard deviation, kurtosis, and skewness). These features can be computed directly using the open-source Scipy

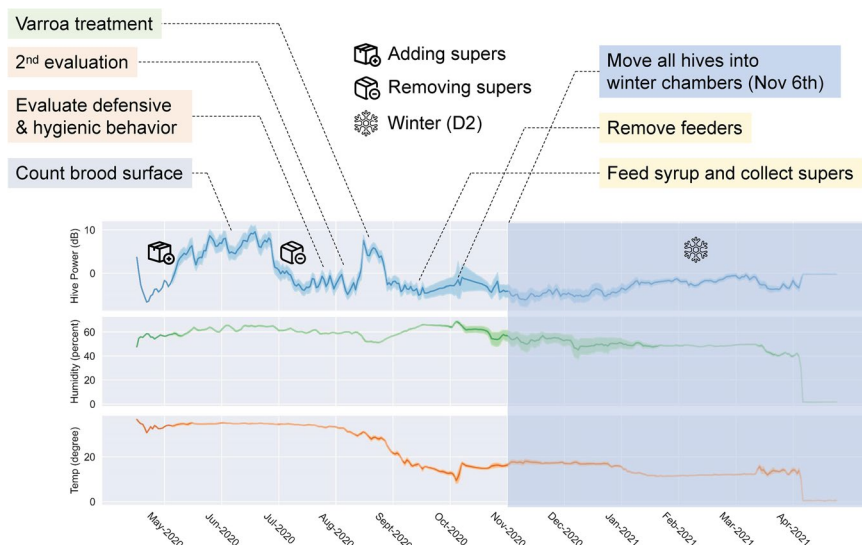


Fig. 8 Multi-sensor data from April 15th, 2020 to April 14th, 2021. Human evaluations are annotated on top of the plots. The lines are formed by connecting daily data points averaged across all hives. The shaded area represents one standard deviation from the average power of all hives.

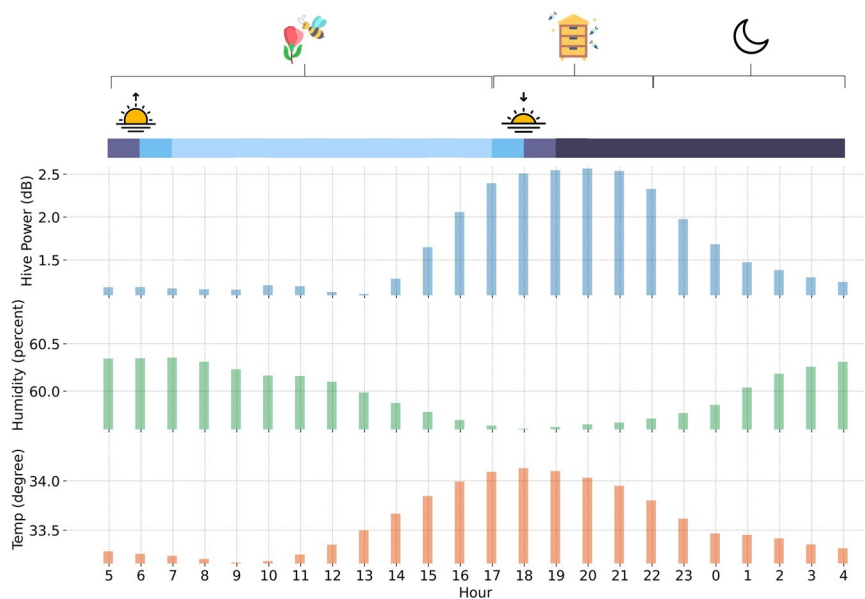


Fig. 9 24-hour changes of multi-sensor data together with the sunrise/sunset time. During the summer months, the sunrise time varies from 5AM to 7AM and the sunset time from 7PM to 9PM. The honeybee circadian clock is illustrated at the top of the plot, suggesting changes in bee behaviors across different times of the day.

library⁴⁹. In the end, 24 features were computed per sample, which encapsulate information about the hive status close to the evaluation dates. These features were then fed into a linear Support Vector Machine (SVM) for classification. We followed a random data partition with a train-test-ratio of 7:3 and used 3-fold cross-validation to obtain the optimal classifier. Since the majority of the samples had FoB > 20, we over-sampled the minority class to achieve a 1:1 ratio during training, and employed the balanced accuracy as the evaluation metric. As a result, the best classifier achieved a balanced accuracy of 65.8%. Fig. 13 reports the test results in a confusion matrix.

Likely affected by the class imbalance, the classifier was found better at recognizing larger colonies (True Positive Rate = 84.5%) and performed relatively worse with smaller colonies (True Negative Rate = 47.1%). While our results demonstrate the potential of using acoustic features for estimating the colony population, the described model only performs a rough estimation of the range of the colony population. Whether such modelling approach can be applied in real-world to assist with the population evaluation, on the other hand, requires a more systematic evaluation. It is hoped that the release of the MSPB dataset will allow machine learning experts to develop new tools to address this in future works.

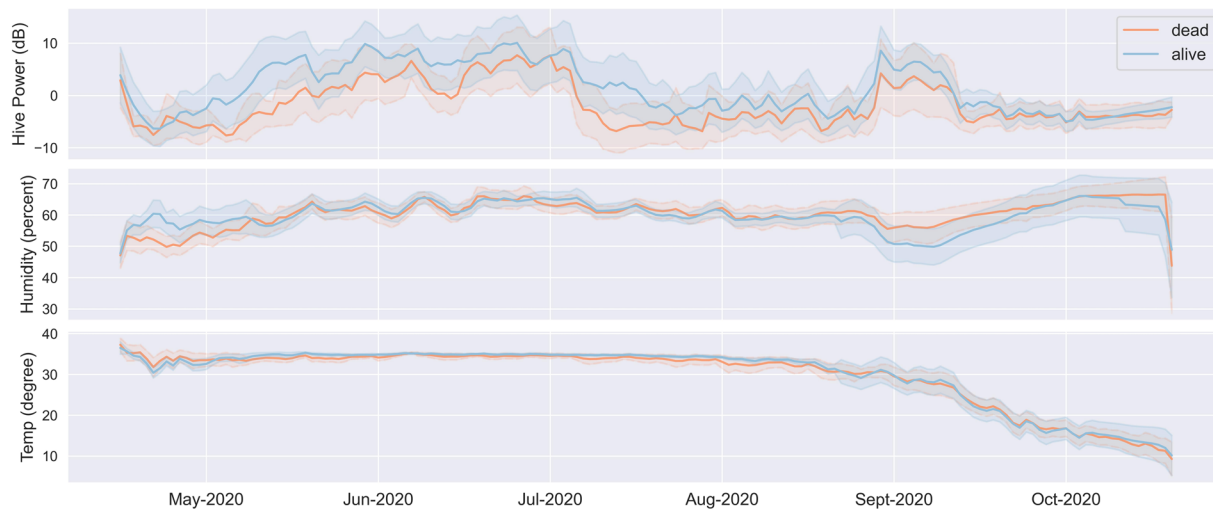


Fig. 10 Difference between hives that survived and failed in winter 2020–2021 based on multi-sensor data from April, 2020 to November, 2020. Data collected during wintering are not included in the plot. A statistical significant difference was found between the average hive power of failed and alive hives, while no significance was found for the other two modalities.

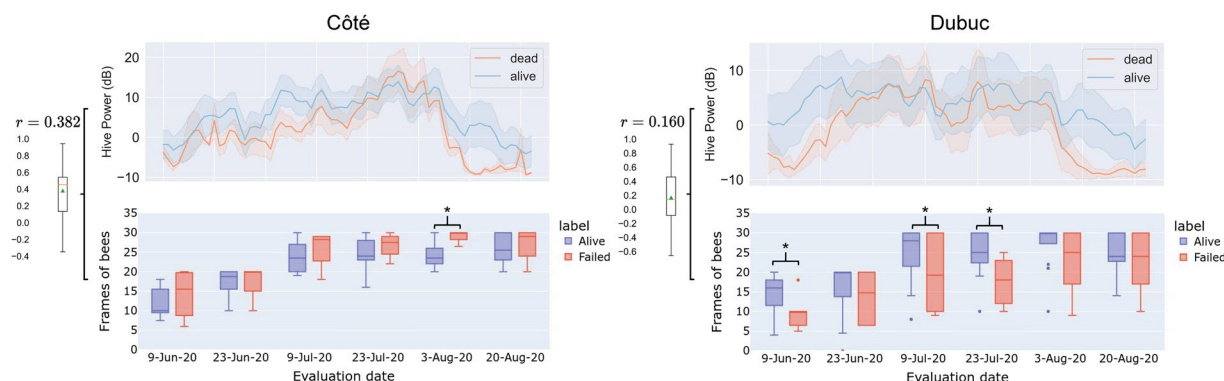


Fig. 11 Hive power and colony population distribution for survived and failed hives from the two apiaries. Statistical significance is indicated by the asterisks. Box plots on the side reflect the distribution of Spearman's correlation coefficient values calculated between the hive power and colony population.

Identification of the presence of queen cells. Queen cell is a special type of cell with an elongated shape, where a larva develops and matures into a new queen⁵⁰. There are usually two distinct reasons for the existence of queen cells, namely supersedure and swarming. Supersedure cells are made when the queen is old, ill, or missing, hence a new queen is needed to replace the old queen. Swarming cells, on the other hand, occur when the colony reproduces in the summer when resources are abundant and the colony is fully developed. As a result of swarming, a portion of the colony will leave with the existing queen, while the new queen will be raised and stay at the current hive. The identification of the presence of queen cells is crucial for beekeepers, since timely management is needed for both swarming and supersedure conditions.

Similar to the bi-weekly human evaluations conducted in our study, traditional methods rely on opening the hive and conducting visual observations (e.g., brood frames, population, etc.). Such methods are advantageous in terms of accuracy, but can be time-consuming and may result in missing the best window of time for taking actions, especially when managing large groups of hives located distantly. Here, we demonstrate the potential of using sensor data to help identify the existence of queen cells. In the MSPB dataset, there are six hives that were observed with queen cells, the details are provided in Table 6. Since the evaluations were performed bi-weekly, it is difficult to pinpoint the precise starting and end dates of the queen cells. We therefore label the days between two weeks before the observation of queen cells and the dates when new queen or queenless were found as positive (i.e., queen cells likely exist), which contains six time regions (see last column in Table 6). For the negative class, we use the days within the conjectured time range of queen cell existence from the hives without queen cells. In general, the positive samples represent the days when the queen cells likely existed in the hive, whereas the negative samples represent normal days. The feature extraction and ML model pipeline are similar to the ones implemented for the winter survivability prediction task, where the statistics (mean, std, kurtosis, skewness) are calculated from the daily-averaged multi-modal sensor data within the six time regions. These features

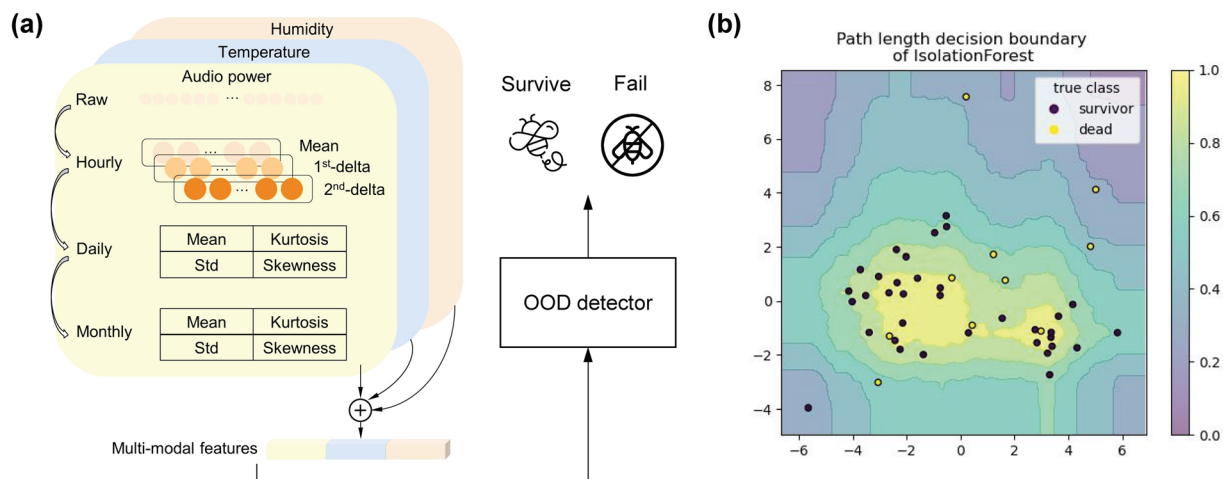


Fig. 12 Early-detection of beehive winter survival by using out-of-distribution detection models: (a) feature extraction from multi-modal sensor data and downstream detection, and (b) decision boundary obtained from an isolation forest detector, where majority of the failed hives are found deviated from the distribution center, though a few overlapped with the alive hives.

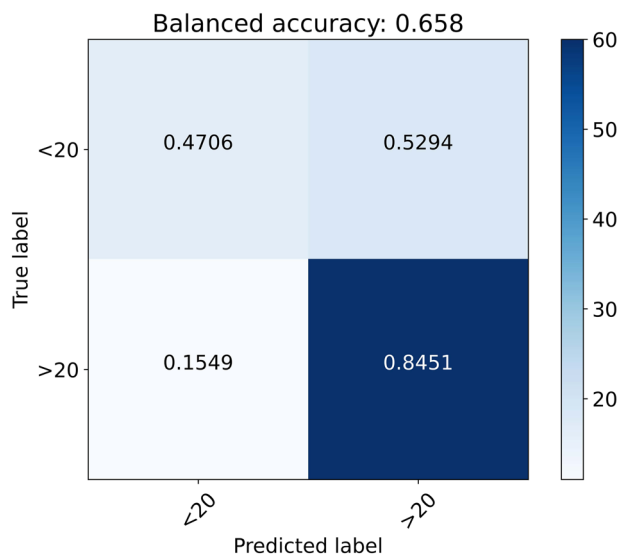


Fig. 13 Confusion matrix for colony population size classification using acoustic features. Hives with over 20 FoB and below 20 FoB were separated into two classes. A linear SVM classifier obtained a balanced accuracy of 65.8%.

Modality	Audio (A)	Temperature (T)	Humidity (H)	A+T	A+H	T+H	A+T+H
AUC-ROC	0.669	0.574	0.551	0.656	0.671	0.580	0.730

Table 5. Winter survivability prediction performance using different modalities and their combinations. Note that temperature and humidity data are collected by the multi-modal sensor placed at the bottom chamber, hence reflecting the environment inside the hives instead of the outside environment.

are then fed into an isolation forest classifier to detect if a sample belongs to the positive (queen cells) or the negative (normal) class. Figure 14(a) visualizes the projected features in a 2-dimensional space, where 5 out of the 6 hives with queen cells (marked as orange crosses) are found distant from the distribution center, indicating that the multi-sensor features are manifested differently when queen cells exist. Figure 14(b) shows the performance of the classification model. An average AUC-ROC of 0.646 is achieved across 5 repetitions, demonstrating the potential of using multi-sensor data for detecting the presence of queen cells.

Hive ID	Evaluation notes	Possible time range of queen cell existence
202046	emergency cells (2020/08/03), and queenless (2020/08/20)	07/19–08/03
202054	requeening cells (2020/06/09), and new queen (2020/06/23)	05/25–06/23
202213	queen cells (2020/06/22) and new virgin queen (2020/07/07)	06/07–07/07
202202	emergency cells (2020/08/06) and queenless (2020/08/21)	07/22–08/21
202209	queen cells (2020/06/09;2020/06/20;2020/07/07) and queenless (2020/08/21)	05/25–08/21
202210	queen cells (2020/06/22;2020/07/07) and new queen (2020/08/13)	06/06–08/13

Table 6. IDs of the hives that were found with queen cells, along with the human evaluation notes and conjectured possible time range of the queen cell existence. Since human evaluations were performed bi-weekly, the queen cell starting dates should be within 2 weeks prior to the date when queen cells were found.

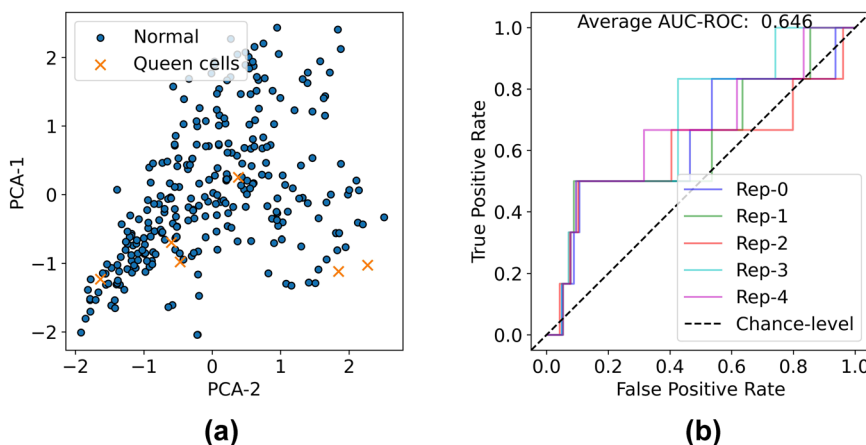


Fig. 14 Queen cell identification results obtained by a multi-modal detection system: (a) multi-modal features projected to a 2-dimensional space using PCA; and (b) ROC curves achieved from 5 rounds of model evaluations.

Usage Notes

File reading and post-processing. All data were shared as the comma separated values (CSV) files. The file reading can be easily done with Python using the Pandas library. The Python scripts used for post-processing, feature extraction, and machine learning tasks demonstrated in this paper can be found at our Github repository, listed in the section “Code Availability” below.

Potential new use cases. The present paper showcased the use of multimodal sensor data for only a few different hive monitoring tasks to validate the quality of the MSPB dataset. The data, however, is much richer and there are several new use cases that can lead to new insights about honeybee behavior. For example, while multiple modalities were collected, their fusion and the importance of each modality can be further investigated. Moreover, the rich phenotypic labels of the dataset can allow for many new ML applications to be developed, including but not limited to detection of *Varroa* infestation, automated characterization of hygienic and defensive behaviors, or honey production prediction. It is hoped that the release of this dataset will allow for such explorations to be taken by the research community.

Code availability

The code used for technical validation is available at our Github repository (<https://github.com/MuSAELab/MSPB>). Other details of hive management can be found at the dataset repository⁴² as well as the project webpage (<https://zhu00121.github.io/MSPB-webpage/>).

Received: 21 November 2023; Accepted: 29 July 2024;

Published online: 09 August 2024

References

1. CAAS, I. Good beekeeping practices for sustainable apiculture. *FAO: Food And Agriculture Organization Of The United Nations*. (2020).
2. Klein, A. *et al.* Importance of pollinators in changing landscapes for world crops. *Proceedings Of The Royal Society B: Biological Sciences* **274**, 303–313 (2007).
3. Aizen, M. & Harder, L. The global stock of domesticated honey bees is growing slower than agricultural demand for pollination. *Current Biology* **19**, 915–918 (2009).
4. Abdollahi, M., Giovenazzo, P. & Falk, T. Automated beehive acoustics monitoring: A comprehensive review of the literature and recommendations for future work. *Applied Sciences* **12**, 3920 (2022).
5. Aronstein, K. & Murray, K. Chalkbrood disease in honey bees. *Journal Of Invertebrate Pathology* **103**, S20–S29 (2010).

6. Gaubert, J., Giovenazzo, P. & Derome, N. Individual and social defenses in *Apis mellifera*: a playground to fight against synergistic stressor interactions. *Frontiers In Physiology* **14** (2023).
7. Brodschneider, R. *et al.* Others Multi-country loss rates of honey bee colonies during winter 2016/2017 from the COLOSS survey. *Journal Of Apicultural Research* **57**, 452–457 (2018).
8. Gray, A. *et al.* Others Loss rates of honey bee colonies during winter 2017/18 in 36 countries participating in the COLOSS survey, including effects of forage sources. *Journal Of Apicultural Research* **58**, 479–485 (2019).
9. Gray, A. *et al.* Others Honey bee colony winter loss rates for 35 countries participating in the COLOSS survey for winter 2018–2019, and the effects of a new queen on the risk of colony winter loss. *Journal Of Apicultural Research* **59**, 744–751 (2020).
10. Zhu, Y. *et al.* Early prediction of honeybee hive winter survivability using multi-modal sensor data. *2023 IEEE International Workshop On Metrology For Agriculture And Forestry (MetroAgriFor)*. pp. 657–662 (2023).
11. Zhang, T., Zmyslony, S., Nozdrenkov, S., Smith, M. & Hopkins, B. Semi-supervised audio representation learning for modeling beehive strengths. Preprint at <https://doi.org/10.48550/arXiv.2105.10536> (2021).
12. Terenzi, A., Ortolani, N., Nolasco, I., Benetos, E. & Cecchi, S. Comparison of feature extraction methods for sound-based classification of honey bee activity. *IEEE/ACM Transactions On Audio, Speech, And Language Processing* **30**, 112–122 (2021).
13. Cetin, U. The effects of temperature changes to bee losses. *Uludag Bee J* **4**, 171–174 (2004).
14. Southwick, E. Metabolic energy of intact honey bee colonies. *Comparative Biochemistry And Physiology Part A: Physiology* **71**, 277–281 (1982).
15. Seeley, T. *Honeybee Ecology: A Study of Adaptation in Social Life*. (Princeton University Press, 2014).
16. Human, H., Nicolson, S. & Diemann, V. *Apis mellifera scutellata*, regulate humidity in their nest? *Naturwissenschaften* **93**, 397–401 (2006).
17. Oertel, E. Relative humidity and temperature within the beehive. *J. Econ. Entomol* **42**, 528–531 (1949).
18. Ferrari, S., Silva, M., Guarino, M. & Berckmans, D. Monitoring of swarming sounds in bee hives for early detection of the swarming period. *Computers And Electronics In Agriculture* **64**, 72–77 (2008).
19. Abou-Shaara, H., Owayss, A., Ibrahim, Y. & Basuny, N. A review of impacts of temperature and relative humidity on various activities of honey bees. *Insectes Sociaux* **64**, 455–463 (2017).
20. Kirchner, W. Acoustical communication in honeybees. *Apidologie* **24**, 297–307 (1993).
21. Pastor, K. & Seeley, T. The brief piping signal of the honey bee: begging call or stop signal? *Ethology* **111**, 775–784 (2005).
22. Uthoff, C., Homsy, M. & Bergen, M. Acoustic and vibration monitoring of honeybee colonies for beekeeping-relevant aspects of presence of queen bee and swarming. *Computers And Electronics In Agriculture* **205**, 107589 (2023).
23. Nolasco, I. *et al.* Audio-based identification of beehive states. *ICASSP 2019-2019 IEEE International Conference On Acoustics, Speech And Signal Processing (ICASSP)*. 8256–8260 (2019).
24. Kim, J., Oh, J. & Heo, T. Acoustic scene classification and visualization of beehive sounds using machine learning algorithms and Grad-CAM. *Mathematical Problems In Engineering* **2021**, 1–13 (2021).
25. Zgank, A. Bee swarm activity acoustic classification for an IoT-based farm service. *Sensors* **20**, 21 (2019).
26. Zgank, A. IoT-based bee swarm activity acoustic classification using deep neural networks. *Sensors* **21**, 676 (2021).
27. Terenzi, A., Cecchi, S., Orcioni, S. & Piazza, F. Features extraction applied to the analysis of the sounds emitted by honey bees in a beehive. *2019 11th International Symposium On Image And Signal Processing And Analysis (ISPA)*. pp. 03–08 (2019).
28. Cecchi, S. *et al.* A preliminary study of sounds emitted by honey bees in a beehive. *Audio Engineering Society Convention 144*. (2018).
29. Jordan, M. & Mitchell, T. Machine learning: Trends, perspectives, and prospects. *Science* **349**, 255–260 (2015).
30. Kulhanek, K. *et al.* Others A national survey of managed honey bee 2015–2016 annual colony losses in the USA. *Journal Of Apicultural Research* **56**, 328–340 (2017).
31. Giovenazzo, P. & Dubreuil, P. Evaluation of spring organic treatments against *Varroa destructor* (Acari: Varroidae) in honey bee *Apis mellifera* (Hymenoptera: Apidae) colonies in eastern Canada. *Experimental And Applied Acarology* **55**, 65–76 (2011).
32. Delaplane, K., Van Der Steen, J. & Guzman-Novoa, E. Standard methods for estimating strength parameters of *Apis mellifera* colonies. *Journal Of Apicultural Research* **52**, 1–12 (2013).
33. Maucourt, S., Fortin, F., Robert, C. & Giovenazzo, P. Genetic parameters of honey bee colonies traits in a Canadian selection program. *Insects* **11**, 587 (2020).
34. Spivak, M. & Downey, D. Field assays for hygienic behavior in honey bees (Hymenoptera: Apidae). *Journal Of Economic Entomology* **91**, 64–70 (1998).
35. Spivak, M. & Reuter, G. Honey bee hygienic behavior. *American Bee Journal* **138**, 283–286 (1998).
36. Guzmán-Novoa, E., Prieto-Merlos, D., Uribe-Rubio, J. & Hunt, G. Relative reliability of four field assays to test defensive behaviour of honey bees (*Apis mellifera*). *Journal Of Apicultural Research* **42**, 42–46 (2003).
37. Guzman-Novoa, E. *Elemental genetics and breeding for the honeybee*. (University of Guelph, 2007).
38. Imdorf, A., Charrière, J., Kilchenmann, V., Bogdanov, S. & Fluri, P. Alternative strategy in central Europe for the control of *Varroa destructor* in honey bee colonies. *Apiacta* **38**, 258–278 (2003).
39. Diemann, V. *et al.* Others Standard methods for varroa research. *Journal Of Apicultural Research* **52**, 1–54 (2013).
40. De Jong, D., Roma, D. & Goncalves, L. A comparative analysis of shaking solutions for the detection of *Varroa jacobsoni* on adult honeybees. *Apidologie* **13**, 297–306 (1982).
41. Borba, R. *et al.* Phenomic analysis of the honey bee pathogen-web and its dynamics on colony productivity, health and social immunity behaviors. *Plos One* **17**, e0263273 (2022).
42. Zhu, Y. *et al.* MSPB: a longitudinal multi-sensor dataset with phenotypic trait measurements from honey bees. *Zenodo* <https://doi.org/10.5281/zenodo.11398835> (2024).
43. Moore, D. Honey bee circadian clocks: behavioral control from individual workers to whole-colony rhythms. *Journal Of Insect Physiology* **47**, 843–857 (2001).
44. Eban-Rothschild, A. & Bloch, G. Circadian rhythms and sleep in honey bees. *Honeybee Neurobiology And Behavior: A Tribute To Randolph Menzel*. pp. 31–45 (2011).
45. Kaiser, W. Busy bees need rest, too: behavioural and electromyographical sleep signs in honeybees. *Journal Of Comparative Physiology A* **163**, 565–584 (1988).
46. Jones, J., Myerscough, M., Graham, S. & Oldroyd, B. Honey bee nest thermoregulation: diversity promotes stability. *Science* **305**, 402–404 (2004).
47. Ferland, J. *et al.* Others Canadian Association of Professional Apiculturists Statement on Honey Bee Wintering Losses in Canada. (2022).
48. Abdollahi, M., Henry, E., Giovenazzo, P. & Falk, T. The Importance of Context Awareness in Acoustics-Based Automated Beehive Monitoring. *Applied Sciences* **13**, 195 (2022).
49. Virtanen, P. *et al.* Others SciPy 1. *Nature Methods* **17**, 261–272 (2020).
50. Scheiner, R. *et al.* Others Standard methods for behavioural studies of *Apis mellifera*. *Journal Of Apicultural Research* **52**, 1–58 (2013).
51. Kulyukin, V. Audio, image, video, and weather datasets for continuous electronic beehive monitoring. *Applied Sciences* **11**, 4632 (2021).
52. Masaquiza, D. *et al.* Hygienic behavior of *Apis mellifera* and its relationship with *Varroa destructor* infestation and honey production in the central highlands of Ecuador. *Insects* **12**, 966 (2021).
53. Nolasco, I. & Benetos, E. To bee or not to bee: Investigating machine learning approaches for beehive sound recognition. Preprint at <https://doi.org/10.48550/arXiv.1811.06016> (2018).

Acknowledgements

The authors acknowledge funding from NSERC via their Alliance program (ALLRP 548872-19), as well as Nectar Technologies Inc and the Centre de recherche en sciences animales de Deschambault for the support with data collection.

Author contributions

Y.Z. contributed to the initial drafting of the manuscript and conducted the experiments in the technical validation. M.A. contributed to data preparation and the initial drafting of the background section. S.M. contributed to phenotypic trait measurements collection, acquisition, and interpretation. N.C. contributed to sensor data acquisition and verified data records. H.G. contributed to the feedback to experimental results. P.G. and T.F. contributed in conceptualizing the study and study design. All authors contributed to the critical revision of the manuscript. All authors had full access to all the data in the study and took responsibility for the decision to submit this draft for publication.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024