Characterizing the suckling behavior by video and 3D-accelerometry in humpback whale calves on a breeding ground (#66429)

First submission

Guidance from your Editor

Please submit by **1 Nov 2021** for the benefit of the authors (and your $200 publishing discount) .

✓ **Structure and Criteria**
Please read the 'Structure and Criteria' page for general guidance.

! **Custom checks**
Make sure you include the custom checks shown below, in your review.

▦ **Raw data check**
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the materials page.

7 Figure file(s)
3 Table file(s)
4 Raw data file(s)

! Custom checks

Vertebrate animal usage checks

- ! Have you checked the authors ethical approval statement?
- ! Were the experiments necessary and ethical?
- ! Have you checked our animal research policies?

Field study

- ! Have you checked the authors field study permits?
- ! Are the field study permits appropriate?

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

🗋 You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- ✓ Clear, unambiguous, professional English language used throughout.
- ✓ Intro & background to show context. Literature well referenced & relevant.
- ✓ Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- ✓ Figures are relevant, high quality, well labelled & described.
- ✓ Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- ✓ Original primary research within Scope of the journal.
- ✓ Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- ✓ Rigorous investigation performed to a high technical & ethical standard.
- ✓ Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- ℹ Impact and novelty not assessed. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- ✓ All underlying data have been provided; they are robust, statistically sound, & controlled.
- ✓ Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Characterizing the suckling behavior by video and 3D-accelerometry in humpback whale calves on a breeding ground

Maevatiana N Ratsimbazafindranahaka Corresp., 1, 2, 3 , **Chloé Huetz** [2] , **Aristide Andrianarimisa** [3] , **Joy S Reidenberg** [4] , **Anjara Saloma** [1] , **Olivier Adam** [2, 5] , **Isabelle Charrier** [2]

[1] Association Cétamada, Barachois Sainte Marie, Madagascar

[2] Institut des Neurosciences Paris-Saclay, Université Paris-Saclay, CNRS, Saclay, France

[3] Département de Zoologie et Biodiversité Animale, Université d'Antananarivo, Antananarivo, Madagascar

[4] Center for Anatomy and Functional Morphology, Icahn School of Medicine at Mount Sinai, New York, United States

[5] Institut Jean Le Rond d'Alembert, Sorbonne Université, Paris, France

Corresponding Author: Maevatiana N Ratsimbazafindranahaka
Email address: maevatiana.ratsimbazafindranahaka@universite-paris-saclay.fr

Getting maternal milk through nursing is vital for all new-born mammals. Despite its importance, nursing has been poorly documented in humpback whales (Megaptera novaeangliae) as it is difficult to observe underwater without disturbing the whales, and usually impossible to observe from a ship. We attempted to observe nursing from the calf's perspective by placing CATS cam tags (video camera, hydrophone and auxiliary sensors: depth sensor, 3D accelerometer, 3D gyroscope) on three humpback calves in the Sainte Marie channel, Madagascar, Indian Ocean, during the breeding seasons. This method minimized any potential disturbance from human presence. A total of 10.52 hours of video recordings were collected, with the corresponding auxiliary data. Video recordings were manually analyzed and correlated with the auxiliary data, allowing us to extract different kinematic features including the depth rate (DR), speed, fluke stroke rate (FSR), overall body dynamic acceleration (ODBA), pitch, roll, and roll rates. We found that suckling events lasted 18.8±8.8 s (N = 34) on average and were performed mostly during dives, at 19±6.7 m depth. Suckling events represented 1.7% of the total observation time. During suckling, the calves were seen at a 30-45° pitch angle relative to the midline their mother's body and were always observed rolling either to the right or to the left. From our dataset, we found that suckling behavior was characterized by a high absolute mean roll and low mean speed. Kinematic features were used for supervised machine learning in order to subsequently detect the suckling activities automatically. These findings open new opportunities for further investigation of suckling behavior in humpback whales and in baleen whales in general.

Characterizing the suckling behavior by video and 3D-accelerometry in Humpback whale calves on a breeding ground

Maevatiana Nokoloina Ratsimbazafindranahaka[1,2,3], Chloé Huetz[1], Aristide Andrianarimisa[2], Joy S. Reidenberg[4], Anjara Saloma[3], Olivier Adam[1,5], Isabelle Charrier[1]

[1] Institut des Neurosciences Paris-Saclay, Université Paris-Saclay, CNRS, Saclay, France
[2] Département de Zoologie et Biodiversité Animale, Université d'Antananarivo, Antananarivo, Madagascar
[3] Association Cétamada, Barachois Sainte Marie, Madagascar
[4] Center for Anatomy and Functional Morphology, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[5] Institut Jean Le Rond d'Alembert, Sorbonne Université, Paris, France

Corresponding Author:
Maevatiana Nokoloina Ratsimbazafindranahaka[1,2,3]
Institut des Neurosciences Paris-Saclay (NeuroPSI), UMR 9197 CNRS
Campus CEA Saclay, 151 route de la Rotonde, 91400 Saclay, FRANCE
Email address: maevatiana.ratsimbazafindranahaka@universite-paris-saclay.fr

Abstract

Getting maternal milk through nursing is vital for all newborn mammals. Despite its importance, nursing has been poorly documented in humpback whales (Megaptera novaeangliae) as it is difficult to observe underwater without disturbing the whales, and usually impossible to observe from a ship. We attempted to observe nursing from the calf's perspective by placing CATS cam tags (video camera, hydrophone, and auxiliary sensors: depth sensor, 3D accelerometer, 3D gyroscope) on three humpback calves in the Sainte Marie channel, Madagascar, Indian Ocean, during the breeding seasons. This method minimized any potential disturbance from human presence. A total of 10.52 hours of video recordings were collected, with the corresponding auxiliary data. Video recordings were manually analyzed and correlated with the auxiliary data, allowing us to extract different kinematic features including the depth rate (DR), speed, fluke stroke rate (FSR), overall body dynamic acceleration (ODBA), pitch, roll, and roll rates. We found that suckling events lasted 18.8±8.8 s (N = 34) on average and were performed mostly during dives, at 19±6.7 m depth. Suckling events represented 1.7% of the total observation time. During suckling, the calves were seen at a 30-45° pitch angle relative to the midline of their mother's body and were always observed rolling either to the right or to the left. From our dataset, we found that suckling behavior was characterized by a high absolute mean roll and low mean speed. Kinematic features were used for supervised machine learning in order to

subsequently detect the suckling activities automatically. These findings open new opportunities for further investigation of suckling behavior in humpback whales and in baleen whales in general.

Introduction

All female mammals feed their offspring with maternal milk to ensure their progeny's development and survival during their early dependent life stages (Balshine, 2012). The maternal behavior associated to the transfer of milk to the young is referred as 'nursing'. From the offspring's perspective, this behavior can be referred to as 'suckling', i.e., to obtain milk from the mammary gland (breast) through the mammary papilla (nipple or teat) (Hall, Hudson & Brake, 1988). This can involve an active role of mouth movements from the young to stimulate nipple erection, compress the breast's lactiferous sinuses, and generate negative pressure (suction) along with a pulling motion on the teat in order to withdraw the milk, or a passive role in which milk is squirted into the open mouth, with no stimulating movements of the offspring's mouth.

Nursing behavior is an integral part of mammalian reproductive behavior (Gittleman & Thompson, 1988). The duration of lactation (the period of time during which milk can be produced; Hall, Hudson & Brake, 1988), and the frequency of nursing, vary greatly among mammals and are related to the developmental characteristics of the young and their environmental constraints (Oftedal, Boness & Tedman, 1987; Oftedal, 1993).

Mammals have breasts in a variety of positions and numbers, depending upon taxonomic group. Breasts are always paired, and may occur anywhere along the milk ridge – a line extending from near the axilla (armpit) to the inguinal region (groin or inner thigh). Many mammals have two rows of paired breasts (e.g., four pairs in cats, four-five pairs in dogs, six-seven pairs in pigs), while some only have them in the chest region (e.g., primates, elephants, manatees) or the inguinal region (e.g., horses, giraffes, hippopotamuses). Cetaceans (whales, including dolphins and porpoises) are in the order Cetartiodactyla, and therefore retain hind breasts as is typical of all artiodactyls. While some artiodactyls have two pairs of hind breasts (e.g., cattle), whales only have one pair (similar to goats and sheep).

Whale breasts are positioned ventro-laterally on either side of the midline genital slit (opening of the vestibule vagina). The location of each breast is marked by a short mammary slit in the overlying blubber through which the nipple can be extruded for nursing. The mammary glandular tissue is located deep to the overlying blubber layer, thus providing a hydrodynamically streamlined outline even when the whale's breasts are engorged with milk during lactation. Lactation initially occurs in the winter, when calves are born. As most large whales migrate, lactation that begins on the calving area must continue during the migration to the feeding area, and then again on the return trip.

Humpback whales (*Megaptera novaeangliae*), the subject of this study, are migratory cetaceans. They move between high-latitude (closer to the poles) feeding grounds and low-latitude (closer to the equator) breeding/calving grounds (Clapham, 2018). Similar to other baleen whales, their nursing strategy is constrained by the prolonged fasting of the mother during lactating period, related to the migratory pattern (Oftedal, 1993). A young humpback whale (calf) feeds exclusively on maternal milk during its first six months of life (Clapham, 2018). The energy intake from the milk is crucial in ensuring rapid growth, and therefore allowing the calf to start its first long migration to the feeding area alongside its mother. As nursing must occur throughout the migration path, mother-calf pairs make slow progress in part due to presumed slower swimming or frequent pauses that allow the calf to suckle. After its first six months of life, the calf starts to feed independently in the feeding area but still continues to get nutrition from milk until complete weaning at 10 to 12 months old (Clapham, 2018). This coincides with the mother-calf pair returning to the calving area.

Despite its importance, nursing (and suckling) behavior has been poorly documented in humpback whales because it is difficult to observe with accuracy and certainty. Indeed, it can occur at varying depths from the sub-surface (< 5 m) to deep dives (up to 64 m), as suggested by camera-observed nursing events in feeding areas (Tackaberry et al., 2020). The earlier descriptions of nursing behavior in humpback whales were mostly based on limited surface and sub-surface observations of stationary or slow moving whales, and relied mainly on the positioning of the calf (Glockner & Venus, 1983; Glockner-Ferrari & Ferrari, 1985; Clapham & Mayo, 1987; Morete et al., 2003; Videsen et al., 2017). Clapham and Mayo (1987) for example described nursing as an event between two calf's successive breaths on either sides of the mother's tail stock (caudal peduncle), when the mother is stationary at the surface, and during which the calf is beneath the mother. Videsen et al. (2017) defined nursing events by associating them with the occurrence of 'peduncle dives', defined as repeated dives during which the calf remained submerged beneath the mother's caudal peduncle.

Nevertheless, from these different studies, various nursing modalities have been advanced. While nursing, the mother is either stationary or swimming very slowly at a depth around 10-15 m (Glockner-Ferrari & Ferrari, 1985), or occasionally at the surface with her tail in the air (Morete et al., 2003). The calf is typically positioned vertically (head up, tail down) beneath the mother (Glockner & Venus, 1983). Videsen et al. (2017) suggested that to initiate suckling, the calf uses tactile stimuli on the mother rather than vocalizations. Tactile stimulation likely results in an autonomic reflex causing extrusion of the nipple. The calf can then grasp it with its tongue. The whale calf has an unusual tongue that is lined with lingual papillae (fringes) that help the tongue grasp the nipple, and may even seal the tongue into a tube that facilitates transferring milk into the mouth.

Underwater video is the best method for studying the nursing and suckling behaviors. Various solutions are possible. One option is to use a video camera from a surface platform (boat or kayak), but this has two major drawbacks: 1) it is necessary to get very close to the whales and potentially interfere with their behavior (harassment); 2) visual contact is lost when the whales dive deeply, even when the water is very clear. Another option is to use an underwater drone (Remotely Operated underwater Vehicle – ROV or Autonomous Underwater Vehicle – AUV, Butcher et al., 2021). This method has potential but presently vehicles are not yet agile enough to track whales: too slow, difficult to manoeuver, and have a limited cable length (in the case of ROV). Alternatively, underwater video can be obtained during close approaches by divers (breath-holding or scuba divers). However, this is only possible when the whales are stationary or swimming very slowly, as it would be impossible for a diver to keep pace with cruising whales. Additionally, the presence of divers in close proximity may disturb the whales, and habituation is not a desirable tactic. Scott Portelli (2021, pers. Comm.), Zoidis and Lomac-MacNair (2017) successfully video documented humpback whale suckling behavior thanks to breath-holding divers. In Zoidis and Lomac-MacNair (2017), the recording of milk clouds in the water column allowed the confirmation of the occurrence of suckling. However, only a few events (*N* = 5 from four mother-calf pairs) could be recorded.

The development of animal-borne multi-sensor tags opened the way for remote and accurate recording of whale behavior (Cade et al., 2016), as these may be equipped with camera, accelerometers, gyroscopes, magnetometers, pressure (depth) sensor, temperature sensor, light sensor, and hydrophone, such as Customized Animal Tracking Solutions (CATS) cam tags. These models are particularly well suited as, in addition to kinematic sensors, they also allow the recording of video directly from the calf's perspective, and thus allow confident confirmation of suckling events. Such devices were used by Tackabery et al. (2020) to accurately describe the suckling behavior of humpback whale calves on feeding areas. One disadvantage of using an animal-borne tag is the potential stress generated during the tag attachment phase. The deployment must be done by experienced operators following a strict protocol (slow speed approach, brief deployment time, etc.) to minimizes the disturbance (Stimpert et al., 2012; Saloma, 2018).

In this study, we characterized the suckling behavior of humpback whale calves on their breeding ground (< 3 months old calves) using CATS cam tags. We aimed to describe in detail: 1) how and at what depth the suckling is performed, and its duration and frequency; 2) the behavioral signatures of suckling events using accelerometry and depth data; and 3) whether the different kinematic features extracted from only two sensors (3-axis accelerometer and depth sensor) among all the available sensors in the CATS cam tags were sufficient for a supervised machine learning algorithm to detect suckling events. Supervised machine learning is a type of machine learning in which an algorithm learns from labelled datasets to classify of unforeseen data. Such classification technique would open the opportunity to detect suckling events from data collected by tags without video camera (such as Acousonde or Dtags) or when visual observations are not good enough (during night time or in water with poor visibility).

Materials & Methods

Study area

The study occurred in the Indian Ocean along coast of Madagascar, specifically in the Sainte Marie channel. The channel is located between the Sainte Marie Island (between latitudes 17° 19' and 16° 42' South, and longitudes 49° 48' and 50° 01' East) and the east coast of Madagascar's mainland. The channel is approximately 60 km long and 7-30 km wide. It is relatively shallow: its average depth is about 35 m and the maximal depth is 60 m (Trudelle et al., 2016). The tag deployments were conducted as part of ongoing study on humpback whale mother-calf interactions, during the calving seasons of 2018 and 2019 (between August and September).

Tag specifications

We used CATS cam tags to investigate humpback whale suckling behavior. CATS cams are small and lightweight (~500 g), non-invasive animal-borne multi-sensors tags, attached via suction cups. They contain six auxiliary sensors (3-axis accelerometer, 3-axis gyroscope, 3-axis magnetometer, pressure (depth) sensor, temperature sensor, and light sensor), a hydrophone, and an HD video camera (100° Field of view). A VHF transmitter (ATS F1835B) attached on the tag allows tracking for tag retrieval. The sampling frequency was set at 10 Hz for the magnetometer, gyroscope, depth, and temperature sensors. Accelerometer sampling rate was set at 400 Hz in 2018 and at 800 Hz in 2019. The camera recorded videos with a 1280 x 720-pixel resolution at 30 frames per second in 2018, and with a 1920 x 1072-pixel resolution at 30 frames per second in 2019. The hydrophone recorded sound at a 48 kHz sampling rate in 2018 and at a 24 kHz sampling rate in 2019 (16-bit resolution).

Tagging procedures

CATS cam tags were deployed on calves accompanied by their mother, from a 6.40 m rigid motor boat, using a 5-m carbon-fiber pole. Deployment were performed by researchers experienced in successfully approaching mother-calf pairs with minimal disturbance. The tags were placed on the back, near the dorsal fin of the animal. Calves were tagged using one of the two approaches described in Stimpert et al. (2012) and Saloma (2018) in order to minimize disturbance to the mother-calf pair. Tagging efforts were terminated if the pair displayed avoidance behavior, or if the calf was not successfully tagged within 30 min. All mother-calf pairs were photo-identified to avoid double-sampling within the calving season. We attributed a relative age to each tagged calf depending on the angle of furl of the dorsal fin, following established methods (newborn versus non-newborn, Cartwright & Sullivan, 2009; Faria et al., 2013; Saloma, 2018).

Tagged animals were not followed immediately after tag deployment to avoid any further disturbance of their behavior. After tag deployment or an aborted attempt, the boat slowly moved away in the opposite direction of the mother-calf pair. The tag was retrieved after few hours or the following day, when it detached itself from the animal (usually as a consequence of rubbing against the mother, jumps, etc.). The VHF tag emitted a continuous signal, facilitating retrieval. All methods and approaches were carried out in accordance with relevant guidelines and regulations in force in Madagascar, and were approved by the Ministry of Fisheries Resources, Madagascar, under the national research and collect permits #28/18-MRHP/SG/DGRHP and #36/19-MAEP/SG/DGPA.

Sensor data analysis

Data from all sensors were downloaded as CSV files and imported into MATLAB (Mathworks) using dedicated scripts (CATS Matlab toolkit, https://github.com/wgough/CATS-Methods-Materials). Raw accelerometer data were downsampled to obtain a common sampling rate of 10·Hz across all sensors. Sensors reading were rotated to match the calf's orientation frame using established methods, and animal pitch and roll (in degree) were calculated (Johnson & Tyack, 2003; Cade et al., 2016; Tackaberry et al., 2020). Animal forward speed (speed hereafter, in meter/second), sampled at 10 Hz, was determined using the tag jiggle recorded in the original high frequency accelerometer data (Cade et al., 2017). Accelerometer data from 2019 were first downsampled to 400 Hz prior to speed calculation for consistency. In addition, we calculated the overall dynamic body acceleration (ODBA, in meter/second^2) as in Wilson et al. (2006), the roll rate (in degree/second), and the depth rate (in meter/second), using custom scripts and the Animal Tag toolbox (http://www.animaltags.org). Furthermore, we determined the fluke stroke pattern using the band-pass filtered pitch (0.2–1 Hz band-pass filter) as in Simon et al. (2012). Stroking was identified when the signal passed from below −3° to above +3° or vice versa within 6.5 s. The used thresholds were determined by visual inspection of the accelerometry plots (Iwata et al., 2021). From the identified fluke strokes, the fluke stroke rate (FSR, in Hertz) was calculated on the basis of half-strokes (López et al., 2015).

Depth data (in meters) helped determine various diving or surface activity phases. We defined diving as any submergence to a depth of >10 m (Stimpert et al., 2012; Saloma, 2018). Dives were further divided into three phases: descent, bottom, and ascent phase. As individual dives could include stops at various depths, the bottom phase was defined as the segment at > 85% of the maximal dive depth for a dive (Stimpert et al., 2012). The descent phase was defined as the segment starting at the surface that immediately preceded the bottom phase. Inversely, the ascent phase was defined as the segment that follows directly the bottom phase and ends at the surface.

Video data analysis and suckling data extraction

We identified and labelled suckling events from the video files using the Behavioral Observation Research Interactive Software (BORIS; Friard & Gamba, 2016). The corresponding depth data was displayed concurrent with the video. We defined a suckling event as a period during which the tip of the calf's snout continuously touched (> 2 s) the mammary slit of its mother and a milk cloud (see video clip in https://www.youtube.com/watch?v=UcyCgiCieFk), even in low density, was observed in the water during the event or upon release of contact (Tackaberry et al., 2020). For each suckling event, we calculated its duration and extracted its corresponding activity phase (descent, bottom, ascent, or surface), together with the average depth, depth rate, speed, FSR, ODBA pitch, roll, and roll rate. In addition to suckling events identification, we also counted the number of suckling dives (dives during which one or several confirmed suckling events were recorded) and the number of non-suckling dives during which the calf is observed staying at least 5 s under the mother, in close proximity (mother visible above the calf) and without any suckling event.

Comparison of suckling with adjacent non-suckling segments

A reference data set for comparison is needed to contrast and identify the behavioral signatures unique to suckling events. To generate such comparison data, we divided non-suckling periods into non-overlapping segments of 20 s (duration comparable to suckling events) and selected all segments that immediately preceded or followed a suckling event. These selected non-suckling segments, referred as "adjacent non-suckling segments", were analyzed for the same characteristics as suckling events (see above). To assess if there was a significant difference between suckling and non-suckling segments, we used a linear mixed-effects models (estimated using REML – Restricted Maximum Likelihood) that included the suckling status (suckling versus adjacent non-suckling) and the activity phase (descent, bottom, ascent or surface phase) as fixed effects, and individuals as random effect (Tackaberry et al., 2020). In the models, we considered the following response variables: mean depth rate, mean speed, mean FSR, mean ODBA, mean pitch, absolute mean roll (absolute value have been used to emphasize any roll deviation from zero), and mean roll rate. The model's reference levels (intercept) corresponded to adjacent bottom non-suckling (non-suckling segment occurring at the bottom phase of dive). All statistical analyses were completed using the R software. Statistical significance level was set to $\alpha = 0.05$.

Testing automatic identification of suckling behavior using supervised machine learning

Data preparation

Contrary to the previous analysis for the identification of the behavioral signatures unique to suckling events (in which only suckling events and non-suckling periods adjacent to suckling were included), the dataset we used in this part of the study included the total periods. As the goal was to identify automatically suckling and non-suckling in non-labelled data, it was crucial to include all data. Thus from the complete dataset, we computed 43 features using 2-sec-sliding windows (blocks), without overlap. These features included mean, minimum, maximum, variance, skewness and kurtosis for depth, depth rate, speed, ODBA, pitch, roll (absolute) and roll rate, and mean for FSR. A class was assigned to each block, following the initial video labelling: (1) 'suckling', the targeted behavior, or (2) 'non-suckling'.

Similar to Jeantet et al. (2020), we segmented the data prior to machine learning implementation. We excluded blocks too close to the surface (< 2 m depth) or of high speed activities (> 2 m s^{-1}). This helped in partially removing noise from the data and also in reducing class imbalance. Suckling events occur mostly at depth (Tackaberry et al., 2020). Given the size of the mothers (> 11 m in length and > 2 m diameter at the umbilical section level in adult whales, as determined by estimates in the field), the generally observed nursing configuration (calf generally below the mother), and speed (Zoidis & Lomac-MacNair, 2017; Tackaberry et al., 2020), we were confident that no nursing events would be recorded at < 2 m depth or at > 2 m s^{-1} speed. We checked the validity of these threshold assumptions in our data to ensure that indeed no suckling events were partially or entirely removed. Validating these parameters ensured that this segmentation could be automatically performed on unknown data, with a very low probability to remove suckling events.

Identification of the adequate classifier

In order to define an appropriate classification model (machine learning model) for associating the suckling status of the calf with the corresponding patterns of different kinematic features, we trained three types of supervised machine learning algorithms — (i) K-Nearest Neighbors (KNN), (ii) Decision tree, and (iii) Ensemble classifiers — using the MATLAB toolbox *Statistics and Machine Learning*. We repeatedly performed a 60:40 holdout splits on our data (i.e., 60% of the data as training set and 40% of the data as testing set) while maintaining class ratios ($N = 30$ runs): the training set was used for training and the classifier's efficiency was evaluated on the remaining unseen data. Preliminary model selection and hyperparameter tuning for each classifier type was performed using a Bayesian Optimization approach (with 5-fold cross-validation). We followed the workflow described in the associated documentation (https://fr.mathworks.com/help/stats/bayesian-optimization-workflow). With the *Statistics and Machine Learning* toolbox's Bayesian Optimization algorithm, we optimized the different types of supervised machine learning algorithms across a selection of classification models and hyperparameter values in order to pre-select the best classification models that were suitable for our dataset. For Ensemble classifiers, the selection of classification models included Bootstrap Aggregation (Bagging), Random Forest (Bag), Random Subspace (Subspace), Adaptive Boosting for Binary Classification (AdaBoostM1), Adaptive Boosting for Multiclass Classification (AdaBoostM2), Gentle Adaptive Boosting (GentleBoost), Adaptive Logistic Regression (LogitBoost), Linear Programming Boosting (LPBoost), Least-Squares Boosting (LSBoost), Robust Boosting (RobustBoost), Random Undersampling Boosting (RUSBoost), and Totally Corrective Boosting (TotalBoost). For KNN classifiers, it included Cityblock, Chebychev, Correlation, Cosine, Euclidean, Hamming, Jaccard, Mahalanobis, Minkowski, Seuclidean, and Spearman metrics-based KNN. For Decision trees, it included classifiers that are based on a Gini's diversity index (gdi), twoing rule and deviance splitting criterion. These different models are detailed in the aforementioned documentation. The number of runs we chose ($N = 90$ runs, 30 for each type of classifier) and the number of iterations for the Bayesian optimization ($N = 200$) were chosen to optimize results stability and processing speed.

For each trial and each class, we calculated five evaluation metrics for making a final decision on the classification model to be retained: Sensitivity (true positive rate, hit rate or recall), Precision (positive predictive value), False Positive Rate (FPR), *F*-score and Global accuracy. These metrics were calculated as follows and used to select the best and most adapted classification model (definitions as in Jeantet et al., 2020):

The Sensitivity measures the ability to detect one behavior among other behaviors:

$$Sensitivity = \frac{TP}{TP + FN} \qquad (1)$$

The Precision measures the ability to correctly identify a behavior:

$$Precision = \frac{TP}{TP + FP} \qquad (2)$$

The False Positive Rate measures the rate of wrongly considering other behaviors as the behavior of interest. It is related to the Specificity, which is the ability to avoid wrongly considering other behaviors as the behavior of interest:

$$FPR = 1 - Specificity \quad \text{with} \quad Specificity = \frac{TN}{TN + FP} \qquad (3)$$

The *F*-score measures the accuracy in classifying a behavior. It is the harmonic mean of precision and sensitivity:

$$F - score = \frac{2TP}{2TP + FP + FN} \qquad (4)$$

The Global accuracy measures the ability to correctly identify all behaviors as a whole:

$$Global\ accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (5)$$

These formulas use the following abbreviations: TP (true positive), TN (true negative), FN (false negative), and FP (false positive).

To assess the potential influence of training set size reduction on the performance of the elected model, we tested whether changing the training set size changed the performance. We ran a series of trainings with different holdout splitting of the data (training-testing): 60:40 (as in the original model selection process), and 50:50, 40:60, 30:70, 20:80, and 10:90. The process was repeated 30 times for each splitting. A general workflow of the whole machine learning process we followed is presented in Fig. 1.

Generalization across individuals

In order to assess whether the model is able to generalize across different individual whale calves (i.e., whether we can use data obtained from other individuals to identify the suckling behavior of one individual), we also tested a leave-one-out split, i.e., we trained the model with the highest performance on *N*-1 calves' data and then tested its performance on the unseen remaining calf data. We followed the same machine learning workflow outlined in Fig. 1.

Aside from possible influence of training set size (e.g., the recorded data from the individuals used for training is relatively small), lack of generalization ability across individuals may result from a dataset with high inter-individual variability of suckling behavior characteristics. To address this potential issue, we tested whether excluding features that contribute most to inter-individual difference (if any) in the suckling blocks would improve the performance of the supervised machine learning. To identify the features that contribute most to inter-individual difference, we trained a random forest algorithm to classify individuals based only on the suckling blocks (R software package randomForest; Liaw & Wiener, 2002). Features that contributed the most to the differentiation by individual were identified as those that had a high

Gini index in the random forest. We re-ran the leave-one-out design, as described above, but excluding the features identified by using the Gini index. We confirmed the reduction of the inter-individual differences of the suckling blocks by examining the error rate (Out-of-the-bag error rate) of the random forest algorithm (trained to classify individuals). A reduction of inter-individual difference would result in the increase of the error rate. A low error rate indicated that the inter-individual difference is evident enough and thus the random forest algorithm is able to make very good classifications.

Results

Tag deployments

Four calves were tagged with CATS cam tag in the Sainte Marie channel: one during the calving season of 2018 (identified as Calf1) and three during the calving season of 2019 (identified as Calf2, Calf3, and Calf4). All of them had unfurled dorsal fins, indicating that they were not newborn calves (Cartwright & Sullivan, 2009). The deployment on Calf4 was not suitable for our study as the tag (thus the camera) was pointing toward the side of the calf (i.e., away from the mouth), and was thus excluded from our analyses. Deployment information including date and time is indicated in Table 1. The three deployments provided a total of 10.52 hours of usable video and auxiliary data.

General description of suckling

In all our video recordings, each time the calf's snout continuously touched (> 2 s) the mammary slit, a milk cloud could be observed during or after the contact. These events were all defined as suckling. In total, we identified 34 suckling events (Fig. 2). Detailed summary statistics per individual are presented in Table 1. Throughout this study, mean values are presented with their corresponding standard deviation (SD) in the format mean±SD. Suckling events lasted 18.8±8.8 s on average. Most of them constituted a sequence of 2-6 successive events spaced less than a minute apart (*N* = 30, dispersed in 10 series). They represented 1.7% of the total observation time.

The majority of suckling events occurred while diving, during the descent and the bottom phases (*N* = 9 for descent suckling and *N* = 20 for bottom suckling, Table 1), at a mean depth of 19±6.7 m (min = 8.9 m, max = 32.6 m). Only five suckling events occurred at the surface, at a mean depth of 6.2±2.7 m (min = 2 m, max = 8.4 m). No suckling event occurred during ascent phases.

While calves were estimated to be at 30-45° pitch angle relative to the midline of their mother's body during suckling in the videos, the recorded pitch (pitch angle relative to the horizontal) was fairly low on average (< 15°). During all of the observed events, the calf was always leaning to one side (to the right or to the left). The leanings were clearly recorded in the roll data as a sustained deviation from zero during suckling (Fig. 2). All series of suckling events were characterized by a distinct alternation between suckling with right leaning and suckling with left leaning, except one, in which the first two successive events were both performed on the same leaning side (Calf3, Fig. 2). In eleven suckling events (one for Calf1, four for Calf2 and six for Calf3), we were able to visually confirm that during a suckling event, the calf's mouth covered only one mammary teat. In three of these events, the tongue was clearly visible and displayed a rhythmic movement (Fig. 3, see also example footage in https://www.youtube.com/watch?v=nyqhb9BemYI). For these events with confirmed covered teat, leaning to the right corresponded to a mouth covering the right teat (*N* = 3) and leaning to the left corresponded to a mouth covering the left teat (*N* = 8, Fig. 2).

Relation between calf's positioning and suckling during dive

Compared to non-suckling dives during which calves remained at least once beneath the mother for 5 s (at least), suckling dives were less common (Table 2). In other words, the positioning (beneath the mother) and the suckling behavior were not necessarily related.

Behavioral signatures of suckling

The characteristics of suckling events and adjacent non-suckling segments are summarized in Fig. 4. The results of the linear mixed-effect models are presented in Table 2. Within our model, for depth rate, the effect of suckling and the interaction between surface and suckling were all statistically non-significant and close to zero (Suckling: $\beta = 0.08$, $P = 0.291$, Suckling*Surface: $\beta = -0.08$, $P = 0.438$, Table 2). In other words, when suckling at the bottom of dive or at surface, calves exhibited a low depth rate that is comparable to adjacent bottom non-suckling segments.

Unsurprisingly, the effect of descent was statistically significant ($P < 0.001$), substantial, and positive ($\beta = 0.49$, Table 2). On the other hand, the interaction effect of suckling on descent was statistically significant ($P = 0.001$) and moderate, although negative ($\beta = -0.33$, Table 2). These indicated that the depth rate was high during the dive descent phase when the calf was not suckling, which is expected, but was significantly lower when suckling.

With respect to speed, a moderate, negative, and statistically significant effect of suckling was observed ($\beta = -0.33$, $P < 0.001$, Table 2). Overall the speed was lower during suckling events. Although the interaction effect of suckling on descent was statistically significant and positive ($\beta = 0.22$, $P = 0.043$, Table 2), it was lower than the aforementioned main effect of suckling on speed.

Suckling did not have any significant effect on FSR (for all factors: low β and $P > 0.05$, Table 2). It also did not have any significant effect on ODBA (for all factors: low β and $P > 0.05$, Table 2). In other words, when suckling, the calves displayed a FSR and an ODBA that were comparable to adjacent bottom non-suckling during descent, bottom, and surface suckling.

For pitch, the effect of suckling and the interaction effect of surface on suckling were weak and statistically non-significant (Suckling: $\beta = 4.24$, $P = 0.36$, Suckling*Surface: $\beta = -0.85$, $P = 0.836$, Table 2). The pitch exhibited by calves during both bottom and surface suckling were close to the pitch displayed during adjacent bottom non-suckling (segment during which calves are mostly and naturally in a horizontal posture, i.e., pitch close to zero). Unsurprisingly, within the studied dive phases, only descent had a statistically significant ($P = 0.002$) and substantial effect on the pitch ($\beta = -16.1$, Table 2). The interaction effect of suckling on descent, although statistically non-significant ($P = 0.081$), was positive and non-negligible ($\beta = 10.91$). During descent phase of dive, calves also had a posture close to horizontal when suckling.

Over all the considered characteristics, the specificity of calves' absolute roll with respect to suckling was the most obvious. The effect of suckling on absolute roll was statistically significant ($P < 0.001$) and very strong ($\beta = 44.42$, Table 2). All other interaction effects were statistically non-significant and weak, compared to the effect of suckling on absolute roll (Suckling*Descent: $\beta = -3.34$, $P = 0.678$, Suckling*Surface: $\beta = -12.85$, $P = 0.127$, Table 2). The recorded absolute roll is particularly high during suckling, regardless of the activity phase. Finally, suckling did not have any significant effect on roll rate (for all factors: low β and $P > 0.05$, Table 2). The calves displayed a roll rate that was comparable to adjacent bottom non-suckling during descent, bottom and surface suckling.

Best classifier for automatic identification of suckling

A total of 6930 behavioral blocks of 2 s duration (4158 for training and 2772 for testing, 60:40 holdout splitting) were obtained from the datasets after removing blocks too close to the surface and those of high speed activities (see Segmentation in the Method section). No suckling blocks were affected by our threshold for removing blocks too close to the surface and those of high speed activities. The class "suckling" represented 4.2% of the data and the class "non-suckling" represented 95.8%. In the preliminary model selection, three types of supervised machine learning algorithms were evaluated on these data: Ensemble classifiers, KNN classifiers and Decision trees. A Bayesian optimization algorithm was run on each classifier type (Ensembles, KNN and Decision trees). For each type of algorithms, an optimization run pre-selected one model as the best one by seeking to minimize the classification error. Fig. 5 presents all 7 pre-selected models and how often each model was selected by the Optimization procedure over the 90 runs (30 for each type of classifier).

All pre-selected classification models had a global accuracy > 96.39% on unseen data (test set). They all identified suckling events with a low False Positive Rate (FPR < 1%), except for GentleBoost and decision tree models (FPR > 1% but not exceeding 2%). The likelihood of misclassifying an event as suckling when it should have been classified as non-suckling is thus extremely low for most of the pre-selected models. As shown in Fig. 5, the GentleBoost model and all decision trees lay on the bottom left quadrant of the figure, indicating that they had a lower Precision and Sensitivity with regard to suckling, in contrast to AdaBoostM1, Bag and all KNN models. The AdaBoostM1, an Ensemble classifier, was the model presenting the highest Sensitivity, Precision and *F*-score (red circle symbol in Fig. 5A) and was therefore the best model for correctly identifying suckling events. Examples of classification resulting from the AdaBoostM1 are presented in Fig. 5B and Fig. 5C. In these examples, the AdaboostM1 model only misclassified non-suckling blocks as suckling in very rare cases (False positives, 10 in Fig. 5B and null in Fig. 5C – no red cross). However, it misclassified suckling blocks as non-suckling slightly more often (False negatives, 34 in Fig. 5B and 3 in Fig. 5C – blue crosses). Nevertheless, it detected most of the real suckling blocks (True positives, 81 in Fig. 5B and 16 in Fig. 5C – blue circles).

As shown in Fig. 5D, while there was typical decrease in the performance of the elected model (AdaBoostM1) when we reduced the training set size, the model still detected substantial amount of suckling blocks and had a relatively good precision even when using only 10% of the data as training set (mean at 10:90 split for suckling: sensitivity = 0.50±0.08, precision = 0.80±0.09, *F*-score = 0.61±0.06, $N = 30$). However, the results seemed more variable (high SD) when test set was small (Fig. 5D).

Generalization of the supervised machine learning across individuals

We evaluated the generalization ability of the AdaBoostM1 model, the best identified classifier in the context of suckling identification, using a leave-one-out design. The goal was to assess whether we can use data from other individuals to detect the suckling behavior of a new individual. Given that we had three individuals, there were only three possible splitting combinations: (1) blocks from Calf2 and Calf3 as training set (5018 blocks, 4.8% suckling and 95.2% non-suckling) and blocks from Calf1 as testing set (1912 blocks, 2.6% suckling and 97.4% non-suckling), (2) blocks from Calf1 and Calf3 as training set (6496 blocks, 3.7% suckling and 96.3% non-suckling) and blocks from Calf2 as testing set (434 blocks, 10.6% suckling and 89.4% non-suckling), and (3) blocks from Calf1 and Calf2 as training set (2346 blocks, 4% suckling and 96% non-suckling) and blocks from Calf3 as testing set (4584 blocks, 4.2% suckling and 95.8% non-suckling).

Results of the leave-one-out design are presented on Fig. 6 (open symbols). The first and second combination (open square and open diamond symbols) were more located at the bottom right part of the Sensitivity versus Precision plot. This indicated that the quantity of the detected suckling blocks was low (low Sensitivity) when using these two combination but the quality was still good (relatively high Precision). The third combination (open triangle symbol) was more located at the bottom left of the plot. This indicated that both the quantity and quality of the classification were bad (low Sensitivity and low Precision). Globally, it can be said that while some good quality classification can be obtained, it still depended on the combination used.

As inter-individual variation of the suckling blocks' kinematic features may have influenced the generalization ability of the AdaBoostM1, we assessed its performance when the features that contributed the most to inter-individual difference were excluded. We performed random forest classification of the individuals to check the existence of inter-individual differences on suckling blocks, analyzed the features' contribution to the differentiation, and identified a set of features to be excluded in order to reduce the inter-individual difference. The random forest algorithm indicated that there was effectively a strong inter-individual difference. Indeed, the classification error rate, indicated by the Out-of-the-bag error rate, was very low (Out-of-the-bag error rate = 3.79%). As shown in the Fig. 7A, the random forest algorithm was able to correctly predict to which individual each suckling blocks belongs to in most of the cases (correspondence between true class and predicted class). The Fig. 7B shows the Gini index of each feature. A high Gini index indicated that the feature played great role in the classification (in other words, in the individual differentiation). In the plot, we noticed a sharp drop in importance after the 9th feature. We thus chose to exclude the first 9 features (Gini index > 4): the mean depth, minimum speed, minimum depth, mean pitch, maximum pitch, minimum pitch, maximum depth, mean speed, and the maximum speed. The Fig. 7C shows the new resulting classification of the suckling blocks by individuals, when these features were not included. The classification error rate increased notably (Out-of-the-bag error rate = 23.45 %). The random forest algorithm confused the individual attribution of the suckling blocks more often (more non-corresponding true class and predicted class), indicating that we successfully reduced partially the inter-individual differences. We observed a substantial increase in performance when we re-ran the leave-one-out design using the restricted list of features (mean FSR, minimum depth rate, depth rate variance, maximum absolute roll, maximum roll rate, maximum ODBA, roll rate variance, minimum absolute roll, maximum depth rate, mean depth rate, mean ODBA, mean absolute roll, speed variance, mean roll rate, ODBA variance, depth variance, minimum ODBA, absolute roll skewness, roll rate kurtosis, speed skewness, roll rate skewness, absolute roll kurtosis, absolute roll variance, depth skewness, pitch kurtosis, ODBA skewness, depth kurtosis, speed kurtosis, depth rate skewness, pitch skewness, depth rate kurtosis, ODBA kurtosis, minimum roll rate, pitch variance) (Fig. 6, blue symbols). Indeed, the Global accuracy was higher than when including all features, and the False Positive Rate for suckling blocks decreased. Most importantly, the quantity of detected suckling blocks increased (higher Sensitivity) for the first and third combination. Also the quality of the classification improved globally (increased Precision), even for the combination that had a very poor Precision at first.

Discussion

Our study characterized the suckling behavior in humpback whale calves less than 3 months old using CATS cam tags on one calving ground in the South Western Indian Ocean. Several important aspects of this vital behavior were investigated: (i) duration and frequency, (ii) occurrence in the water column, (iii) modalities, (iv) behavioral signatures, and (v) the possibility to use only accelerometer and depth-derived data to perform an automatic detection of suckling events.

Advantage of using animal-borne camera-equipped multi-sensor tags

Previously, positioning and posture of the calves have been used frequently as a proxy for determining the occurrence of nursing or suckling in humpback whales (Glockner & Venus, 1983; Glockner-Ferrari & Ferrari, 1985; Clapham & Mayo, 1987; Morete et al., 2003; Videsen et al., 2017; Zoidis & Lomac-MacNair, 2017). As previously shown by Tackaberry et al. (Tackaberry et al., 2020) with older calves (studied in their feeding grounds), we found that young calves were also positioned frequently under the mother and in close proximity when diving but they were actually not suckling. The suckling dives (i.e., dives during which suckling was visually confirmed by evidence of a milk cloud), appeared to be much rarer in comparison to non-suckling dives. Consequently, great care must be taken when associating calf positioning with suckling, as these are not necessarily associated.

Animal-borne multi-sensor tag equipped with video camera is a very efficient device for studying suckling behavior, since it can deliver a view from calf's perspective and thus confirm evidence of suckling. This allowed differentiation of events in which the calf started to intentionally touch its mother's mammary slit, from events when it was only positioning itself under the mother. If visual observations were performed from a different perspective or at distance (e.g., surface ship-based, or aerial drone-based), such differentiation between suckling and non-suckling events would not be possible, therefore leading to overestimations of suckling behaviors.

In our study, the view offered by the video camera allowed clear definition of suckling events initiated by physical contact between the calf's snout and the mammary teat. Such contact did not happen every time the calf was under the mother. Suckling behavior was further confirmed by the presence of a milk cloud, as in Tackaberry et al. (2020). This method greatly improved describing suckling behavior with confidence.

Duration and frequency of suckling events

Suckling events were brief and rare (18.8±8.8 s on average and < 2% of observation time). They were shorter in duration compared to those observed by divers in another breeding area (30.6±17 s on average, Zoidis & Lomac-MacNair, 2017). This difference might be related to the method used, as discussed above. The average suckling time reported in the feeding area is quite similar to what we observed (23±7 s, Tackaberry et al., 2020). The associated suckling frequency was, however, much lower in the feeding area (0.3% of the time for an observation between 0745 and 1700 hours) compared to what we observed in the calving area (1.7% between 1000 and 1800 hours). This expected difference suggests suckling rate varies with age (greater for young calves in the breeding grounds versus older calves in the feeding grounds), as is the case with several species of pinnipeds (Oftedal, Boness & Tedman, 1987).

Occurrences of suckling in the water column

As in Tackaberry et al. (2020), most of the suckling events occurred at depth and during dives. Humpback whales may favor nursing at depth to help the calf in maintaining the suckling posture, as the latter would be more buoyant at surface, or for facilitating the thermoregulation of the adult female (Videsen et al., 2017). Regarding the latter, we noticed that the rare surface suckling events only occurred at the end of the day, when the sun was less intense, suggesting a possible link between surface nursing events and females' thermoregulation behavior. Further investigations on suckling during the night are needed for comparison, as most videos were recorded during daylight only (this study; Tackaberry et al., 2020). Nursing behavior might show different patterns at night for the reasons cited above.

The maximum suckling depth recorded (32.6 m) was much shallower than the maximum reported by Tackabery et al. (2020). However, this does not necessarily relate to any physical limits of young calves. Rather, it might be an effect of the environment. Indeed, the average depth of the Sainte Marie channel in Madagascar is 35 m (Trudelle et al., 2016), and in our video recordings it was common to see the seabed when the whales dove to around 30 m depth. All suckling events during dives occurred during the descent and bottom phases only, unlike in Tackaberry et al. (2020) where they also observed suckling events during the ascent phase. One possible explanation is that very young calves found in breeding grounds have less breath capacities (Saloma, 2018) than older calves found in feeding areas. Thus, during the ascent phase of their dive, their priority would be to reach the surface to breathe than to suckle.

Suckling modalities

With respect to posture, as estimated from the video recordings, the calf positioned itself at 30-45° pitch angle relative to the midline of the mother's body when suckling. This is in agreement with past observations (Glockner-Ferrari & Ferrari, 1985; Zoidis & Lomac-MacNair, 2017). However, the pitch recorded by the tags (pitch angle relative to the horizontal) during suckling was always relatively low on average (< 15°). This suggests that the mother is generally oriented facing slightly downward when nursing its calf, even during phases other than the descent (during which the mother does not otherwise have to lean downward).

One of the most innovative findings of our study was the calf's tendency to lean continuously to the side when suckling. Such rolling behavior might be related to the anatomy of the mammary gland, in particular the orientation of the nipple as it is extruded from the mammary slit. If each breast is positioned to that the teat is directed at an angle relative to the midsagittal plane, then the calf would be forced to match that angle with its mouth in order to stimulate the nipple to extrude and then grasp the nipple with its tongue. Video recordings show the calf's tongue is directed laterally, visible in the oral gape of the mouth on one side of the head. The calf must roll to one side in order to align the lateral aspect of its mouth with the mammary slit (rolling left to align its right oral gape against the mother's left breast, and vice versa).

There was also the clear pattern of alternation from one side to the other between successive suckling events. Our data confirm the correspondence between the leaning side and the suckled teat in nine events. The calves rolled to the left to suckle on the left teat and rolled to the right to suckle on the right teat. Why, then, must the calf alternate sides? Perhaps the milk supply of each breast is limited, forcing the calf to suckle both sides in order to obtain a sufficient volume for satiation. This would ensure that the mother's breasts continue to lactate evenly on both sides. Another reason may be related to the milk production/storage/ejection system. It is widely accepted that milk is voluntarily ejected by the mother into the calf's mouth in cetaceans (Slijper, 1966). It is possible that the amount of milk the mother can continuously eject is limited, and the gland may need a refractory period during which it must reset before the next ejection can occur. If so, then alternating which side to suckle maximizes feeding for the calf, while increasing milk delivery efficiency for the mother. In this scenario, the currently suckled breast performs milk ejection, while the previously suckled breast refills in preparation for the next ejection. This would thus force the calf to alternate between the two mammary glands to get enough milk during the successive nursing bouts. In-depth anatomical investigation is needed to answer this question, as there is scant literature on whale lactation or breast anatomy.

Regarding the anatomy of suckling, it is important to note that in a few cases, rhythmic movements of the calf's tongue were observed during suckling events. This suggests that the calf actively participates in directing milk into its mouth. It is unclear whether these movements are stimulatory to the mother's "let down" (nipple erection and milk ejection) reflex, serve to "strip" the nipple (as occurs in many land mammals to squeeze milk out), or create a piston-like suction to draw out milk (perhaps in combination with milk ejection from the mother). The lateral aspect of the calf's rostral tongue is comprised of elongated marginal papillae that may serve to grasp the nipple. This latching-on-nipple function has been proposed for newborn tongues in other whale species (Kastelein & Dubbeldam, 1990; Shindo et al., 2008; Ferrando et al., 2010; Kienle et al., 2015). These marginal papillae may also "zipper" together to form a tube for channeling milk into the mouth, compensating for the calf's lack of lips and cheeks to seal and contain the nipple and the extruded milk. Alternatively, the rhythmic movements of the tongue may be only an artefact produced by the calf's swimming movements as it undulates its body to maintain its suckling position while the tongue maintains constant contact with the nipple. Further study of the calf tongue may provide additional insights to its function.

Behavioral signatures of suckling events

The comparison of the data derived from accelerometer and depth sensor for suckling and adjacent non-suckling events revealed mainly two distinctive characteristics of the suckling behavior: a high absolute roll and a low speed, for an effort (FSR and ODBA) comparable to adjacent non-suckling events. For suckling occurring during the descent phase, there were two additional notable specificities: a low depth rate and a pitch close to zero. This is in sharp contrast with descent non-suckling, where calves showed a high descent rate on average and were directed downwards.

To our knowledge, no study to date has highlighted the characteristics of suckling behavior in terms of roll recorded by the accelerometer. The sustained roll deviance from zero during suckling events is directly related to the aforementioned leaning pattern observed in the videos: leaning to the side facilitates access a particular mammary slit.

With respect to the kinematics, the fact that the speed is lower while still deploying some physical effort (FSR and ODBA) when suckling is consistent with the results of Tackaberry et al. (2020). Even though the pair makes little or no forward movement during suckling, the calf still has to maintain actively the suckling posture to stay in physical contact with the mammary slit. Concerning the descent suckling, the relatively low depth rate may help the calf in maintaining the suckling posture.

Automatic detection of suckling using accelerometer and depth-derived data

In the field of ethology, the advances in machine learning have offered the opportunity to classify behaviors within a complex database (Valletta et al., 2017). Given the stereotyped traits of suckling behavior in humpback whale calves, we tested whether it was possible to apply machine learning algorithms on labelled accelerometer and depth sensor data to automatically discern suckling from non-suckling events. Although accelerometer data have been already largely coupled with supervised machine learning to detect behaviors in various species (Nathan et al., 2012; Carroll et al., 2014; Ladds et al., 2016; Jeantet et al., 2020), this is a first attempt to use it for humpback whales and in the framework of a suckling behavior study.

The biggest challenge in the automatic classification of suckling behavior is probably the class imbalance. Indeed, suckling events are naturally rare (this study; Tackaberry et al., 2020). It is argued that unbalanced data tend to bias predictions in favor of the majority class. This bias is problematic in situations in which missing the minority class case is worse than misclassifying a majority class (Leevy et al., 2018). In the case of suckling versus non-suckling classification, this bias is not a significant concern since there is more interest in minimizing the False Positive Rate for the targeted behavior (i.e. suckling). Such a conservative measure is generally adopted in behavioral detection (Nathan et al., 2012; Carroll et al., 2014; Tennessen et al., 2019). Reporting non-suckling as suckling would be costlier than reporting suckling as non-suckling, as it would induce an overestimation of suckling events. On the other hand, the best model we found for identifying suckling events, the AdaBoostM1, is a model suitable for unbalanced binary classifications (Galar et al., 2011; Leevy et al., 2018). Thus, class imbalance is less of a concern. In order to avoid evaluation bias, the evaluation process was guided primarily by the minimization of the False Positive Rate rather than the Global accuracy, since the latter does not distinguish between the numbers of correctly classified examples of different classes, and can be misleading in the framework of an unbalanced dataset (Galar et al., 2011).

With an initial 60:40 holdout training-testing split, the average Sensitivity and Precision of the AdaBoostM1 model with respect to suckling class were of 0.74 and 0.92 respectively, and the FPR was < 0.01. In other words, the model was able to detect the vast majority of suckling blocks, and about 9 out of 10 blocks reported as suckling corresponded to visually confirmed suckling events.

The Leave-one-out design (blocks from two individuals as training set and the one remaining individual as a testing set), allowed testing whether data from two individuals can be used to detect suckling events in a novel individual. In this design, it was expected that the performance would drop notably (lower Sensitivity and Precision). Indeed, the inter-individual difference has been shown to generally penalize the performance of automatic classification models (Vázquez Diosdado et al., 2015; Ladds et al., 2016). When we reduced the inter-individual difference in terms of suckling by excluding features that introduce substantial inter-individual differences, the performance returned to a reasonable level, regardless of the combination (i.e., the model generalized better and became more robust). Indeed, out of every 10 blocks reported as suckling, at least 6 were confirmed to be correct (Precision > 0.60, and up to 1). The associated Sensitivity was relatively low, indicating that several suckling blocks were not detected. However, the priority goal is to identify suckling blocks that correspond with confidence to suckling (maximizing the Precision), even in a small number. A complete suckling event (start and end) can be determined from initiation and termination of roll deviation, as this was shown to correspond to the suckling blocks with high certainty.

As more data is added to the learning set, we can expect better performance from the models. Indeed, we showed that the size of the training set can slightly influence the accuracy of the results. Similarly, we can also expect that with more individuals, the constraints and restrictions related to inter-individual differences would also be reduced (Ladds et al., 2016).

These results on automatic identification of suckling events are very promising for rapid evaluation of larger data sets, and will open new opportunities for investigating limited data sets. First, the model may be adjusted to accommodate for incomplete observations (altered/missing video file, night-time recording) using only two sensors (3-axis accelerometer and depth sensors). Secondly, the model may be applied to data from multi-sensor tags that lack visual support (video recordings lacking or too dark, tags without camera such as Acousonde and Dtags, etc.).

While these methods will broaden the opportunities in the domain of humpback whale mother-calf behavioral studies (and eventually may be applied to studying other large whales), care must be taken in these early stages. Until the acquisition of large training data (more individuals) occurs, researchers should be aware of the limitations of each trained model. For example, possible differences in gaits and behavioral pattern, with respect to age (Noren, Biedenbach & Edwards, 2006; Saloma, 2018), may hinder the generalization ability of the models. For our trained model, we caution that it should be conservatively applied to less than 3-month old non-newborn calves in a shallow breeding area. Nevertheless, a concrete application example of the automatic identification of suckling behavior would be the investigation of the behavioral time budget of a mother-calf pair over 24 hours (day and night).

Conclusions

Our results provide new knowledge on the suckling behavior of humpback whale calves in their calving ground. Our descriptions, based on videos from the calf's perspective, accelerometer data, and depth data, complement previous studies based on surface and sub-surface observations (Glockner & Venus, 1983; Glockner-Ferrari & Ferrari, 1985; Clapham & Mayo, 1987; Morete et al., 2003; Videsen et al., 2017; Zoidis & Lomac-MacNair, 2017). We found that suckling is more frequent compared to what has been reported in the feeding area (Tackaberry et al., 2020), suggesting a variation of the suckling rate with the calf’s age and underlining the importance of the suckling behavior in the calving area. We also confirmed that most suckling events occur at depth, highlighting the importance of using multi-sensor tags equipped with a video camera in order obtain reliable observations. Finally, we found that suckling behavior is highly stereotyped, especially characterized by a continuously sustained roll deviating significantly from zero and a low speed, and can be detected automatically by supervised machine learning. These findings open new opportunities for the investigation of suckling and nursing behavior in humpback whales, will improve our understanding of mother-young social bond and interactions, and will be useful in studying other baleen whales for which such vital behavior is still undiscovered.

Acknowledgements

We would like to warmly thank Aina F. Ramanampamonjy, Léo Duperret, Mandrindra O. Rakotovao and Nyal Mueenuddin who contributed to the data collection.

References

Balshine S. 2012. *Patterns of parental care in vertebrates*. Oxford University Press Oxford. DOI: 10.1093/acprof:oso/9780199692576.001.0001.

Butcher P, Colefax A, Gorkin R, Kajiura S, López N, Mourier J, Purcell C, Skomal G, Tucker J, Walsh A, Williamson J, Raoult V. 2021. The Drone Revolution of Shark Science: A Review. *Drones* 5:8. DOI: 10.3390/drones5010008.

Cade DE, Barr KR, Calambokidis J, Friedlaender AS, Goldbogen JA. 2017. Determining forward speed from accelerometer jiggle in aquatic environments. *Journal of Experimental*

Biology 221:jeb170449. DOI: 10.1242/jeb.170449.

Cade DE, Friedlaender AS, Calambokidis J, Goldbogen JA. 2016. Kinematic Diversity in Rorqual Whale Feeding Mechanisms. *Current Biology* 26:2617–2624. DOI: 10.1016/j.cub.2016.07.037.

Carroll G, Slip D, Jonsen I, Harcourt R. 2014. Supervised accelerometry analysis can identify prey capture by penguins at sea. *Journal of Experimental Biology* 217:4295–4302. DOI: 10.1242/jeb.113076.

Cartwright R, Sullivan M. 2009. Behavioral ontogeny in humpback whale (Megaptera novaeangliae) calves during their residence in Hawaiian waters. *Marine Mammal Science* 25:659–680. DOI: 10.1111/j.1748-7692.2009.00286.x.

Clapham PJ. 2018. Humpback Whale. In: *Encyclopedia of Marine Mammals*. Elsevier, 489–492. DOI: 10.1016/B978-0-12-804327-1.00154-0.

Clapham PJ, Mayo CA. 1987. Reproduction and recruitment of individually identified humpback whales, Megaptera novaeangliae , observed in Massachusetts Bay, 1979–1985. *Canadian Journal of Zoology* 65:2853–2863. DOI: 10.1139/z87-434.

Faria M-A, DeWeerdt J, Pace F, Mayer F-X. 2013. Observation of a humpback whale (Megaptera novaeangliae) birth in the coastal waters of Sainte Marie Island, Madagascar. *Aquatic Mammals* 39:296. DOI: 10.1578/AM.39.3.2013.296.

Ferrando T, Caresano F, Ferrando S, Gallus L, Wurtz M, Tagliafierro G. 2010. The tongue morphology and lingual gland histochemistry of Ligurian Sea odontocetes. *Marine mammal science* 26:588–601.

Friard O, Gamba M. 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in ecology and evolution* 7:1325–1330. DOI: 10.1111/2041-210X.12584.

Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F. 2011. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42:463–484. DOI: 10.1109/TSMCC.2011.2161285.

Gittleman JL, Thompson SD. 1988. Energy allocation in mammalian reproduction. *American zoologist* 28:863–875. DOI: 10.1093/icb/28.3.863.

Glockner-Ferrari DA, Ferrari MJ. 1985. *Individual identification, behavior, reproduction, and distribution of humpback whales, Megaptera novaeangliae, in Hawaii.*

Glockner DA, Venus SC. 1983. Identification, growth rate, and behaviour of humpback whale (Megaptera novaeangliae) cows and calves in the waters off Maui, Hawaii, 1977--79: 223--258 in Payne, R.(ed), Communication and behavior of whales.

Hall WG, Hudson R, Brake SC. 1988. Terminology for use in investigations of nursing and suckling. *Developmental psychobiology* 23:89–91. DOI: 10.1002/dev.420210107.

Iwata T, Biuw M, Aoki K, Miller PJO, Sato K. 2021. Using an omnidirectional video logger to observe the underwater life of marine animals: Humpback whale resting behaviour. *Behavioural processes* 186:104369. DOI: 10.1016/j.beproc.2021.104369.

Jeantet L, Planas-Bielsa V, Benhamou S, Geiger S, Martin J, Siegwalt F, Lelong P, Gresser J, Etienne D, Hiélard G, others. 2020. Behavioural inference from signal processing using animal-borne multi-sensor loggers: a novel solution to extend the knowledge of sea turtle ecology. *Royal Society open science* 7:200139. DOI: 10.1098/rsos.200139.

Johnson MP, Tyack PL. 2003. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE journal of oceanic engineering* 28:3–12. DOI:

10.1109/JOE.2002.808212.

Kastelein RA, Dubbeldam JL. 1990. Marginal papillae on the tongue of the harbour porpoise (Phocoena phocoena), bottlenose dolphin (Tursiops truncatus) and Commerson’s dolphin (Cephalorhynchus commersonii). *Aquatic Mammals* 15:158–170.

Kienle SS, Ekdale EG, Reidenberg JS, Deméré TA. 2015. Tongue and hyoid musculature and functional morphology of a neonate gray whale (Cetacea, Mysticeti, Eschrichtius robustus). *The Anatomical Record* 298:660–674. DOI: 10.1002/ar.23107.

Ladds MA, Thompson AP, Slip DJ, Hocking DP, Harcourt RG. 2016. Seeing it all: evaluating supervised machine learning methods for the classification of diverse otariid behaviours. *PloS one* 11:e0166898. DOI: 10.1371/journal.pone.0166898.

Leevy JL, Khoshgoftaar TM, Bauder RA, Seliya N. 2018. A survey on addressing high-class imbalance in big data. *Journal of Big Data* 5:1–30. DOI: 10.1186/s40537-018-0151-6.

Liaw A, Wiener M. 2002. Classification and regression by randomForest. *R news* 2:18–22.

López LMM, Miller PJO, De Soto NA, Johnson M. 2015. Gait switches in deep-diving beaked whales: Biomechanical strategies for long-duration dives. *Journal of Experimental Biology* 218. DOI: 10.1242/jeb.106013.

Morete ME, Freitas A, Engel MH, Pace III RM, Clapham PJ. 2003. A novel behavior observed in humpback whales on wintering grounds at Abrolhos Bank (Brazil). *Marine Mammal Science* 19:694–707.

Nathan R, Spiegel O, Fortmann-Roe S, Harel R, Wikelski M, Getz WM. 2012. Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *Journal of Experimental Biology* 215:986–996. DOI: 10.1242/jeb.058602.

Noren SR, Biedenbach G, Edwards EF. 2006. Ontogeny of swim performance and mechanics in bottlenose dolphins (Tursiops truncatus). *Journal of Experimental Biology* 209:4724–4731. DOI: 10.1242/jeb.02566.

Oftedal OT. 1993. The adaptation of milk secretion to the constraints of fasting in bears, seals, and baleen whales. *Journal of dairy science* 76:3234–3246. DOI: 10.3168/jds.S0022-0302(93)77660-2.

Oftedal OT, Boness DJ, Tedman RA. 1987. The behavior, physiology, and anatomy of lactation in the pinnipedia. In: *Current mammalogy*. Springer, 175–245. DOI: 10.1007/978-1-4757-9909-5_6.

Saloma A. 2018. Humpback whale (Megaptera novaeangliae) mother-calf interactions. Université d’Antananarivo & Université Paris-Saclay.

Shindo J, Yamada TK, Yoshimura K, Kageyama I. 2008. Morphology of the tongue in a newborn stejneger’s beaked whale (Mesoplodon stejnegeri). *Okajimas folia anatomica Japonica* 84:121–124. DOI: 10.2535/ofaj.84.121.

Simon M, Johnson M, Madsen PT. 2012. Keeping momentum with a mouthful of water: behavior and kinematics of humpback whale lunge feeding. *Journal of Experimental Biology* 215:3786–3798. DOI: 10.1242/jeb.071092.

Slijper EJ. 1966. Functional morphology of the reproductive system in Cetacea. In: *Whales, dolphins, and porpoises*. 277–319. DOI: 10.1525/9780520321373-017.

Stimpert AK, Mattila D, Nosal E-M, Au WWL. 2012. Tagging young humpback whale calves: methodology and diving behavior. *Endangered Species Research* 19:11–17. DOI: 10.3354/esr00456.

Tackaberry JE, Cade DE, Goldbogen JA, Wiley DN, Friedlaender AS, Stimpert AK. 2020. From a calf's perspective: humpback whale nursing behavior on two US feeding grounds. *PeerJ* 8:e8538. DOI: 10.7717/peerj.8538.

Tennessen JB, Holt MM, Hanson MB, Emmons CK, Giles DA, Hogan JT. 2019. Kinematic signatures of prey capture from archival tags reveal sex differences in killer whale foraging activity. *Journal of Experimental Biology* 222. DOI: 10.1242/jeb.191874.

Trudelle L, Cerchio S, Zerbini AN, Geyer Y, Mayer F-X, Jung J-L, Hervé MR, Pous S, Sallée J-B, Rosenbaum HC, others. 2016. Influence of environmental parameters on movements and habitat utilization of humpback whales (Megaptera novaeangliae) in the Madagascar breeding ground. *Royal Society open science* 3:160616. DOI: 10.1098/rsos.160616.

Valletta JJ, Torney C, Kings M, Thornton A, Madden J. 2017. Applications of machine learning in animal behaviour studies. *Animal Behaviour* 124:203–220. DOI: 10.1016/j.anbehav.2016.12.005.

Vázquez Diosdado JA, Barker ZE, Hodges HR, Amory JR, Croft DP, Bell NJ, Codling EA. 2015. Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system. *Animal Biotelemetry* 3:15. DOI: 10.1186/s40317-015-0045-8.

Videsen SKA, Bejder L, Johnson M, Madsen PT. 2017. High suckling rates and acoustic crypsis of humpback whale neonates maximise potential for mother--calf energy transfer. *Functional Ecology* 31:1561–1573. DOI: 10.1111/1365-2435.12871.

Wilson RP, White CR, Quintana F, Halsey LG, Liebsch N, Martin GR, Butler PJ. 2006. Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *Journal of Animal Ecology* 75:1081–1090. DOI: 10.1111/j.1365-2656.2006.01127.x.

Zoidis AM, Lomac-MacNair KS. 2017. A note on suckling behavior and laterality in nursing humpback whale calves from underwater observations. *Animals* 7:51. DOI: 10.3390/ani7070051.

Figure 1

Machine learning workflow for supervised classification of suckling events observed in free ranging humpback whale calves

The pie charts represent the class distribution (suckling versus non-suckling).

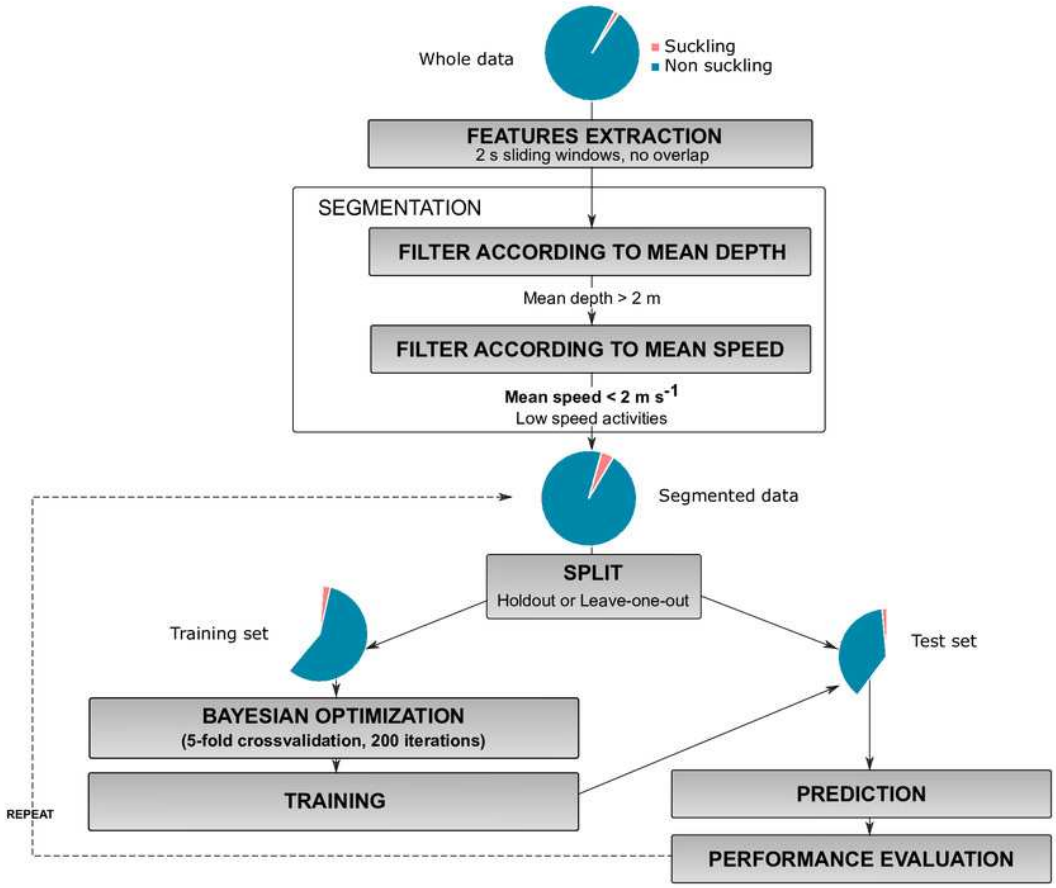

Whole data
Suckling
Non suckling
FEATURES EXTRACTION
2 s sliding windows, no overlap
SEGMENTATION
FILTER ACCORDING TO MEAN DEPTH
Mean depth > 2 m
FILTER ACCORDING TO MEAN SPEED
Mean speed < 2 m s-1
Low speed activities
Segmented data
SPLIT
Holdout or Leave-one-out
Training set
Test set
BAYESIAN OPTIMIZATION
(5-fold crossvalidation, 200 iterations)
TRAINING
REPEAT
PREDICTION
PERFORMANCE EVALUATION

Figure 2

Depth, pitch, roll, speed, FSR and ODBA during and around suckling events for each calf.

Suckling events are identified in red in the depth profile and by the yellow box in the pitch(blue)/roll(yellow) profiles. Yellow box placed on top indicate that the calf was observed leaning to the right side on the corresponding video. Yellow box placed on the bottom indicate that the calf was rather observed leaning to the left side. Events during which the teat suckled by the calf was clearly evident on the corresponding video are marked with (R) or (L): (R) for right teat and (L) for left teat.


Calf1
Calf2
Calf3
Depth (m)
Pitch/Roll (°)
Speed (m s⁻¹)
FSR (Hz)
ODBA (m s⁻²)
(L)
(R)
12:57:00
12:58:00
12:59:00
13:00:00
13:01:00
12:09:00
12:10:00
12:11:00
12:12:00
12:36:00
12:37:00
12:38:00
12:39:00
13:09:00
13:10:00
13:11:00
13:12:00
13:13:00
13:14:00
13:16:00
13:17:00
13:19:00
13:20:00
16:54:00
16:55:00
16:56:00
17:41:00
17:42:00
17:43:00
17:44:00
17:45:00
17:46:00
17:47:00

Figure 3

Screen capture of a video footage of calf suckling on the right teat of its mother.

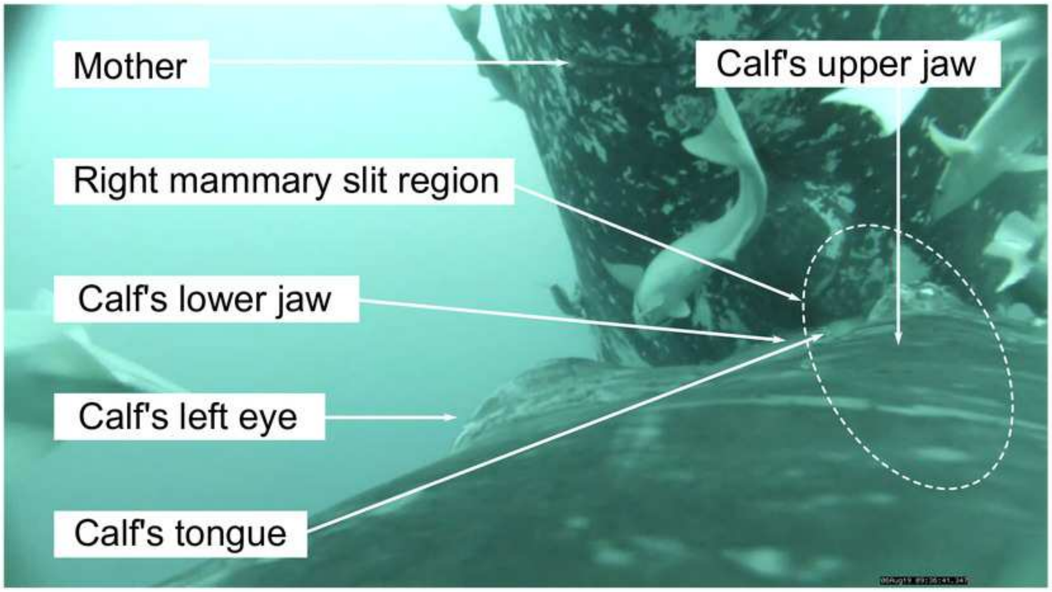

Mother
Calf's upper jaw
Right mammary slit region
Calf's lower jaw
Calf's left eye
Calf's tongue

Figure 4

Comparison of suckling and adjacent non-suckling events with respect to activity phases.

Mean and median are indicated by diamond marks and bold horizontal line respectively. No ascent suckling has been observed.

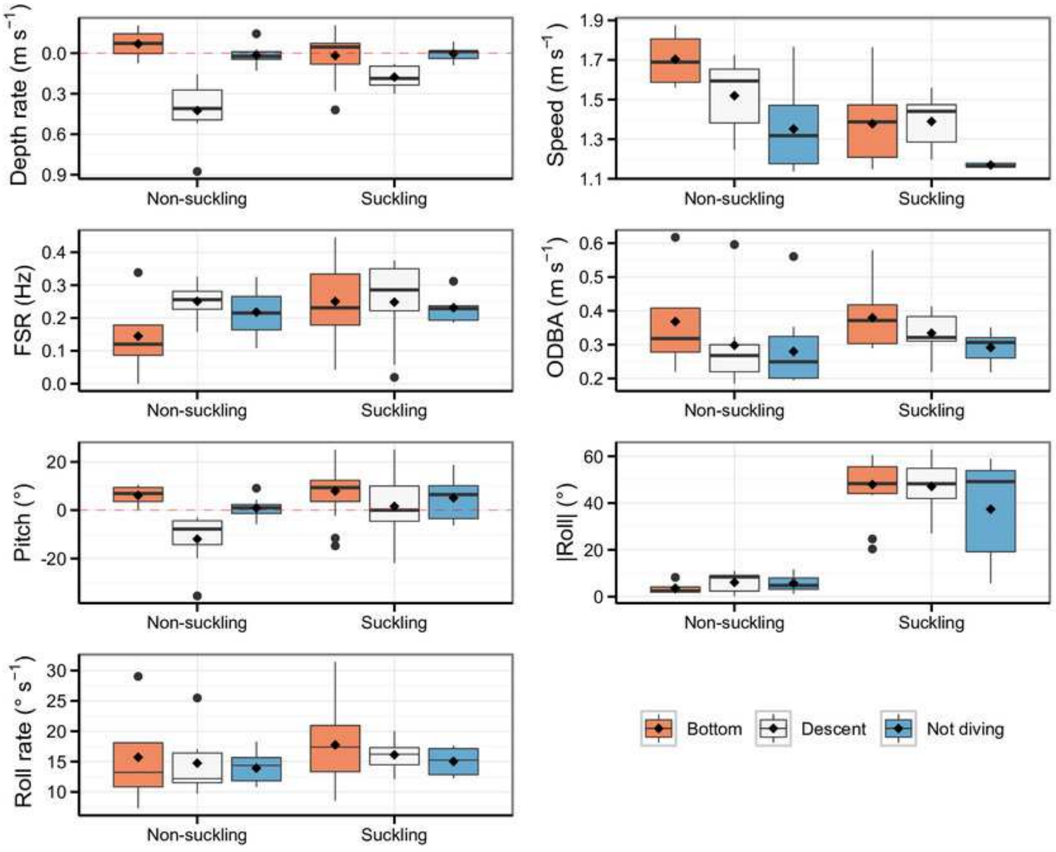

Depth rate (m s^{-1})
0.0
0.3
0.6
0.9
Non-suckling
Suckling
Speed (m s^{-1})
1.9
1.7
1.5
1.3
1.1
Non-suckling
Suckling
FSR (Hz)
0.4
0.3
0.2
0.1
0.0
Non-suckling
Suckling
ODBA (m s^{-1})
0.6
0.5
0.4
0.3
0.2
Non-suckling
Suckling
Pitch (°)
20
0
-20
Non-suckling
Suckling
|Roll| (°)
60
40
20
0
Non-suckling
Suckling
Roll rate (° s^{-1})
30
25
20
15
10
Non-suckling
Suckling
Bottom
Descent
Not diving

Figure 5

Models’ performance in automatically identifying suckling blocks.

A 60:40 holdout training-testing split was used. (A) Sensitivity versus Precision plot of 7 models pre-selected using a Bayesian optimization approach. The symbols show the mean values and the SD. The models were pre-selected in the optimization procedure over 90 runs (30 for each type of classifier). (B) Example of global results from automatic identification of suckling behavior using the AdaBoostM1 model. (C) Example of results from automatic identification of suckling behavior using the AdaBoostM1 model for Calf1. (D) Performance of the AdaBoostM1 model when the training set size is reduced. The symbols show the mean values and the SD (*N* = 30 runs).

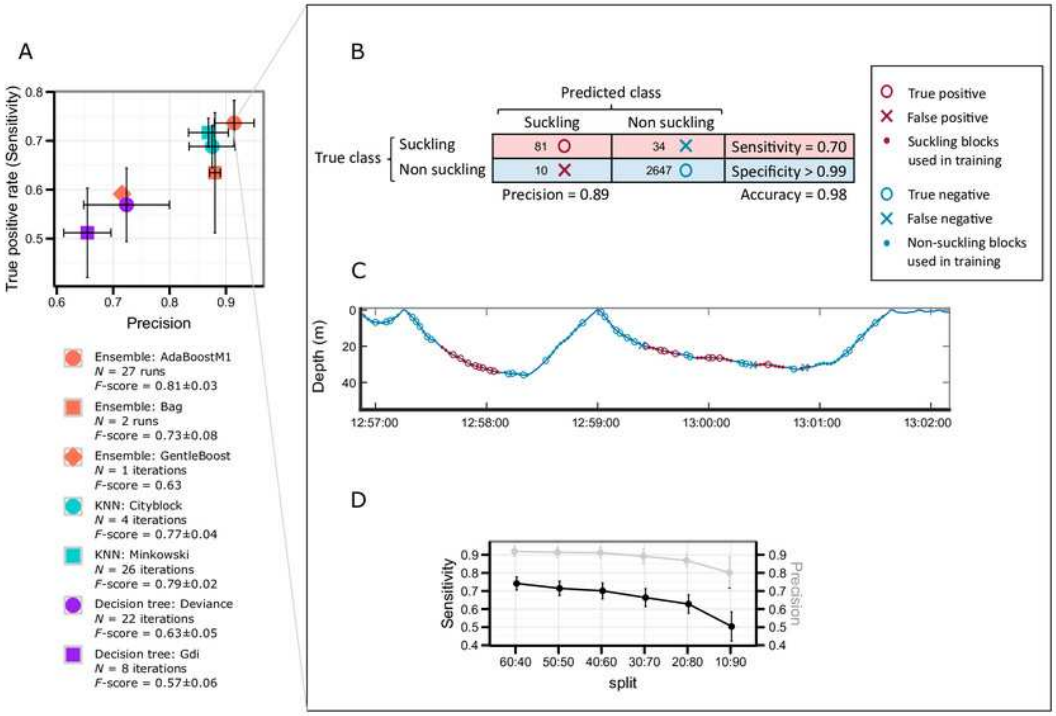

A
True positive rate (Sensitivity)
0.8
0.7
0.6
0.5
0.6
0.7
0.8
0.9
Precision
Ensemble: AdaBoostM1
N = 27 runs
F-score = 0.81±0.03
Ensemble: Bag
N = 2 runs
F-score = 0.73±0.08
Ensemble: GentleBoost
N = 1 iterations
F-score = 0.63
KNN: Cityblock
N = 4 iterations
F-score = 0.77±0.04
KNN: Minkowski
N = 26 iterations
F-score = 0.79±0.02
Decision tree: Deviance
N = 22 iterations
F-score = 0.63±0.05
Decision tree: Gdi
N = 8 iterations
F-score = 0.57±0.06
B
Predicted class
Suckling
Non suckling
True class
Suckling
Non suckling
81
34
10
2647
Sensitivity = 0.70
Specificity > 0.99
Precision = 0.89
Accuracy = 0.98
True positive
False positive
Suckling blocks used in training
True negative
False negative
Non-suckling blocks used in training
C
Depth (m)
0
20
40
12:57:00
12:58:00
12:59:00
13:00:00
13:01:00
13:02:00
D
Sensitivity
0.9
0.8
0.7
0.6
0.5
0.4
Precision
60:40
50:50
40:60
30:70
20:80
10:90
split

Figure 6

Sensitivity versus precision plot obtained from AdaBoostM1 model for suckling events using leave-one-out splits (data from two individuals used for training while one individual is kept unseen for test).

The plot shows the classifier's performance when all features were included (open symbols) versus when the features that contributed most to the inter-individual variation in terms of suckling were excluded (blue symbols). For comparison, mean results from 60:40 holdout split (purple symbol) and the worst scenario (red cross symbol) are showed.

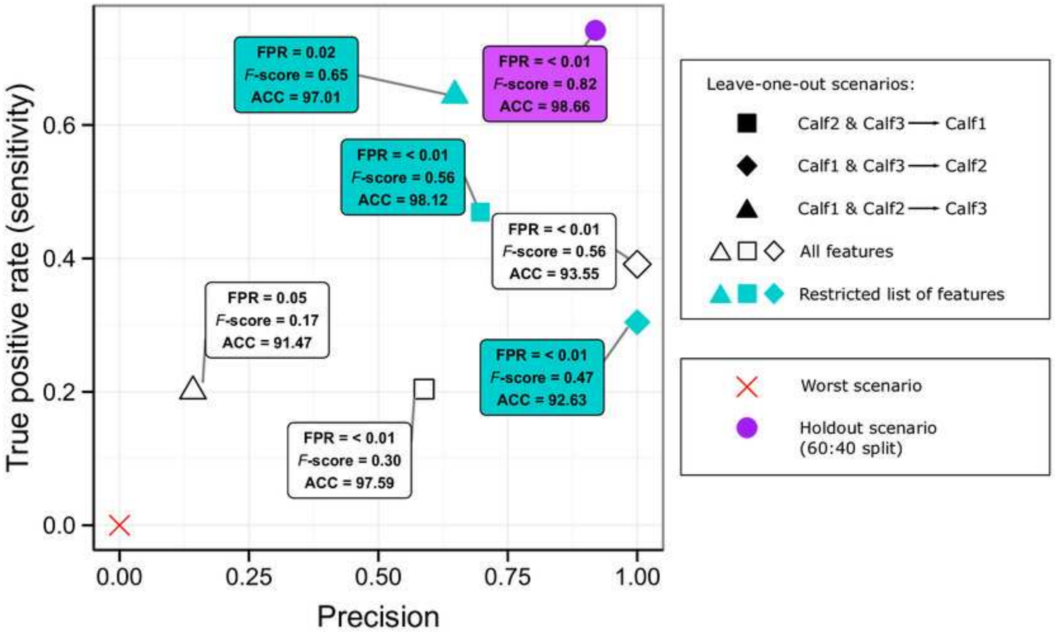
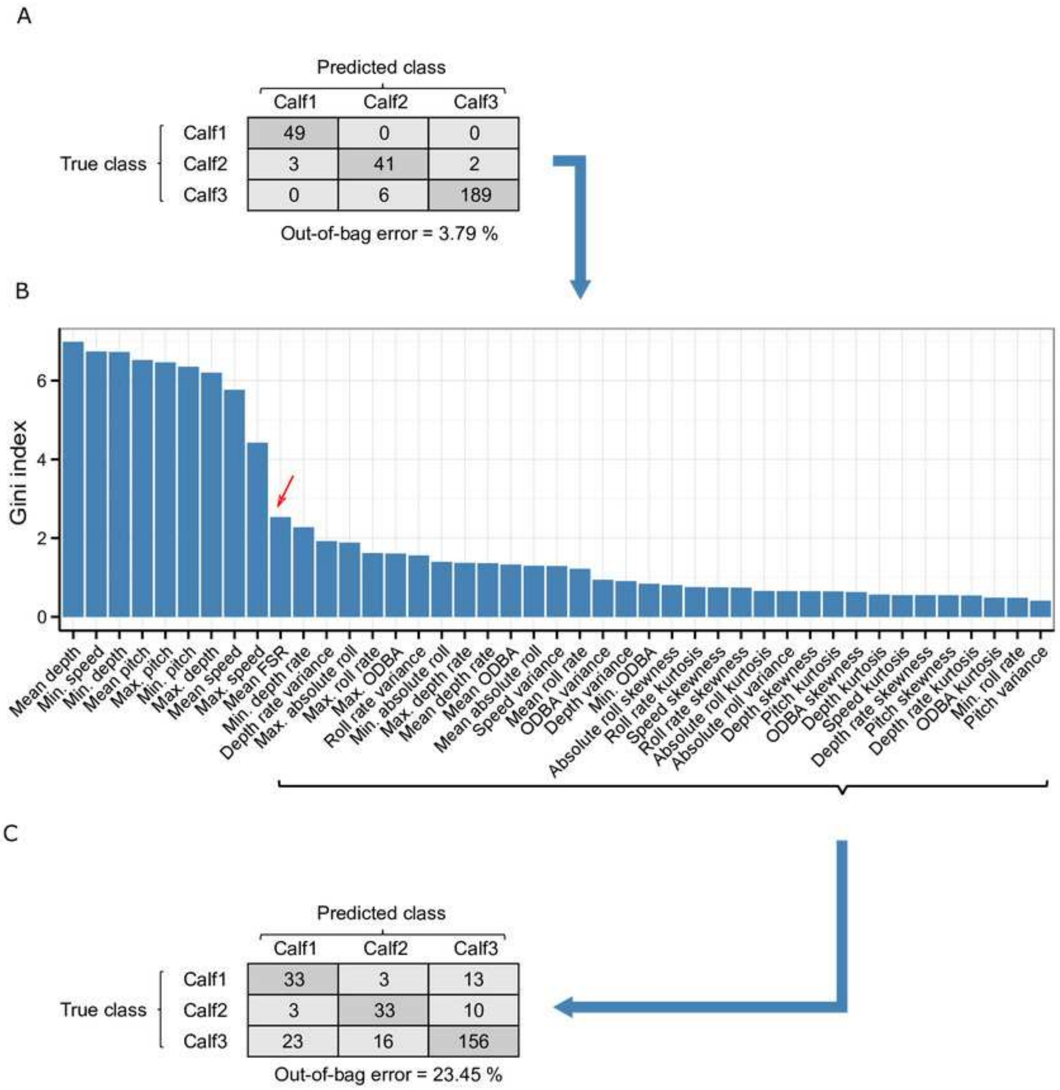

FPR = 0.02
F-score = 0.65
ACC = 97.01
FPR = < 0.01
F-score = 0.82
ACC = 98.66
FPR = < 0.01
F-score = 0.56
ACC = 98.12
FPR = < 0.01
F-score = 0.56
ACC = 93.55
FPR = 0.05
F-score = 0.17
ACC = 91.47
FPR = < 0.01
F-score = 0.47
ACC = 92.63
FPR = < 0.01
F-score = 0.30
ACC = 97.59
True positive rate (sensitivity)
0.0
0.2
0.4
0.6
0.00
0.25
0.50
0.75
1.00
Precision
Leave-one-out scenarios:
Calf2 & Calf3 → Calf1
Calf1 & Calf3 → Calf2
Calf1 & Calf2 → Calf3
All features
Restricted list of features
Worst scenario
Holdout scenario
(60:40 split)

Figure 7

Analysis of the inter-individual difference between suckling blocks.

(A) Result of a random forest classification (confusion matrix) of the suckling blocks by individuals when all features were used. (B) Importance of the features in the classification by individuals. A sharp drop of the Gini index is noticeable after the 9th feature (red arrow). (C) Result of a random forest classification (confusion matrix) of the suckling blocks by individuals when the first 9 most important features were excluded.



A

True class \ Predicted class	Calf1	Calf2	Calf3
Calf1	49	0	0
Calf2	3	41	2
Calf3	0	6	189

Out-of-bag error = 3.79 %

C

True class \ Predicted class	Calf1	Calf2	Calf3
Calf1	33	3	13
Calf2	3	33	10
Calf3	23	16	156

Out-of-bag error = 23.45 %

Table 1(on next page)

Characteristics of all suckling events for each calf.

Date-times are presented in the format DD/MM/YYYY hhmm – hhmm, and corresponds to date, tag attachment time, and tag detachment time. The letters d/b/a indicate the dive phase: descent/bottom/ascent. The letters R/L indicate the rolling side of the calf: Right/Left. The budget represents the proportion of time that the animal was observed suckling. The remaining values are presented following the format mean±SD (min ; max). Note the use of absolute value for roll, to emphasize the roll deviation from zero (regardless of the side).

ID	Date-time	Occurrence	N (d/b/a) – (R/L)	Duration (s)	Budget (%)	Depth (m)	Depth rate (m s^{-1})	Speed (m s^{-1})	FSR (Hz)	Pitch (°)	|Roll| (°)	Roll rate (° s^{-1})	ODBA (m s^{-2})
Calf1	14/09/2018 1134-1418	During dive	5 (2/3/0) – (3/2)	22±9.1 (8.8 ; 34.2)	1.2	28.2±3.8 (22.3 ; 31.8)	0.12±0.23 (-0.21 ; 0.42)	1.46±0.05 (1.39 ; 1.51)	0.22±0.06 (0.14 ; 0.29)	-8±13 (-22 ; 11)	38±12 (20 ; 48)	16±4 (12 ; 20)	0.39±0.09 (0.31 ; 0.54)
Calf2	06/08/2019 1057-1234	During dive	4 (1/3/0) – (3/1)	25±16.7 (6.2 ; 44.9)	1.7	25.1±5.8 (19.3 ; 32.6)	0.04±0.19 (-0.14 ; 0.28)	1.63±0.1 (1.56 ; 1.76)	0.16±0.07 (0.06 ; 0.23)	15±9 (4 ; 25)	51±8 (44 ; 63)	25±7 (16 ; 32)	0.5±0.07 (0.41 ; 0.58)
Calf3	09/08/2019 1003-1802	During dive	20 (6/14/0) – (10/10)	16.5±5.8 (4.8 ; 25.5)	1.4	15.43±3.78 (8.9 ; 21)	0.06±0.14 (-0.12 ; 0.3)	1.31±0.14 (1.15 ; 1.67)	0.28±0.12 (0.02 ; 0.45)	8±9 (-9 ; 25)	50±10 (25 ; 63)	16±4 (9 ; 24)	0.33±0.06 (0.22 ; 0.44)
		At surface	5 (-/-/-)– (3/2)	19.6±10 (7.9 ; 35)	0.4	6.2±2.7 (2 ; 8.4)	0±0.07 (-0.09 ; 0.09)	1.17±0.02 (1.15 ; 1.19)	0.23±0.05 (0.19 ; 0.31)	5±10 (-6 ; 19)	37±24 (6 ; 59)	15±3 (13 ; 17)	0.29±0.05 (0.22 ; 0.35)

Table 2(on next page)

Comparison of the number of non-suckling dives during which calves remained under the mother for at least 5 s at least once, and the number of suckling dives.

Individuals	Number			
	Non-suckling dives		Suckling dives	Total dives
	Calf observed staying in close proximity beneath the mother at least 5 s during the dive	Calf never observed staying in close proximity beneath the mother at least 5 s during the dive		
Calf1	17	18	2	37
Calf2	5	2	3	10
Calf3	20	21	9	50

Table 3(on next page)

Summary table of the mixed effect models for the characteristics of suckling events.

The models included the suckling status and diving phase as fixed effect and individuals as random effect (reference level = bottom and adjacent non-suckling, i.e. adjacent bottom non-suckling). Significant *P* (< 0.05 in this study) are marked in bold. CI: Confidence interval. SE: Standard Error of estimate.

Response	**Fixed effect**	**Effect estimate (β)**	**95% CI**	**SE**	**t(48)**	***P***
Depth rate	(Intercept)	-0.05	[-0.19, 0.10]	0.07	-0.66	0.511
	Suckling	0.08	[-0.07, 0.23]	0.08	1.06	0.291
	Descent	0.49	[0.32, 0.65]	0.09	5.61	< **0.001**
	Surface	0.08	[-0.08, 0.24]	0.08	0.98	0.326
	Suckling*Descent	-0.33	[-0.53, -0.13]	0.10	-3.21	**0.001**
	Suckling*Surface	-0.08	[-0.29, 0.13]	0.11	-0.78	0.438
Speed	(Intercept)	1.78	[1.58, 1.99]	0.10	17.38	**< .001**
	Suckling	-0.33	[-0.49, -0.18]	0.08	-4.19	**< .001**
	Descent	-0.22	[-0.40, -0.04]	0.09	-2.4	**0.016**
	Surface	-0.36	[-0.52, -0.19]	0.08	-4.2	**< .001**
	Suckling*Descent	0.22	[0.01, 0.44]	0.11	2.07	**0.038**
	Suckling*Surface	0.21	[-0.01, 0.43]	0.11	1.84	0.066
FSR	(Intercept)	0.14	[0.05, 0.24]	0.05	2.9	**0.004**
	Suckling	0.11	[0.00, 0.21]	0.05	1.94	0.053
	Descent	0.11	[-0.02, 0.23]	0.06	1.7	0.09
	Surface	0.07	[-0.04, 0.19]	0.06	1.26	0.208
	Suckling*Descent	-0.11	[-0.25, 0.04]	0.07	-1.46	0.145
	Suckling*Surface	-0.09	[-0.24, 0.06]	0.08	-1.21	0.227
ODBA	(Intercept)	0.42	[0.30, 0.53]	0.06	6.9	**< .001**
	Suckling	0	[-0.09, 0.10]	0.05	0.04	0.97
	Descent	-0.09	[-0.20, 0.02]	0.06	-1.63	0.104
	Surface	-0.09	[-0.20, 0.01]	0.05	-1.81	0.071
	Suckling*Descent	0.04	[-0.09, 0.17]	0.07	0.6	0.55
	Suckling*Surface	0.04	[-0.10, 0.17]	0.07	0.54	0.592
Pitch	(Intercept)	1.53	[-12.16, 15.22]	6.98	0.22	0.827
	Suckling	4.24	[-4.83, 13.30]	4.62	0.92	0.36
	Descent	-16.1	[-26.45, -5.75]	5.28	-3.05	**0.002**
	Surface	-3.75	[-13.37, 5.87]	4.91	-0.76	0.445
	Suckling*Descent	10.91	[-1.34, 23.17]	6.25	1.75	0.081
	Suckling*Surface	-0.85	[-13.61, 11.91]	6.51	-0.13	0.896
|Roll|	(Intercept)	3.31	[-7.37, 13.98]	5.45	0.61	0.544
	Suckling	44.42	[32.88, 55.96]	5.89	7.54	**< .001**
	Descent	2.61	[-10.60, 15.81]	6.74	0.39	0.699
	Surface	2.17	[-10.13, 14.46]	6.27	0.35	0.73
	Suckling*Descent	-3.34	[-19.01, 12.34]	8.00	-0.42	0.677
	Suckling*Surface	-12.85	[-29.06, 3.36]	8.27	-1.55	0.12
Roll rate	(Intercept)	16.7	[11.29, 22.12]	2.76	6.04	**< .001**
	Suckling	2.07	[-3.11, 7.25]	2.64	0.78	0.433
	Descent	-1.38	[-7.30, 4.53]	3.02	-0.46	0.647
	Surface	-1.73	[-7.23, 3.77]	2.81	-0.62	0.538
	Suckling*Descent	-0.26	[-7.27, 6.76]	3.58	-0.07	0.943
	Suckling*Surface	-0.14	[-7.43, 7.14]	3.72	-0.04	0.969

1