Received: 19 March 2020 | Accepted: 31 August 2020

DOI: 10.1111/2041-210X.13491

RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Seek and learn: Automated identification of microevents in animal behaviour using envelopes of acceleration data and machine learning

Pritish Chakravarty[1] | Gabriele Cozzi[2,3] | Hooman Dejnabadi[4] | Pierre-Alexandre Léziart[1,5] | Marta Manser[2,3] | Arpat Ozgul[2,3] | Kamiar Aminian[1]

[1]School of Engineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[2]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zürich, Switzerland

[3]Kalahari Research Centre, Kuruman River Reserve, Van Zylsrus, South Africa

[4]EddySonix, Orbe, Switzerland

[5]Sciences Industrielles de l'Ingénieur, Ecole Normale Supérieure de Rennes, Rennes, France

Correspondence
Pritish Chakravarty
Email: chakravartypritish@gmail.com

Funding information
Swiss National Science Foundation, Grant/Award Number: CR32I3_159743

Handling Editor: Robert B. O'Hara

Abstract

1. Animal-borne accelerometers have been used across more than 120 species to infer biologically significant information such as energy expenditure and broad behavioural categories. While the accelerometer's high sensitivity to movement and fast response times present the unprecedented opportunity to resolve fine-scale behaviour, leveraging this opportunity will require overcoming the challenge of developing general, automated methods to analyse the nonstationary signals generated by nonlinear processes governing erratic, impulsive movement characteristic of fine-scale behaviour.
2. We address this issue by conceptualising fine-scale behaviour in terms of characteristic microevents: impulsive movements producing brief (<1 s) shock signals in accelerometer data. We propose a 'seek-and-learn' approach: a novel microevent detection step first locates where shock signals occur ('seek') by searching for peaks in envelopes of acceleration data. Robust machine learning ('learn') employing meaningful features then separates microevents. We showcase the application of our method on tri-axial accelerometer data collected on 10 free-living meerkats *Suricata suricatta* for four fine-scale foraging behaviours – searching for digging sites, one-armed digging, two-armed digging and head jerks during prey ingestion. Annotated videos served as groundtruth, and performance was benchmarked against that of a variety of classical machine learning approaches.
3. Microevent identification (μEvId) with eight features in a three-node hierarchical classification scheme employing logistic regression at each node achieved a mean overall accuracy of >85% during leave-one-individual-out cross-validation, and exceeded that of the best classical machine learning approach by 8.6%. μEvId was found to be robust not only to inter-individual variation but also to large changes in model parameters.
4. Our results show that microevents can be modelled as impulse responses of the animal body-and-sensor system. The microevent detection step retains only informative regions of the signal, which results in the selection of discriminative features

that reflect biomechanical differences between microevents. Moving-window-based classical machine learning approaches lack this prefiltering step, and were found to be suboptimal for capturing the nonstationary dynamics of the recorded signals. The general, automated technique of μEvId, together with existing models that can identify broad behavioural categories, provides future studies with a powerful toolkit to exploit the full potential of accelerometers for animal behaviour recognition.

KEYWORDS

accelerometer, behaviour recognition, fine-scale behaviour, foraging, machine learning, meerkat, microevent, signal envelope

1 | INTRODUCTION

Fine-scale information obtained through animal-attached sensors recording at high frequency has proven to be invaluable for understanding behaviour, the environment and the relationship between the two (Kays et al., 2015; Ropert-Coudert & Wilson, 2005) across more than 120 species as of 2013 (Brown et al., 2013). By measuring both a dynamic component of movement indicative of activity intensity (Müller & Schrader, 2003; Wilson et al., 2006), and a static component due to Earth's gravity indicative of posture (Hansson et al., 2001), accelerometers have enabled inference of broad behavioural categories such as locomotion, resting and foraging (e.g. Chakravarty, Cozzi, et al., 2019; Nathan et al., 2012). Due to their fast response times and high sensitivity to movement (Béliveau et al., 1999), accelerometers are capable of capturing rapid changes in acceleration during animal motion. This ability presents the unprecedented opportunity to resolve fine-scale animal behaviour. However, leveraging this opportunity will require overcoming the challenge of developing general, automated methods to analyse the complex signals generated by fine-scale behaviour.

Fine-scale behaviour within broad categories – for example, food search, food capture and food ingestion within the broad context of foraging – usually involves brief, abrupt, situation-specific manoeuvres which we shall henceforth call 'microevents'. Some examples of microevents are bites or head thrusts during prey capture in hooded seals *Cystophora cristata* (Suzuki et al., 2009), digging in ground squirrels *Spermophilus elegans* (Zegers, 1981) and ground pecking or scratching in red junglefowl *Gallus gallus* L. (Dawkins, 1989). Not only are microevents common, but the ability to identify them can have important consequences for biological inference based on classified behaviour, for example, the duration and frequency of bites during grazing can vary with type of vegetation consumed (Wilson, Holton, et al., 2018).

Sudden muscle activation by the animal or induced sensor impact during a microevent is akin to an impulse given to the animal body-and-sensor system, and the brief (<1 s) shock signal recorded by the accelerometer akin to the system's impulse response. The impulse response of a dynamical system measures the time profile of the effect of shocks at a given point in time on the expected future values of variables in the system (Pesaran & Shin, 1998). Knowing the impulse response is invaluable since it completely characterises a linear, time-invariant system (Phillips et al., 2003). Sensor-bearing animal bodies, however, are nonlinear, time-varying systems: the impulse response varies with time. This is in part due to body tissue viscoelasticity (Hill, 1938), but perhaps mainly due to nonlinear movement artefacts arising from sensors that are attached to the skin (e.g. for tape-, clamp- or glue-on tags) or loosely suspended (e.g. collar, leg bracelet; Brown et al., 2013). The local mobility of skin-attached markers in human biomechanics studies, termed soft tissue artefact, is known to lead to errors in recorded trajectories that can be of the same order of magnitude as the underlying joint motions (Camomilla et al., 2017). Contact forces arising from collisions of loosely suspended sensors with the body during dynamic motion vary nonlinearly with the shape, surface conditions and mechanical properties of the contacting bodies (Koshy et al., 2013). These nonlinearities yield nonstationary signals that are difficult to analyse using common techniques such as statistical or Fourier analysis.

Existing methods to recognise animal behaviour from acceleration data fall under two broad categories: automated and semiautomated. In automated recognition, descriptive features are first computed from a moving window of fixed size that is slid across acceleration data. Different behaviours are then separated using thresholds (e.g. Viviant et al., 2010; Watanabe & Takahashi, 2013) or machine learning algorithms (e.g. Martiskainen et al., 2009; Nathan et al., 2012). Features are commonly statistical summaries of acceleration data, such as mean, standard deviation and skewness, but have also been derived from signal spectra computed through wavelet transformation (Sakamoto et al., 2009). The preset window size for feature computation in these methods precludes identification of individual microevents that may have occurred within a window. Missing this information might have important consequences for biological inference from classified behaviour, for example, the duration and frequency of sheep bites during grazing can be variable according to the type of vegetation that individuals consume (Wilson, Holton, et al., 2018). Semiautomated recognition involves manually characterising signal patterns followed by classifying using a decision tree (Wilson, Neate, et al., 2018). While this approach does not have the limitation of requiring a fixed-size moving window to operate, the authors report that it requires appreciable investment in time and understanding (Wilson, Neate, et al., 2018). Further, the same type of microevent may generate different waveforms on different occasions depending on time-varying body-sensor

impact characteristics, which further complicates the task of characterising waveforms manually.

Despite the complexity of microevent-generated acceleration signals, however, one common trait is that a microevent would produce a transient acceleration signal of limited duration, possibly containing multiple peaks of varying amplitudes and inter-peak intervals. Such acceleration impulse responses have been observed in diverse situations: (a) in animal studies, for example, mouth openings during prey capture attempts in marine predators (e.g. Iwata et al., 2012; Suzuki et al., 2009), bites during sheep grazing (Wilson, Neate, et al., 2018) and eating in chipmunks (Hammond et al., 2016), (b) motion of flexible multibody robotic systems (Hariharesan & Barhorst, 1999), (c) in humans during falls (Bourke et al., 2007) and fidgeting (Esseiva et al., 2018) and (d) fault-induced changes in vibration patterns of defective bearings (McFadden & Smith, 1984). The field of 'bearing condition monitoring', which deals with the measurement of vibration using accelerometers for the detection of defects in rolling element bearings, offers two key ideas that can help automate the detection of microevent signal patterns. The first idea is defining 'events' as transient waves consisting of a group of peaks (Tandon & Choudhury, 1999). The second idea is to smooth acceleration signals using envelopes. A signal envelope is defined as a curve that encloses the signal by outlining its local extremes in a smooth manner (Johnson Jr. et al., 2011); it is used to smooth the high-frequency impulse response arising from fault-induced impacts in order to detect the moments (or frequency) at which impacts occur (McFadden & Smith, 1984).

In this study, we present a method for automatically identifying behavioural microevents that combines considerations of the physics of animal body-and-sensor systems, the technique of signal enveloping, feature engineering and robust machine learning. In our method, we: (a) automatically detect microevents using a novel algorithm based on signal enveloping, (b) engineer meaningful features to characterise microevents and (c) use robust machine learning algorithms to classify microevents. We showcase the application of our method on data collected on 10 wild-living meerkats *Suricata suricatta*, where fine-scale foraging behaviour such as searching, one-armed and two-armed digging, and prey ingestion were groundtruthed against annotated videos. We compare the performance of microevent identification (μEvId) against that of a variety of classical machine learning approaches, and present an analysis of the robustness of μEvId performance to changes in model parameters.

2 | MATERIALS AND METHODS

2.1 | Microevent identification

2.1.1 | Detecting microevents in recorded acceleration

Our microevent (μE) detection method involves: (a) selecting the accelerometer axis on which μE's will be detected, (b) computing signal amplitude and (c) segmenting the amplitude to find locations and spans of μE's. Custom MATLAB code (microevent) to detect microevents, along with a vignette with pseudocode explaining its usage (*vignette_microevent*), are supplied as Supplementary Information.

Axis selection and pre-filtering

The first step is to choose one axis of the accelerometer – out of surge (anterior-posterior), sway (medial-lateral) and heave (inferior-superior) – that best records μE's. For instance, the surge axis has been shown to be sensitive to fine-scale behaviours such as head movements during prey capture (e.g. Suzuki et al., 2009; Viviant et al., 2010). This step is followed by pre-filtering data from this axis. Low-pass filtering to remove unwanted sensor- and analogue-to-digital signal quantisation-generated noise (Usui & Amidror, 1982) should suffice for most applications, for instance when bursts of activity associated with a μE of interest are separated by short periods of relative inactivity. If, however, the μE of interest is performed during other dynamic activity of similar intensity, for instance, prey-capture attempts during swimming in marine predators (e.g. Viviant et al., 2010), pre-filtering should involve band-pass or band-stop filtering to suppress the signal component corresponding to locomotion. The latter works because the impulse-like μE is composed of several frequencies whereas locomotion is carried out in a relatively tight frequency range for a given species.

Computing signal amplitude

Raw upper and lower envelopes of the filtered acceleration signal s_{acc} are first computed using the maximum and minimum of s_{acc} in a moving window of size M_{env} samples centred at each sample of s_{acc} (Figure 1a). A low-pass filter of cut-off frequency f_c is then applied to obtain smoothed upper (U_{acc}) and lower (L_{acc}) envelopes. Finally, signal amplitude A_{acc} is computed according to Equation 1:

$$A_{acc} = \frac{U_{acc} - L_{acc}}{2}. \quad (1)$$

Effective values for M_{env} and f_c can be chosen if one considers high-speed locomotion frequency f_{hiFreq} (e.g. the frequency of running in land animals) to be indicative of the maximal rate at which successive microevents can be performed (i.e. number of microevents per second). Thus, $f_c \geq f_{hiFreq}$, with higher values being more conservative. The choice of M_{env} involves two trade-offs. On the one hand, one would aim to simplify microevent detection by maximally smoothing A_{acc} so that multiple peaks in s_{acc} corresponding to a single μE are coalesced into a single peak in A_{acc}. On the other hand, to resolve a signal of frequency f_{hiFreq}, one needs to sample at least once every $1/(2 \times f_{hiFreq})$ seconds (Nyquist-Shannon sampling theorem). This trade-off can be balanced by choosing M_{env} to be equal to $f_s/(2 \times f_{hiFreq})$ samples (rounded to nearest integer), where f_s is the accelerometer's sampling frequency.

Segmenting amplitude to find microevents

Positive peaks in A_{acc} are first found and designated as possible μE candidates (Figure 1b). The span of each μE candidate is defined as being the portion of A_{acc} starting from μE's onset to its end. To find

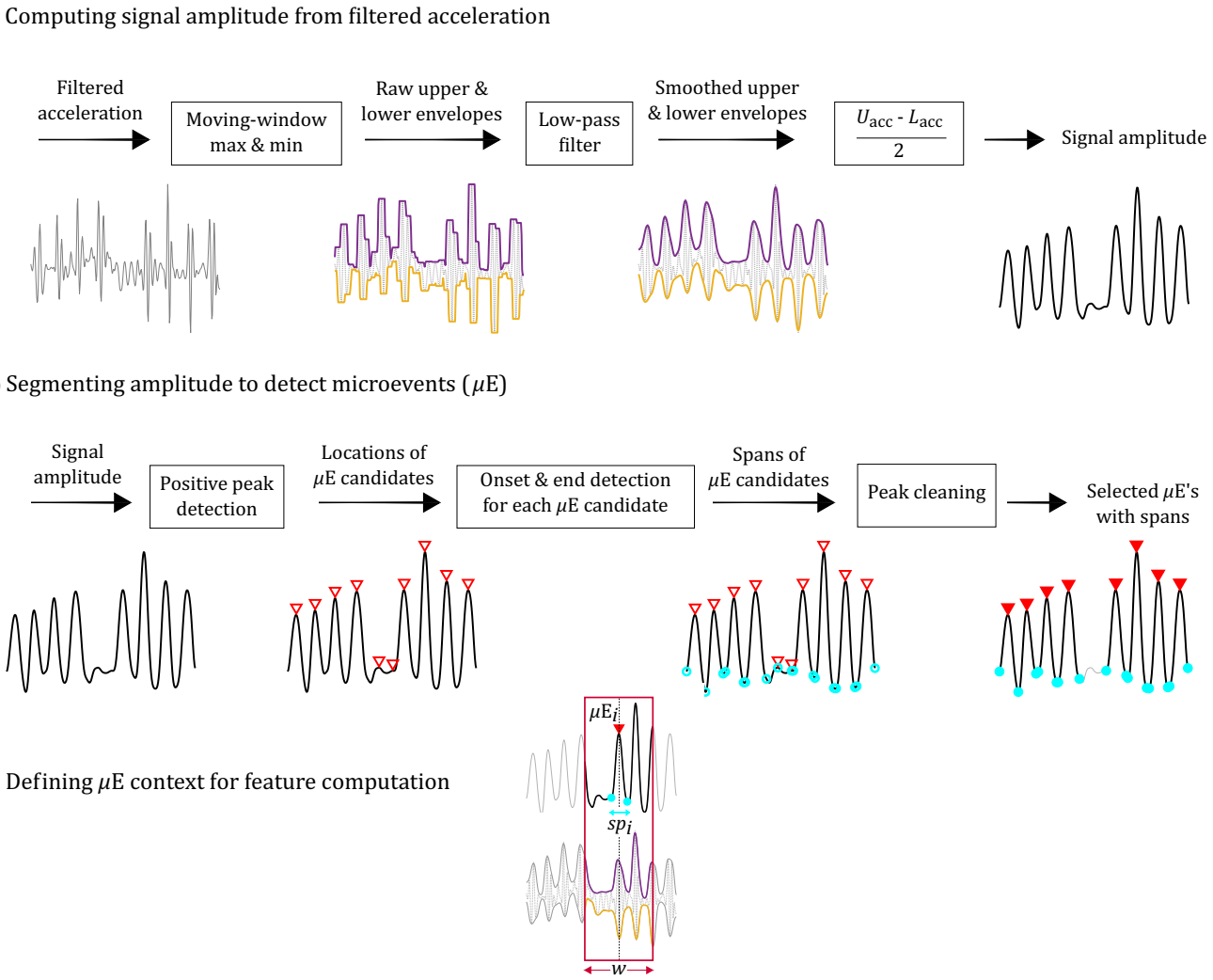

FIGURE 1 Microevent detection and characterisation. (a) Smoothed upper (U_{acc}, purple) and lower (L_{acc}, yellow) envelopes of filtered acceleration (grey) are first computed, and half of their sample-wise difference defined as signal amplitude (black), (b) signal amplitude (black) is segmented into individual microevents (solid red triangles) with spans given by each microevent's onset and end (solid cyan discs) after selecting from the pool of microevent candidates (locations: unfilled red triangles; spans: unfilled cyan circles) and removing non-event zones (grey), and (c) two neighbourhoods were considered for characterising microevents through feature computation: local, given by microevent μE_i's span (solid cyan discs), and contextual, given by a window (dark red box) of length w seconds centred at μE_i's location (solid red triangle). Features were computed from four signals per accelerometer axis: acceleration signal, its upper and lower envelopes, and amplitude

estimates of the onset and end of each μE candidate, two thresholds are used – an amplitude threshold th_{amp}, and a standard deviation threshold th_{std} computed on a moving window of size M_{seg} (can be chosen to be equal to M_{env} by default) samples – in conjunction with negative peak detection (Appendix S1). The two thresholds are used to filter out regions of A_{acc} that belong to 'non-event zones', that is, zones which cannot contain μE's of interest because A_{acc} is either nearly constant, slowly rising or slowly decaying in such regions.

2.1.2 | Characterising detected microevents using features

Both local and contextual data from each axis are used in this step to characterise each microevent μE_i (Figure 1c) through feature computation. Local data are defined as being delimited by the span sp_i of μE_i, and contextual data by a fixed window of length w seconds centred at μE_i's amplitude peak. Local data capture short time-scale manoeuvres performed by the animal during μE_i, and contextual data capture the longer time-scale, background movement during which μE_i occurred. Note that while the microevent detection step (cf. Section 2.1.1) uses data from a single axis of the accelerometer, feature computation can leverage data from all three axes. We drove the feature development process by first identifying four general feature categories: biomechanics (e.g. posture, intensity), frequency (e.g. periodicity, dominant frequency), coordination (e.g. coefficient of correlation between data from different axes) and pattern (e.g. signal tortuosity, number of zero-crossings, microevent's peak shape, statistical features). We created a non-exhaustive list of 200 candidate features across

these four broad categories (Appendix S2) based on data from filtered acceleration, upper and lower envelopes, and signal amplitude for each accelerometer axis. Feature selection is then carried out by choosing the most discriminative feature from each sub-category (e.g. using rankfeatures, © MATLAB R2019a). This approach helps achieve an order-of-magnitude reduction in the number of features – from the total number of individual features (200) to the number of sub-categories (~10) – while ensuring that meaningful features are selected.

2.1.3 | Classifying microevent types using machine learning

Robust, readily available machine learning algorithms are used to separate microevent types in feature space. Classifiers can either be used directly (e.g. Random Forest, Nathan et al., 2012) or combined in a user-defined hierarchical tree (Chakravarty, Cozzi, et al., 2019).

2.2 | Case study: Foraging in Kalahari meerkats

2.2.1 | Data collection and groundtruthing

Fieldwork was conducted as described in Chakravarty, Cozzi, et al. (2019). Data from 10 adult meerkats foraging in their natural habitat were collected at the Kalahari Meerkat Project (Clutton-Brock & Manser, 2016) from August 2016 to November 2017. The animals were habituated to a stage where it was possible to follow them within a distance of 1 m. Individuals bore collars equipped with an inertial measurement unit (IMU; adapted version of Physilog IV, GaitUp SA, Switzerland) containing a tri-axial accelerometer recording at a sampling frequency of f_s = 100 Hz/axis with a range of ± 16 g (±156.96 m/s^2) and 16-bit resolution. The size of the collar case (IMU and battery) was 35 mm × 29 mm × 19 mm, and overall weight was <25 g. Accelerometer calibration was conducted prior to each recording session according to Ferraris et al. (1995). Each recording session lasted about three continuous hours in the morning during which each individual was filmed using an electronically synchronised handheld video camera recording at 25 frames/s (Chakravarty, Maalberg, et al., 2019). Videos were annotated by playing them in Solomon Coder (version: beta 17.03.22) and noting down start and end times of behaviours of interest (in Microsoft® Excel®, 2013).

2.2.2 | Fine-scale meerkat foraging behaviours, and classification scheme

A hierarchical model to recognise coarse-scale meerkat behaviour such as vigilance, resting, foraging and running has previously been developed (Chakravarty, Cozzi, et al., 2019). Here, we aimed to expand upon this model by resolving the broad category of 'foraging' into the following four finer behaviours:

1. *Searching*: the animal walks with variable speed, often turning and pausing, with head pointing downwards, looking from side to side and palpating the ground with the front paws for suitable digging sites or prey items. This behaviour appeared to be the default state during foraging.
2. *One-armed digging* (*1AD*): after finding a suitable digging site, the animal digs using one forelimb at a time. The forelimbs either cycle back and forth alternately, or the same forelimb is used repeatedly to prod when direct access is difficult.
3. *Two-armed digging* (*2AD*): after finding a suitable digging site, the animal uses both forelimbs together to pierce the ground, scoop up sand and push it back out through the gap between its hind limbs.
4. *Chewing*: after finding and manoeuvring a prey item into its mouth, the animal proceeds to either chew it directly for small prey, or ingest it gradually for large prey by snapping with multiple, irregular jerking motions of the head.

Video clips of each behaviour are provided in Supplementary Information. Video annotation for these four behaviours served as groundtruth. We defined a three-node hierarchical classification scheme (Figure 2) to separate chewing from the other three activities at the first node, then digging (1AD and 2AD) from searching at the second node, and finally 1AD from 2AD at the third node. The rationale behind this classification structure was that firstly, chewing primarily involved head motion whereas the other three activities primarily involved limb motion. Secondly, 1AD and 2AD are both forms of 'digging' whereas searching does not involve digging.

2.2.3 | Applying the microevent identification algorithm

Axis selection and signal amplitude computation

We chose the surge axis for microevent detection. We filtered surge acceleration with a Butterworth low-pass filter of order 4 and cut-off frequency 25 Hz to remove noise. In line with the considerations in Section 2.1.2, and given that meerkat running frequency was observed to be ~4–5 Hz (Chakravarty, Cozzi, et al., 2019), we computed envelopes with M_{env} = 11 samples, and smoothed them using an M_{env} th-order FIR low-pass filter (implemented using fir1 in MATLAB R2019b) with cut-off frequency f_c = 5 Hz (see purple/yellow curves in Figure 1a for smoothed envelopes).

Amplitude segmentation

For amplitude segmentation, in line with arguments presented in Section 2.1.1, we chose an amplitude threshold of th_{amp} = 0.1 g (0.98 m/s^2), and standard deviation threshold of th_{std} = 0.01 g (0.098 m/s^2) with a moving window of size M_{seg} = 11 samples (= M_{env} by default). The choices of th_{amp} and th_{std} are conservative ones, designed to select the

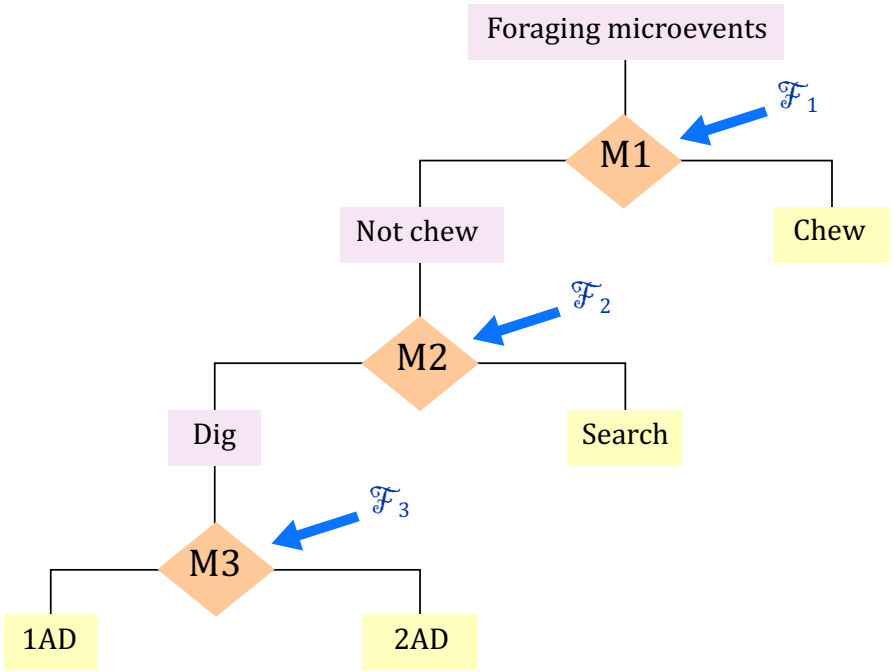

FIGURE 2 Microevent classification scheme. A hierarchical, tree-like classification scheme using one machine learning algorithm (M_i, $i \in \{1,2,3\}$) and feature set ($\mathcal{F}_i$, $i \in \{1,2,3\}$) per node was used to identify microevents occurring during fine-scale meerkat foraging behaviours. 1AD, one-armed digging; 2AD, two-armed digging

maximum number of microevents from a signal while rejecting spurious peaks and zones in the signal where short durations of static activity occur, for example, pauses between successive head jerks during chewing, or between successive bursts of 1AD.

Feature development and selection

We chose a window of size $w = 1.5$ s for feature computation to ensure that the window would be long enough to include sufficient context before and after the microevent, but short enough to capture enough data since meerkat behaviours are typically short-lived. We computed 200 features from tri-axial acceleration across nine sub-categories (cf. Section 2.1.3): posture, intensity, periodicity, dominant frequency, band-specific frequency content, microevent amplitude-peak characteristics, coordination between surge and heave, tortuosity and statistical summaries. We selected the best features from each sub-category and node (Figure 2) using five filter feature-ranking methods (Appendix S3). Further, we computed nine sets of features and compared them by varying the microevent neighbourhood (local only, contextual only, both local and contextual), and signals from which features were computed (acceleration only, envelope-based only, both acceleration- and envelope-based; Table S2). Finally, to compare results with fewer features, we computed an additional set of results by using only the features selected for the first node at all three nodes.

Classification

At each node, we tested five machine learning algorithms (names in parentheses refer to function names used to implement these algorithms in MATLAB R2019b): Naïve-Bayes (NB; implementation: fitcnb, default settings), Linear Discriminant Analysis (LDA; implementation: fitcdiscr, default settings), Logistic Regression (LR; implementation: fitcecoc with templateLinear('Learner','logistic'), 'Coding' set to 'onevsall', other settings at default values), Support Vector Machine with linear kernel (SVM; implementation: fitcecoc with templateLinear('Learner','svm'), 'Coding' set to 'onevsall', other settings at default values) and Decision Tree (DT; implementation: fitctree, default settings). Five model combinations, M1-M2-M3 (Figure 2) resulted from these upon application of the same algorithm at each node, for example, LR-LR-LR.

2.2.4 | Cross-validation

To ensure that the inevitable imbalance in the amount of data collected per individual did not bias model training, we balanced the dataset so that each behavioural class had an equal amount of data from each individual (Appendix S4). For this, we resampled the dataset through a combination of oversampling rare classes using the Synthetic Majority Oversampling Technique (SMOTE) and randomly undersampling over-represented classes (Chawla et al., 2002; Appendix S4). Thereafter, we performed leave-one-individual-out cross-validation (LOIO) on the balanced dataset. Models were trained using data pooled over all individuals except one, and tested on data from the individual left out. This process was repeated until each individual had been the 'test' individual once. To prevent specific outputs of random number generation in SMOTE from potentially biasing results, we averaged results over an ensemble of 10 balanced datasets for each cross-validation run. In this process, the 10 resulting confusion matrices for each test individual (one confusion matrix per balanced dataset) were aggregated, and 13 model performance metrics were computed as in Chakravarty, Cozzi, et al., 2019 (Appendix S5): three

behaviour-specific metrics (sensitivity, precision and specificity) for each of the four behaviours and overall model accuracy. Finally, the effect of dataset balancing was assessed by comparing LOIO results obtained with the balanced dataset with those obtained with the imbalanced dataset.

2.2.5 | Alternative classification strategies: Benchmarking against classical machine learning approaches

We benchmarked microevent identification performance against sliding window-based classification using classical machine learning (CML). We considered two machine learning algorithms (names in parentheses refer to function names used to implement these algorithms in MATLAB R2019b): Random Forest (RF; implementation: TreeBagger, 500 trees, other settings at default values), and Support Vector Machine with linear kernel (SVM; implementation: fitcecoc with templateLinear('Learner','svm'), 'Coding' set to 'onevsall', other settings at default values). With these, we considered two feature families: (a) the 38 summary statistics presented in Nathan et al., 2012 (Table S5) – such as statistical moments (*M*, *SD*), autocorrelation and trend of data from tri-axial acceleration and vectorial norm, and (b) 131 of the 200 features developed in this study (Table S2) that did not require microevent-related information. To ensure that models being compared had the same complexity in terms of number of features employed, we chose the top n_F features from each feature family (Appendix S6), where n_F would be the number of features employed by the best microevent identification model.

Classical machine learning has previously been used with small window sizes (0.3 s) to differentiate fine-scale behaviour such as prey handling from swimming in little penguins *Eudyptula minor* (Carroll et al., 2014). We included this possibility by computing benchmarking results with two window sizes: 1.5 s with 1 s overlap, and 0.3 s with 50% overlap between successive windows. We performed feature selection separately for the two window sizes (Appendix S6). Finally, we computed LOIO classification results with balanced as well as imbalanced datasets.

2.2.6 | Analysis of μEvId's sensitivity to the choice of model parameters

We evaluated μEvId's robustness to changes in the choice of model parameters. Each of the six parameters (M_{env}, f_c, th_{amp}, th_{std}, M_{seg}, w) was varied by up to ± 55% (Table 3) relative to its reference value (11, 5, 0.1, 0.01, 11, 1.5, respectively; cf. Section 2.2.3), and the results with LOIO re-computed (cf. Section 2.2.4). While one parameter was being varied, the other five were kept constant. We quantified μEvId's sensitivity to parameter change by recording the maximum deviation Δ_{max} across mean values of the 13 performance metrics relative to results obtained with reference values. The lower the value of Δ_{max} for a parameter, the higher μEvId's robustness to changes in it, and vice versa.

3 | RESULTS

3.1 | Collected data and detected microevents

Data for all four behaviours were collected for eight of ten individuals (Table 1). The dataset was imbalanced both across individuals and behaviours: (a) 55% of the total duration of recorded behaviours came from three individuals (individuals #3, #6 and #10) and (b) 50% of the duration of chewing data came from one individual (#3), and 39% of all 2AD came from one individual (#6). Recorded acceleration patterns for the microevents (μE's; Figure 3, left column) showed, as expected, multiple peaks with varying amplitudes and inter-peak

TABLE 1 Summary of groundtruth data and detected microevents. The number of seconds and detected microevents for four fine-scale meerkat foraging behaviours are shown here

Individual #	1AD Seconds ($n_{\mu E}$)	2AD Seconds ($n_{\mu E}$)	Search Seconds ($n_{\mu E}$)	Chew Seconds ($n_{\mu E}$)	Total/individual Seconds ($n_{\mu E}$)
1	21.7 (48)	11 (8)	30 (41)	6.5 (5)	69.2 (102)
2	0 (0)	34 (62)	37.2 (31)	20.4 (22)	91.6 (115)
3	60.4 (109)	61.1 (135)	64.8 (92)	327.7 (752)	514 (1,088)
4	38.7 (70)	29.9 (23)	38.4 (73)	10.8 (11)	117.8 (177)
5	74.2 (151)	32.4 (58)	53.9 (92)	52.2 (85)	212.7 (386)
6	61.9 (136)	161.5 (379)	40.5 (63)	34.7 (44)	298.6 (622)
7	27.9 (53)	0 (0)	31.6 (54)	34 (26)	93.5 (133)
8	39.9 (94)	25.1 (39)	22.3 (33)	5.6 (14)	92.9 (180)
9	68.9 (157)	21.3 (29)	23.4 (19)	88.2 (171)	201.8 (376)
10	71.2 (184)	35.3 (22)	61.2 (120)	81 (120)	248.7 (446)
Total	464.8 (1,002)	411.6 (755)	403.3 (618)	661.1 (1,250)	1,940.8 (3,625)

Abbreviations: 1AD, one-armed digging; 2AD, two-armed digging; $n_{\mu E}$, number of detected microevents.

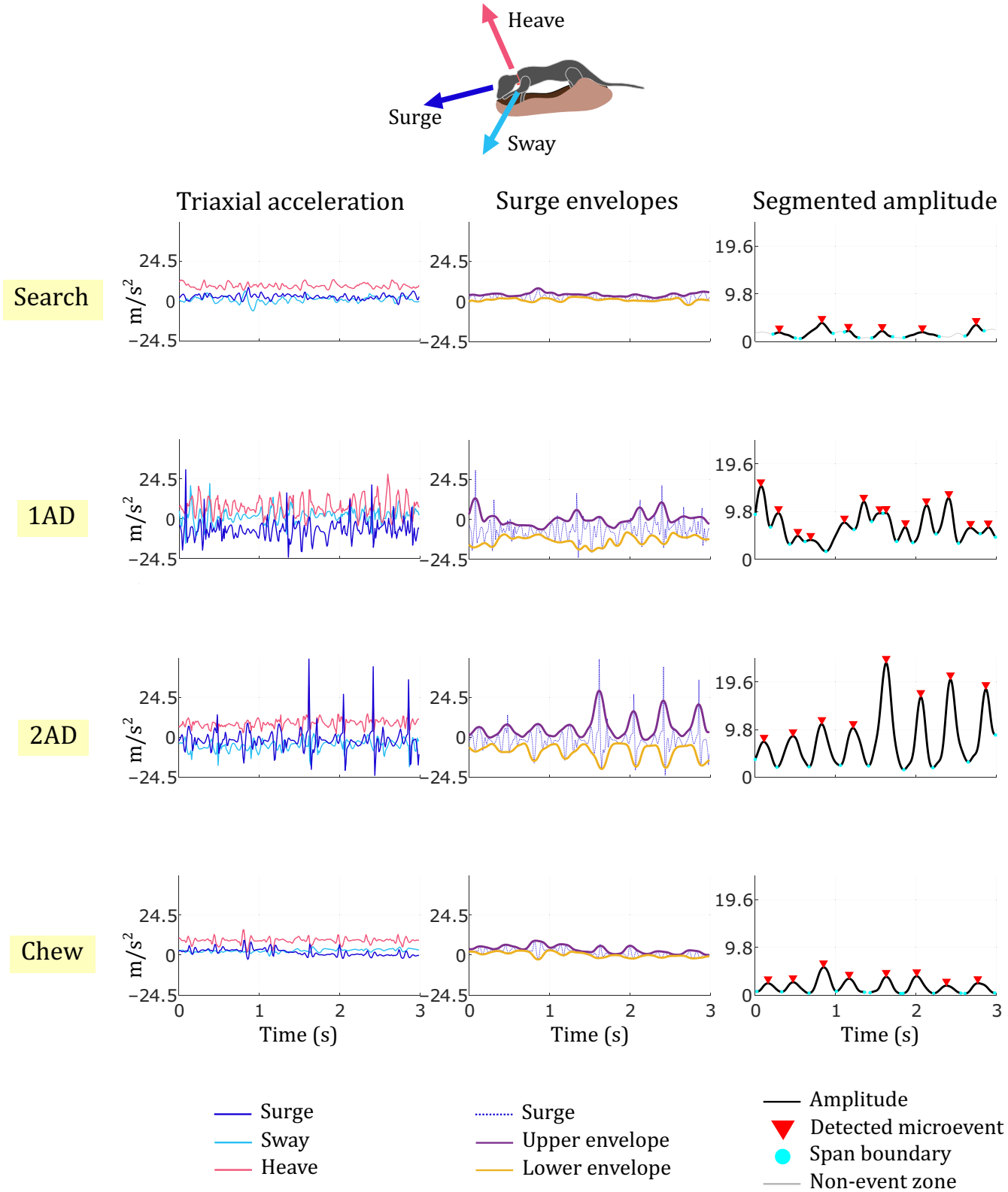

FIGURE 3 Observed signals and detected microevents. Each fine-scale meerkat foraging behaviour (rows) yielded nonstationary acceleration signals (left column). Smoothed upper and lower envelopes of surge acceleration (middle column) enabled computation of signal amplitude, which was segmented to detect microevents (right column). 1AD, one-armed digging; 2AD, two-armed digging

intervals. μE's lasted for <1 s: span lengths were 0.26 ± 0.08 s for 1AD, 0.30 ± 0.09 s for 2AD, 0.28 ± 0.09 s for searching, and 0.30 ± 0.07 s for chewing. 1AD was the 'busiest' activity with 2.1 ± 0.3 μE/s ($M \pm SD$ across individuals) while the rate was lower for the other activities: 1.5 ± 0.4 μE/s for searching, 1.5 ± 0.6 μE/s for 2AD, and 1.5 ± 0.6 μE/s for chewing.

3.2 | Feature selection and data resampling

A set of eight features (Table S3) selected sub-category wise for the first node (and repeated at the two remaining nodes) given as input to an LR-LR-LR model (cf. Section 2.2.3 under '*Classification*') performed the best (Table 2, first row) for microevent identification. These eight features were derived from both local and contextual acceleration-only signals (cf. Section 2.2.3 under '*Feature development and selection*').

For comparison with CML-based approaches, the top $n_F = 8$ features (cf. Section 2.2.5) from Nathan et al's. (2012) feature family (Tables S6 and S8) and those from the feature family developed in this study (Tables S7 and S9) were selected. For LOIO, three individuals had to be left out because of insufficient data for at least one of the four behaviours (Table 1): individuals #2 and #7 did not have data from 1AD and 2AD, respectively, and individual #1 did not have enough chewing samples (SMOTE requires at least six samples from each class). Thereafter, the dataset with data from the seven retained individuals was resampled (cf. Section 2.2.4) and used for LOIO.

3.3 | Performance evaluation and comparison

For microevent identification (μEvId), means of all behaviour-specific performance metrics with μEvId were ≥80%, and mean overall accuracy across individuals was >85% (Table 2). Performance was robust across individuals despite possible variations arising from differences in tag attachment and recorded acceleration magnitudes between individuals, as well as those arising from differences in digging style, sand compactness, digging site and prey size. μEvId performed particularly well for chewing and 2AD, with high accuracy (means of all three behaviour-specific metrics >89% for chewing, and >86% for 2AD) and low standard deviation across individuals (<6% for chewing and <9% for 2AD across all three behaviour-specific performance metrics). Large (>10%) standard deviations were observed for 1AD precision and Search sensitivity. This was primarily due to misclassification of Search as 1AD for a single individual (#4, Table S10). Re-examination of video footage revealed that most of this individual's (#4, Table 1) annotated searching behaviour (68%) actually came from brisk walking/trotting – though #4 looked briefly at the ground while doing so, it seemed unclear whether it was actually searching for food or simply catching up with its group. Thus, the errors obtained for this individual are more a consequence of imperfect annotation rather than limitations of the model. 1AD precision improved to 86.7 ± 7.4% and searching sensitivity to 87.6 ± 7% when this individual's metrics were excluded.

With balanced datasets, μEvId's overall accuracy exceeded that of the best CML model by 8.6% (Table 2). For each CML model, there was at least one behaviour-specific performance metric that was <70%. Further, μEvId was found to be more robust to inter-individual variation: standard deviations were >10% for only two of 13 performance metrics with μEvId, as opposed to eight of 13 performance metrics with the best CML model. For CML models, results using a

TABLE 2 Results with leave-one-individual-out cross-validation. The performance of microevent identification (eight features with logistic regression) was compared with classical machine learning (CML) approaches employing the same number of features. For CML, results are shown for the best-performing combinations of moving-window size (w = 1.5 s and w = 0.3 s) and feature family (*fF1* and *fF2*: sets of top eight features from feature family in Nathan et al., 2012 and this study, respectively) for two different machine learning algorithms (RF: Random Forest; SVM: Support Vector Machine, linear kernel)

		1AD			2AD			Search			Chew			Overall
Model		Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Accuracy (%)
Microevent identification		83.5 ± 7.8	93.8 ± 4.8	83.0 ± 11.4	90.6 ± 8.7	95.0 ± 2.8	86.4 ± 6.2	80.0 ± 19.5	94.7 ± 4.0	84.3 ± 8.7	89.2 ± 5.7	97.6 ± 1.5	92.6 ± 4.3	85.8 ± 5.3
Classical machine learning														
RF	w = 1.5 s, *fF2*	84.4 ± 9.6	92.1 ± 5.7	79.5 ± 11.2	77.8 ± 16.8	91.4 ± 13.6	82.9 ± 16.9	75.4 ± 21.6	92.7 ± 3.7	78.5 ± 8.6	66.7 ± 22.1	91.9 ± 6.7	75.2 ± 12.0	76.1 ± 10.0
SVM	w = 1.5 s, *fF1*	77.7 ± 14.2	93.8 ± 4.9	82.2 ± 10.3	68.6 ± 25.2	89.3 ± 7.9	73.6 ± 13.6	75.9 ± 17.3	91.0 ± 10.3	80.5 ± 17.7	86.6 ± 11.7	95.5 ± 3.8	87.7 ± 9.1	77.2 ± 6.2

TABLE 3 Effect of varying microevent identification model parameters on classification performance. The six parameters used in the microevent identification algorithm were varied and the maximum drop in mean values across thirteen performance metrics – sensitivity, specificity and precision for each of the four behaviours, and overall accuracy – during leave-one-individual-out cross validation used to provide a measure of model robustness to parameter change

Parameter	Description	Reference value	Sweeping range	Units	Max. performance drop (%)
M_{env}	Window size for envelope computation	11	5–17 (±55%)	Samples (@100 Hz)	6.1
f_c	Low-pass cut-off frequency for smoothing envelope	5	2.5–7.5 (±50%)	Hz	2
th_{amp}	Amplitude threshold for segmentation algorithm	0.1	0.05–0.15 (±50%)	g (1g = 9.81 m/s^2)	3.6
th_{std}	Standard deviation (std) threshold for segmentation algorithm	0.01	0.005–0.015 (±50%)	g (1g = 9.81 m/s^2)	0.5
M_{seg}	Window size to compute std in segmentation algorithm	11	5–17 (±55%)	Samples (@100 Hz)	1
w	Window size describing context around main event	1.5	1–2 (±33%)	Seconds	5.7

longer window (w = 1.5 s) were consistently better than those with a short window (w = 0.3 s) regardless of feature family or machine learning algorithm employed (Table S12). In the absence of data balancing, none of the models yielded a mean overall accuracy of >80% (Table S13).

μEvId was found to be robust to large (≥±33%) changes in each of the six model parameters (Table 3); detailed results are provided in Appendix S8. The largest sensitivity was to changes in M_{env}, with Δ_{max} occurring at M_{env} = 5 samples, and performance dropping again when M_{env} > 11 samples (Table S14). This justified the reasoning behind our choice of v = 11 samples (cf. Section 2.2.3). Sensitivity to changes in w was also relatively high, with results generally improving as w increased (Table S19), suggesting that features incorporating greater contextual information were more discriminative. This, along with low sensitivity to changes in amplitude segmentation-related parameters (th_{amp}, th_{std} and M_{seg}), suggests that the choices of window size for envelope computation (M_{env}) and context definition (w) are most critical to microevent identification performance.

4 | DISCUSSION

We modelled brief, impulsive movements (microevents) characteristic of fine-scale animal behaviour as impulses acting on the animal body-and-sensor system, and used signal enveloping to detect the resulting transient impulse responses recorded in the acceleration signal. We then used robust machine learning employing meaningful features to automate the identification of microevents. Using data collected on 10 foraging meerkats in their natural habitat, we demonstrated the efficacy of our method in identifying fine-scale foraging behaviour. We showed that microevent identification (μEvId) using eight features and logistic regression in a hierarchical classification scheme outperformed classical machine learning (CML) approaches in a variety of cases: different feature families, window sizes and machine learning algorithms. μEvId was found to be robust not only to inter-individual variation but also to large changes in model parameters.

4.1 | Advantages of the enveloping operation

Two advantages of enveloping can be observed visually: (a) by effectively 'sandwiching' filtered acceleration between the upper and lower envelopes, enveloping makes μE detection robust to changes in signal mean over time, such as during chewing (Figure 3, last row), and (b) smoothing during enveloping allows μE detection even when microevent-generated signals differ slightly in duration, highly (2×) in amplitude, and in number of constituent peaks, as in 2AD (Figure 3, third row from top).

4.2 | Microevent detection improves class separability by rejecting noninformative regions of the signal

Results from a previous study indicated that there was low separability between different meerkat foraging behaviours in terms of posture, intensity and periodicity (Chakravarty, Cozzi, et al., 2019). Results with CML in this study, too, indicate that even when the number of features is increased (from three in Chakravarty, Cozzi, et al., 2019 to eight in the present study), separability between foraging behaviours remains relatively low. Yet, there is a large difference in performance between μEvId and CML which is not explained by typical factors that influence behaviour recognition performance, such as feature selection, machine learning algorithm, class imbalance and classification scheme. Despite testing several combinations of selected features (two feature families, two window sizes) and machine learning algorithms (Random

Forest, Support Vector Machine), the mean overall accuracy of the best CML model was still lower than that of μEvId by 8.6% (with some overlap in standard deviation). Balancing the amount of data from different individuals and classes prior to cross-validation removed differences arising from class imbalance in both cases (μEvId, CML). Class balancing also nullified the greater ease with which hierarchical classification schemes (e.g. μEvId) can deal with class imbalance compared to directly classifying all behaviours at once (CML; Chakravarty, Cozzi, et al., 2019). Therefore, since no major differences are encountered during the classification process, the microevent detection step conducted prior to classification in μEvId must be driving its improved class separability.

A fundamental methodological difference between the two approaches is that in μEvId, windows (for feature computation) are defined only where microevents are detected ('informative' regions of the signal). Thus, the microevent detection step automatically rejects regions of the signal consisting of low-acceleration movements or smooth changes in acceleration uncharacteristic of microevents ('non-informative' regions; light grey parts of the curve in Figure 1b, right-most plot, and in Figure 3, top-right plot). Further, centring the window at the detected microevent location provides the requisite signal context (before, after) within which the μE's signal occurs. Features are thus computed only from informative regions of the signal since non-informative regions are filtered out. In contrast, in CML, a sliding window does not distinguish between informative and noninformative regions of the signal and, consequently, fails to choose the most discriminative features, leading to poorer class separation. For instance, *cxFFTpeakpowerHeave* (periodicity of heave acceleration) and *cxFFTpowerfracHeave* (fraction of heave signal in frequency range 2.5–5 Hz; Table S3) were both higher for two-armed digging than for the other behaviours. Their selection reflects the visual observation that the trunk of a meerkat engaged in two-armed digging moves rhythmically – within a relatively tight frequency range – downward as the meerkat reaches for more sand to dig out, and upward as it flings dug-out sand backward. Neither of these features were selected by any of the CML models.

4.3 | Limitations

There are two main limitations to the proposed method. First, though μEvId does not impose restrictions on the duration of signal patterns of interest, it does require all patterns to have similar duration, for example, <1 s for μE's. Thus, it cannot be used to identify signal patterns across time-scales, that is, patterns whose durations differ by an order of magnitude, say 1–10 s. This is because envelope computation in the μE detection step requires window size (M_{env}, Table 3) to be fixed: short windows would not be long enough to capture the long patterns of interest, whereas long windows would club together shorter patterns of interest within the same window. Available methods to identify patterns across scales involve signal characterisation by hand (Wilson, Neate, et al., 2018). Second, μEvId is designed to identify shock-like acceleration patterns recorded during short-lived behaviours which, by definition, requires sampling at high frequency. Thus, it may not be possible to apply μEvId to data collected during long-term recordings, where it is common to sample at low frequency (<10 Hz). However, future developments in on-board data processing may allow devices to record at high frequency, apply μEvId in real time, and store only the labels of classified behaviour—this would reduce both the power consumption involved in writing to memory, as well as the amount of storage memory required.

4.4 | Potential for other studies, and future work

The fact that a majority of μEvId's selected features (Table S3) were based on contextual acceleration (Figure 1c) suggests that including data from a time window longer than the signal pattern of interest gives more accurate and robust results than characterising the signal pattern alone. This finding could be leveraged to define robust guidelines for manual waveform characterisation in semi-automated approaches (Wilson, Neate, et al., 2018). In our study, accurate identification of fine-scale behaviour was achieved with a single sensor. This shows that algorithmic advances could possibly be helpful in simplifying tag attachment issues in other studies, for example, by reducing the number of sensors needed to identify rapid head movements during prey capture in Adélie penguins *Pygoscelis adeliae* (Watanabe & Takahashi, 2013) from two (on head and back) to one. Our specificity analysis, that is, the accuracy with which behaviours can be rejected when they are not present, showed that μEvId can accurately separate different types of confounding μE's. High specificity is desirable since misclassifying confounding fine-scale behaviours can alter biological interpretation, for example, detection of prey biting motions in Weddell seals *Leptonychotes weddellii* may be biased by head bobbing associated with vocalisation (Naito et al., 2010). It could be possible to combine our approach with existing approaches: here, sequences of microevents (identified using μEvId) could be combined using Boolean classification (Wilson, Neate, et al., 2018) to identify higher-level, biologically relevant behaviours (possibly composed of multiple types of microevents) with high accuracy. Information on bouts of higher-level behaviour could be useful for other purposes as well. For instance, identification of successive chewing μE's can be used to first define a bout of the higher-level behaviour, 'prey ingestion'. The duration of prey ingestion and feature values of its constituent μE's can then be combined to predict prey size.

5 | CONCLUSIONS

We showed that robust, automated recognition of fine-scale animal behaviours from recorded acceleration can be achieved by conceptualising fine-scale behaviour as being composed of characteristic microevents. Our findings demonstrate that transient acceleration patterns recorded during microevents can be modelled as impulse responses of the animal body-and-sensor system. Different

microevents generate different impulse responses and are carried out in different signal contexts. In our microevent identification model, enveloping detects where impulse responses occur in the signal ('seeking'), and machine learning employing features that characterise both impulse response and signal context separates microevents ('learning'). The microevent detection step acts as a prefilter that selects informative, microevent-generated regions of the signal while rejecting noninformative regions uncharacteristic of microevents. This prefiltering results in the selection of discriminative features that reflect biomechanical differences in the movements carried out by the animal during a microevent. The resulting method was able to identify microevents with high accuracy, and robustness to inter-individual variation. The classical approach of sliding-window-based machine learning cannot separate informative and noninformative regions of the signal, and was found to be suboptimal for capturing the nonstationarity of signals recorded during fine-scale animal behaviour. Microevent identification, together with existing models that can identify broad behavioural categories, provides future studies with a powerful toolkit to exploit the full potential of acceleration data for animal behaviour recognition.

ACKNOWLEDGEMENTS

We thank the Northern Cape Conservation Authority for permission to conduct this research, and the Kalahari Research Trust and Tim Clutton-Brock as director of the Kalahari Meerkat Project for access to facilities and habituated animals in the Kuruman River Reserve, South Africa. This project was supported by the Swiss National Science Foundation (research grant no. CR32I3_159743). During the period of our project, the long-term research on meerkats was supported by a European Research Council Advanced Grant (No. 294494) to Tim Clutton-Brock, by grants from the University of Zurich to Marta Manser, and the MAVA foundation. We thank the field managers, collaborators and assistants for facilitating field work and helping with data collection, and in particular David Gaynor, Tim Vink, Ana Morales González and Héctor Ruiz Villar. We thank Nino Maag for fruitful discussions on defining foraging behaviours, Selin Ersoy for extensive help with video annotation, and Pascal Morel for adapting the Physilog IV for this study. We thank Maiki Maalberg for sub-routines to perform cross-validation in MATLAB, and Jean-Marc Vesin for MATLAB code for envelope computation.

AUTHORS' CONTRIBUTIONS

K.A., P.C. and G.C. developed the research idea, and M.M. and A.O. contributed to refinements; P.C. and G.C. supervised fieldwork; P.C., P.-A.L. and H.D. performed data analyses; P.C. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT

Data deposited in the Dryad Digital Repository https://doi.org/10.5061/dryad.5mkkwh742 (Chakravarty et al., 2020).

ORCID

British Chakravarty https://orcid.org/0000-0002-2975-6253
Gabriele Cozzi https://orcid.org/0000-0002-1744-1940
Marta Manser https://orcid.org/0000-0001-8787-5667
Arpat Ozgul https://orcid.org/0000-0001-7477-2642
Kamiar Aminian https://orcid.org/0000-0002-6582-5375

REFERENCES

Béliveau, A., Spencer, G. T., Thomas, K. A., & Roberson, S. L. (1999). Evaluation of MEMS capacitive accelerometers. *IEEE Design & Test of Computers*, *16*(4), 48–56. https://doi.org/10.1109/54.808209

Bourke, A. K., O'Brien, J. V., & Lyons, G. M. (2007). Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait & Posture*, *26*(2), 194–199. https://doi.org/10.1016/j.gaitpost.2006.09.012

Brown, D. D., Kays, R., Wikelski, M., Wilson, R., & Klimley, A. P. (2013). Observing the unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry*, *1*(1), 20. https://doi.org/10.1186/2050-3385-1-20

Camomilla, V., Dumas, R., & Cappozzo, A. (2017). Human movement analysis: The soft tissue artefact issue. *Journal of Biomechanics*, *62*, 1–4. https://doi.org/10.1016/j.jbiomech.2017.09.001

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*(24), 4295–4302. https://doi.org/10.1242/jeb.113076

Chakravarty, P., Cozzi, G., Dejnabadi, H., Léziart, P.-A., Manser, M., Ozgul, A., & Aminian, K. (2020). Data from: Seek and learn: Automated identification of microevents in animal behaviour using envelopes of acceleration data and machine learning. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.5mkkwh742

Chakravarty, P., Cozzi, G., Ozgul, A., & Aminian, K. (2019). A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods in Ecology and Evolution*, *10*(6), 802–814. https://doi.org/10.1111/2041-210X.13172

Chakravarty, P., Maalberg, M., Cozzi, G., Ozgul, A., & Aminian, K. (2019). Behavioural compass: Animal behaviour recognition using magnetometers. *Movement Ecology*, *7*(1), 28. https://doi.org/10.1186/s40462-019-0172-6

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Clutton-Brock, T. H., & Manser, M. (2016). Meerkats: Cooperative breeding in the Kalahari. *Cooperative Breeding in Vertebrates*, *294*, 317.

Dawkins, M. S. (1989). Time budgets in red junglefowl as a baseline for the assessment of welfare in domestic fowl. *Applied Animal Behaviour Science*, *24*(1), 77–80. https://doi.org/10.1016/0168-1591(89)90126-3

Esseiva, J., Caon, J., Mugellini, E., Khaled, O. A., & Aminian, K. (2018). Feet fidgeting detection based on accelerometers using decision tree learning and gradient boosting. In *Proceedings of the International Conference on Bioinformatics and Biomedical Engineering*, 15–17 September 2018 (pp. 75–84). Springer.

Ferraris, F., Grimaldi, U., & Parvis, M. (1995). Procedure for effortless in-field calibration of three-axis rate gyros and accelerometers. *Sensors and Materials*, *7*, 311.

Hammond, T. T., Springthorpe, D., Walsh, R. E., & Berg-Kirkpatrick, T. (2016). Using accelerometers to remotely and automatically characterize behavior in small animals. *Journal of Experimental Biology*, *219*(11), 1618–1624. https://doi.org/10.1242/jeb.136135

Hansson, G. A., Asterland, P., Holmer, N. G., & Skerfving, S. (2001). Validity and reliability of tri-axial accelerometers for inclinometry in posture analysis. *Medical and Biological Engineering and Computing*, *39*(4), 405–413. https://doi.org/10.1007/BF02345361

Hariharesan, S., & Barhorst, A. A. (1999). Modelling, simulation and experimental verification of contact/impact dynamics in flexible multi-body systems. *Journal of Sound and Vibration*, *221*(4), 709–732.

Hill, A. V. (1938). The heat of shortening and the dynamic constants of muscle. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *126*(843), 136–195.

Iwata, T., Sakamoto, K. Q., Takahashi, A., Edwards, E. W., Staniland, I. J., Trathan, P. N., & Naito, Y. (2012). Using a mandible accelerometer to study fine-scale foraging behavior of free-ranging Antarctic fur seals. *Marine Mammal Science*, *28*(2), 345–357. https://doi.org/10.1111/j.1748-7692.2011.00482.x

Johnson Jr., C. R., Sethares, W. A., & Klein, A. G. (2011). *Software receiver design: Build your own digital communication system in five easy steps*. Cambridge University Press.

Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, *348*(6240), aaa2478. https://doi.org/10.1126/science.aaa2478

Koshy, C. S., Flores, P., & Lankarani, H. M. (2013). Study of the effect of contact force model on the dynamic response of mechanical systems with dry clearance joints: Computational and experimental approaches. *Nonlinear Dynamics*, *73*(1–2), 325–338. https://doi.org/10.1007/s11071-013-0787-x

Martiskainen, P., Järvinen, M., Skön, J. P., Tiirikainen, J., Kolehmainen, M., & Mononen, J. (2009). Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Applied Animal Behaviour Science*, *119*(1–2), 32–38. https://doi.org/10.1016/j.applanim.2009.03.005

McFadden, P. D., & Smith, J. D. (1984). Vibration monitoring of rolling element bearings by the high-frequency resonance technique – A review. *Tribology International*, *17*(1), 3–10. https://doi.org/10.1016/0301-679X(84)90076-8

Müller, R., & Schrader, L. (2003). A new method to measure behavioural activity levels in dairy cows. *Applied Animal Behaviour Science*, *83*(4), 247–258. https://doi.org/10.1016/S0168-1591(03)00141-2

Naito, Y., Bornemann, H., Takahashi, A., McIntyre, T., & Plötz, J. (2010). Fine-scale feeding behavior of Weddell seals revealed by a mandible accelerometer. *Polar Science*, *4*(2), 309–316. https://doi.org/10.1016/j.polar.2010.05.009

Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., & Getz, W. M. (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*(6), 986–996. https://doi.org/10.1242/jeb.058602

Pesaran, H. H., & Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, *58*(1), 17–29. https://doi.org/10.1016/S0165-1765(97)00214-0

Phillips, C. L., Parr, J. M., & Riskin, E. A. (2003). *Signals, systems, and transforms* (p. 209). Prentice Hall.

Ropert-Coudert, Y., & Wilson, R. P. (2005). Trends and perspectives in animal-attached remote sensing. *Frontiers in Ecology and the Environment*, *3*(8), 437–444. https://doi.org/10.1890/1540-9295(2005)003[0437:TAPIAR]2.0.CO;2

Sakamoto, K. Q., Sato, K., Ishizuka, M., Watanuki, Y., Takahashi, A., Daunt, F., & Wanless, S. (2009). Can ethograms be automatically generated using body acceleration data from free-ranging birds? *PLoS One*, *4*(4), e5379. https://doi.org/10.1371/journal.pone.0005379

Suzuki, I., Naito, Y., Folkow, L. P., Miyazaki, N., & Blix, A. S. (2009). Validation of a device for accurate timing of feeding events in marine animals. *Polar Biology*, *32*(4), 667–671. https://doi.org/10.1007/s00300-009-0596-3

Tandon, N., & Choudhury, A. (1999). A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology International*, *32*(8), 469–480. https://doi.org/10.1016/S0301-679X(99)00077-8

Usui, S., & Amidror, I. (1982). Digital low-pass differentiation for biological signal processing. *IEEE Transactions on Biomedical Engineering*, *10*, 686–693. https://doi.org/10.1109/TBME.1982.324861

Viviant, M., Trites, A. W., Rosen, D. A., Monestiez, P., & Guinet, C. (2010). Prey capture attempts can be detected in Steller sea lions and other marine predators using accelerometers. *Polar Biology*, *33*(5), 713–719. https://doi.org/10.1007/s00300-009-0750-y

Watanabe, Y. Y., & Takahashi, A. (2013). Linking animal-borne video to accelerometers reveals prey capture variability. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(6), 2199–2204. https://doi.org/10.1073/pnas.1216244110

Wilson, R. P., Holton, M. D., di Virgilio, A., Williams, H., Shepard, E. L., Lambertucci, S., Quintana, F., Sala, J. E., Balaji, B., Lee, E. S., Srivastava, M., Scantlebury, D. M., & Srivastava, M. (2018). Give the machine a hand: A Boolean time-based decision-tree template for rapidly finding animal behaviours in multisensor data. *Methods in Ecology and Evolution*, *9*(11), 2206–2215. https://doi.org/10.1111/2041-210X.13069

Wilson, R. P., Neate, A., Holton, M. D., Shepard, E. L. C., Scantlebury, D. M., Lambertucci, S. A., di Virgilio, A., Crooks, E., Mulvenna, C., & Marks, N. (2018). Luck in food finding affects individual performance and population trajectories. *Current Biology*, *28*(23), 3871–3877. https://doi.org/10.1016/j.cub.2018.10.034

Wilson, R. P., White, C. R., Quintana, F., Halsey, L. G., Liebsch, N., Martin, G. R., & Butler, P. J. (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*(5), 1081–1090. https://doi.org/10.1111/j.1365-2656.2006.01127.x

Zegers, D. A. (1981). Time budgets of Wyoming ground squirrels, *Spermophilus elegans*. *Great Basin Naturalist*, *41*, 221–228.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Chakravarty P, Cozzi G, Dejnabadi H, et al. Seek and learn: Automated identification of microevents in animal behaviour using envelopes of acceleration data and machine learning. *Methods Ecol Evol*. 2020;00:1–13. https://doi.org/10.1111/2041-210X.13491