

METHODOLOGY

Open Access



Simulation-based validation of activity logger data for animal behavior studies

Jiawei Chen^{1*} , Geoffrey Brown¹ and Adam Fudickar²

Abstract

Bio-loggers are widely used for studying the movement and behavior of animals. However, some sensors provide more data than is practical to store given experiment or bio-logger design constraints. One approach for overcoming this limitation is to utilize data collection strategies, such as non-continuous recording or data summarization that may record data more efficiently, but need to be validated for correctness. In this paper we address two fundamental questions—how can researchers determine suitable parameters and behaviors for bio-logger sensors, and how do they validate their choices? We present a methodology that uses software-based simulation of bio-loggers to validate various data collection strategies using recorded data and synchronized, annotated video. The use of simulation allows for fast and repeatable tests, which facilitates the validation of data collection methods as well as the configuration of bio-loggers in preparation for experiments. We demonstrate this methodology using accelerometer loggers for recording the activity of the small songbird *Junco hyemalis hyemalis*.

Keywords: Simulation, Validation, Activity, Detection, Logger, Video, Sampling, Summarization, Compression, Bird

Background

Annually, animals across the planet make regional to long-distance movements that result in the transport of billions of tons of biomass [1]. Because of their impact on the daily and seasonal dynamic nature of ecosystems, animal movements and especially movement patterns are of great research interest [2]. Bio-logging is useful when direct observation of animals is impractical, or when continuous monitoring over long time periods is desired [3–7]. Due to the potentially large timescales involved in such studies, the direct storage of unprocessed sensor data places high demands on energy and memory, potentially beyond the limits of practicality [8–10]. Various data collection strategies can be employed to reduce resource consumption, but they must be validated to ensure they do not affect the validity of the data.

In this paper we describe a methodology that enables researchers to combine "raw" bio-logger data with video to determine the impact of such strategies, and by extension, visualize the relationship between video and sensor logs of animal behavior. Our bio-logger validation procedure involves collecting continuous, uncompressed sensor data and synchronized video, simulating bio-loggers in software using the recorded sensor data, and evaluating their ability to detect movements shown on video. The benefit of simulation over purely empirical testing is that it allows for faster, more repeatable tests that make more effective use of experiment data, which may be especially attractive to studies involving non-captive animals.

We have also developed a software application, QValiData, to facilitate the experimental procedure by synchronizing video, assisting with video analysis, simultaneously playing back synchronized video and data tracks, and running bio-logger simulations. With the proposed methodology and software tool, we hope to increase confidence in the reliability of bio-loggers that utilize sampling and summarization and improve the

*Correspondence: jiaweichen@alumni.iu.edu

¹ School of Informatics, Computing and Engineering, Indiana University, Bloomington, IN 47405, USA

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

efficiency of validation experiments aimed towards such bio-loggers. We demonstrate the use of our methodology with accelerometer data collected from captive Dark-eyed Juncos (*Junco hyemalis hyemalis*) to validate the summarization strategy in preparation for future activity logging with both captive and free-ranging animals.

Challenges of long-term bio-logging

Despite technological advances, power and memory still impose limits on the endurance of bio-loggers and data quality in long-term experiments. In particular, mass limitations may restrict logger energy budgets, and as a result, the amount of data that can be collected [11]. Excessive mass restricts the range of species for which such loggers are usable, since it may influence animal behavior [5, 12, 13]. For instance, it is common practice in bird studies to limit logger mass to 3–5% of the animal’s body mass to mitigate this possibility [11].

High data rates require increased processor activity, resulting in increased energy usage. In addition, increased storage requirements may necessitate the use of additional storage media, such as external flash memory, which further increases energy demand. Since a battery’s mass is roughly proportional to its capacity (Table 1), increasing battery capacity will inevitably increase total logger mass. For bio-loggers approaching the mass limit of a particular species, the energy and

memory budgets imposed by logger mass may preclude continuous high-speed recording altogether [4, 8, 9, 14].

When runtime extensions cannot be made through hardware modification, the standard approaches to overcoming logger limitations are to rate limit data collection (sampling) [16] and/or store only data summaries (summarization) [8, 10, 14, 17].

Sampling

Sampling (Fig. 1) involves recording full-resolution data in short bursts. Many studies already make use of loggers that utilize sampling at fixed intervals, or “synchronous” sampling [18–20]; however, this method may miss events that occur between sampling periods and will additionally record periods of inactivity. An improvement upon this method is “asynchronous” sampling, which only records when activity of interest is detected by sensors [21]. This increases the likelihood of recording desired movements and more effectively utilizes both energy and storage. Although our methodology can be used to validate either sampling method, asynchronous sampling presents a more interesting use case for the simulation component.

Asynchronous sampling is capable of recording the dynamic aspects of individual movement bouts, which makes it suitable in studies where the movements themselves are of interest [3, 16, 21–23]. Since recording only occurs when a movement of interest is detected, activity-based sampling can lead to significant efficiency improvements when movements of interest are sparse. However, this recording method sacrifices continuity, which may lead to a loss of context, and additionally may miss activity that is necessary for collecting more general parameters such as total energy expenditure [24, 25].

Table 1 Battery energy density of Seiko MS-series rechargeable batteries [15]

Capacity (mAh)	Nominal voltage (V)	Energy storage (Joules)	Mass (g)	Energy density (Joules/g)
3.4	2.5	30.6	0.17	180
5	2.5	45.0	0.23	196
11	2.5	99.0	0.47	211

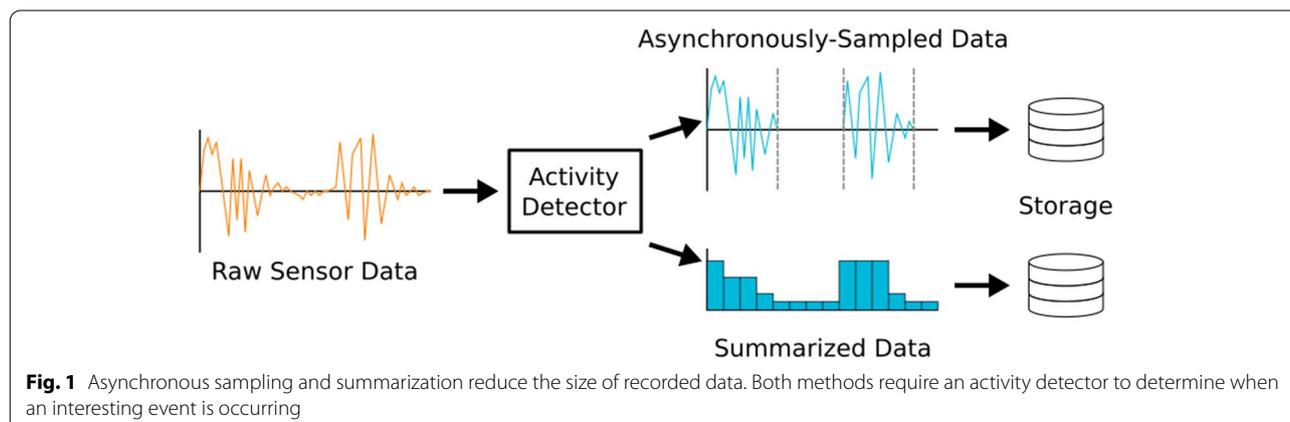


Fig. 1 Asynchronous sampling and summarization reduce the size of recorded data. Both methods require an activity detector to determine when an interesting event is occurring

Summarization

If continuous recording is desired but bio-logger design or experiment requirements preclude full-resolution recording, sensor data can be analyzed on-board the logger and observations extracted, in a process known as “summarization” [10] (Fig. 1). For instance, movement data can be summarized as a numerical value corresponding to frequencies detected within movement data or the level of activity detected [8] (“characteristic” summarization), or even a simple binary value or count representing the presence or absence of particular movements of interest (“behavioral” summarization).

Although summarized data are unable to record the unique dynamics of individual movement bouts, it can provide insight into characteristics and trends of an animal’s activity over an extended period of time [8, 10, 24, 25]. In addition, sampling bio-loggers can be used to augment ethogram studies once behaviors are characterized. Currently, many bio-logging studies focus on developing models to classify various behavior types [6, 19, 20, 26–29]. If further study or quantification of these behaviors is desired, the incorporation of these models into summarization-type bio-loggers can be used to count the occurrences of specific behaviors over long periods of time, allowing for improved recording times in comparison to conventional, continuously sampling loggers.

Activity detection

In either method, there arises the need to determine which portions of data are worth recording, and how to implement an activity detector to reliably differentiate these from the remainder of the data in real-time (Fig. 1). To maximize storage and energy efficiency, activity detection must be sensitive enough to detect all interesting events but also selective enough to avoid recording unnecessarily. In addition, activities that are detected must be recorded at a satisfactory level of detail for a given experiment’s objectives. Since on-board activity detection methods operate unsupervised, and unrecorded data are unrecoverable [3, 6, 10], it is impossible to ascertain their correctness or completeness from recorded data alone. As a result, recording strategies that employ activity detection leave open two questions: do they accurately reflect “raw” sensor data, and what animal behaviors can we infer from these data? We seek to address the first question by validating recording strategies in advance and discuss possible approaches to the second.

Ensuring validity of data

Activity detection methods can be validated in controlled environments, in which the animal under study, with a

logger attached, is closely observed while performing motions of interest that are representative of behavior similar to those exhibited in the field [4, 9]. In addition to the logger, at least one other independent, synchronized source of movement data is required to cross-check the data produced by activity detection. For instance, some experiments derive this validation source from time-stamped direct visual observations or video recordings [3, 4, 9, 19–23, 26, 29–31].

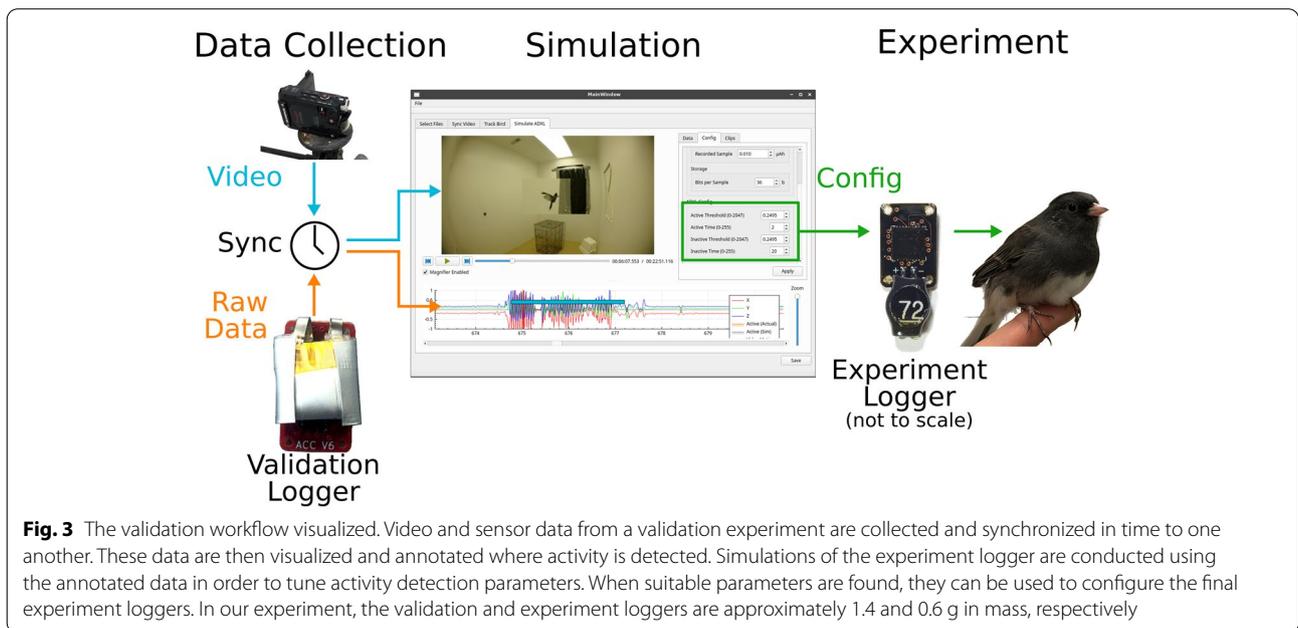
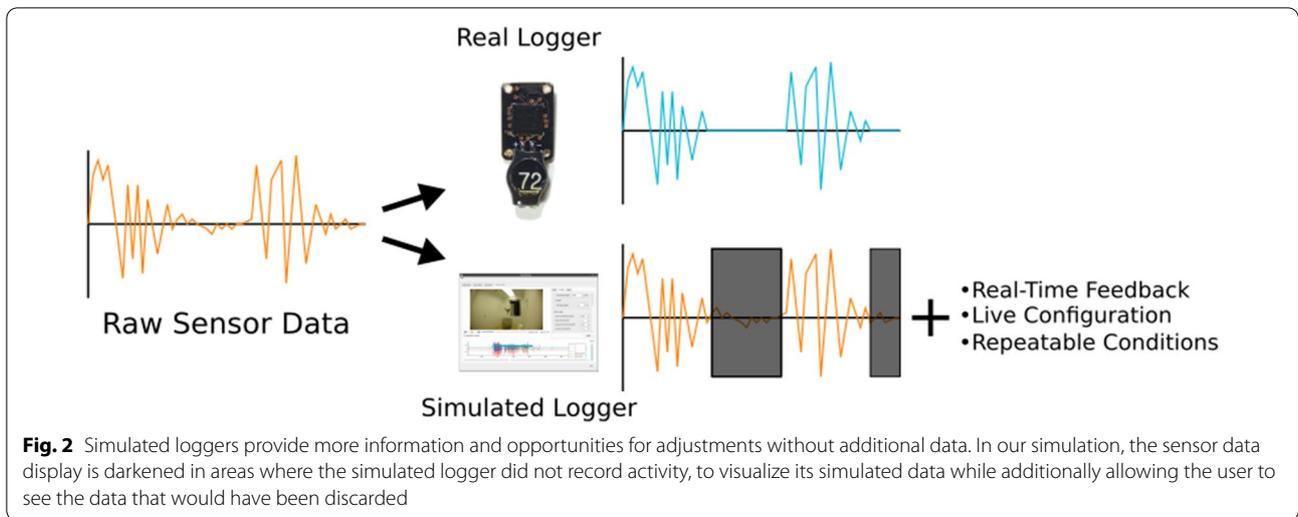
Regardless of observation method, these validation experiments serve to associate motions recorded by the logger with known behaviors detected through an independent means. When particular motions of interest are found, the corresponding sensor signatures can then be characterized to develop a model that can automatically classify future events without human intervention. This model can then be applied to loggers for data collection experiments in the field to perform data analysis and compression *in situ*.

One of the issues with experimental validation is that, since the logger under test often cannot be re-configured once deployed, making adjustments to the logger’s configuration requires a completely new trial. Furthermore, since loggers are often the only source of observation and they discard large portions of this data, incorrect classifications (such as missed events or false positives) may be difficult to diagnose or replicate. Additionally, since animal movement is often variable [32], the types and intensities of certain movements vary across trials and individuals. As a result, many trials are needed to fine-tune activity detection parameters, and the effects of incremental improvements are difficult to quantify.

Methods

Our validation procedure consists of gathering observations of an animal’s movements, associating those observations with raw sensor data of the movements, and then running a series of simulations using the recorded sensor data (Fig. 2) to develop and evaluate the performance of activity detection methods. When a suitable configuration has been found, it is then applied to loggers in an actual experiment. We have developed a software application, QValiData, to manage the data generated by these trials, and to assist in analyzing video, performing video magnification, and simulating bio-loggers. The validation workflow is illustrated in (Fig. 3).

Movement data in our study were collected from a small sparrow, the Dark-eyed Junco. Juncos are common across North America and thrive in captivity, making them ideal for the current study. Since some juncos undergo seasonal migration and in turn exhibit seasonal variability in activity, we were primarily interested in using activity loggers to track trends in their



general activity levels leading up to and during migration periods [33]. Thus, we used the validation procedure to detect periods of significant activity, which can then be applied to summarizing loggers. Although our application was originally targeted towards accelerometer-based activity loggers on small songbirds, this procedure is applicable to other species as well as other kinds of movement loggers, accommodating the wide variety of sensor types currently employed in movement studies [8, 15].

Data collection

The purpose of the data collection phase is to obtain as many examples of desirable activities as possible in a short span of time, in order to capture detailed dynamics of an animal’s movements for the simulation. To obtain continuous, raw sensor data, we developed a custom “validation logger” (Fig. 3) that continuously recorded full-resolution sensor samples at a high rate, at the cost of significantly reduced run time on the order of 100 h, as opposed to several months. Since individual data

collection runs were designed to be less than an hour in duration, the loss of extended run time was an acceptable compromise. Apart from differences in operational behavior, the sensors in the validation logger were identical to those used in the field [4].

In addition to using the validation logger, we recorded the animal's movements with a video camera, which later allowed us to correlate visual observations of motions of interest with sensor values. To ensure that the animal was visible throughout the entire video, we conducted these experiments in a small room and used a camera equipped with a wide-angle lens [34]. The room we chose had solid-colored walls and was brightly lit, which facilitated locating the animal both manually and via computer vision. The animal was then allowed to move throughout the room while being recorded by both the logger and video camera. In some trials, a small remote-controlled vehicle was driven around the room to encourage movement, while in other trials, occasional branch-tapping and movements made by human observers served the same purpose.

For our particular validation experiment, we captured five recordings across four unique individuals, each consisting of 20 to 30 min of video and uncompressed accelerometer data. Video was recorded at 640×480 pixels and 30 frames per second, while accelerometer data were captured at the validation logger's full 100 Hz sample rate (25 Hz effective bandwidth).

Synchronization and simultaneous playback

The need for synchronized, simultaneous data and video playback has been made apparent by other validation experiments and is crucial for the accuracy of video-based validation [31, 35]. However, the camera and sensors often operate at substantially different sampling rates and their internal clocks may be mis-matched, complicating the synchronization process.

To properly conduct video validation, the two sources must be closely matched at all points to ensure that the movements displayed in video correspond to the correct movements in data. QValiData provides an interface for synchronizing and aligning data sources. In the synchronization view, a movable marker in the data plotter, resembling the "pointer" on a standard video player's trackbar, indicates the logger samples that are currently associated with the video's playback position. If the video does not reflect the data at this instant, the user may move the marker to the appropriate location in data. QValiData then computes the necessary time offset to apply to the data. Similarly, the data playback rate, expressed as a ratio of data seconds to video seconds, can be adjusted by moving the "Video End" marker to compensate for differences in clock speeds.

During the validation experiment, the logger was tapped or inverted in front of the camera several times at the beginning and end of each trial, to create movement events that would be easily distinguishable from the bird's natural movements. Depending on sensor type, other methods may need to be employed to create a distinguishable event. The timestamps at which these events occurred could then be used to facilitate the aforementioned synchronization process.

Video analysis

Due to the use of a wide-angle lens, the animal occupied only a small fraction of the video's total field of view, making it potentially difficult to see, even with high-resolution recording. To make the animal more visible, QValiData provides a video tracking tool that can magnify the image around areas of movement, enabling closer inspection of the animal (Fig. 4). In addition, the video tracker annotates the sensor data where these movements occur, facilitating the detection of interesting activity events. The video tracker serves primarily to detect the presence of motion in the video for validation purposes, with simple object tracking for video magnification as a secondary objective. Thus, its main purpose is to augment, rather than fully replace, human annotation, and requires close supervision to monitor its accuracy and correct any errors.

Any motion events that were not detected through computer vision but were still visually identifiable on video by the user were manually added. Movements that were not visible by either computer or human (such as when the bird flies out of view or is obscured by objects) were not annotated. The data were trimmed to exclude periods in which the animal was being handled to avoid false readings.

Video analysis was performed with OpenCV (version 4.1.0-pre) [36]. Video frames first undergo a background-subtraction step, which compares each frame with an



Fig. 4 Magnification allows animals' movements to be more easily seen in wide-angle video

average of the previous frames, revealing areas of movement in the current frame [37]. The resulting image is then smoothed with a Gaussian blur, converted to grayscale, and then analyzed with a contour finder to locate areas of the video with significant change, which may indicate animal movement. Areas with movement were tracked using the built-in CSRT tracker, whose output is then smoothed with a Kalman filter, which also assists in re-tracking in case the animal is temporarily occluded by other objects in the scene [38, 39]. All contiguous frames containing motion in a single activity event are compiled into a “motion path”, which is annotated in both video and data (Fig. 5).

Simulation

After annotations have been made, a series of simulations were performed whereby the raw data were played back to simulated bio-loggers in order to gauge their effectiveness at detecting events of interest. Once a method or set of parameters was deemed satisfactory, the activity detector can then be implemented on bio-loggers slated for use in actual experiments.

Our bio-loggers recorded activity using the ADXL362 sensor, a three-axis MEMS accelerometer configured with an acceleration range of ± 4 g and an effective bandwidth of 25 Hz (100 Hz sample rate with a low-pass anti-aliasing filter) [40]. These sensors are capable of measuring accelerations due to animal movement in three axes with great precision, enabling a wide dynamic range of motion to be captured [9]. In addition, they sense the direction of the constant acceleration due to Earth’s gravity, known as static acceleration, which can be employed to deduce position changes [3, 13, 27, 41].

The ADXL362 contains an integrated hardware-based activity detector. Hardware-based activity detectors offer many adjustable parameters and the ability to detect complex movement patterns [42]. While less customizable than activity detection implemented in the logger’s firmware, hardware-based activity detection offers significantly reduced energy consumption by offloading data processing from the logger’s main processor, allowing it to remain in a low-power state until recording is necessary. The ADXL362 activity detector continually monitors all three accelerometer axes for changes in acceleration, and compares these changes to configurable

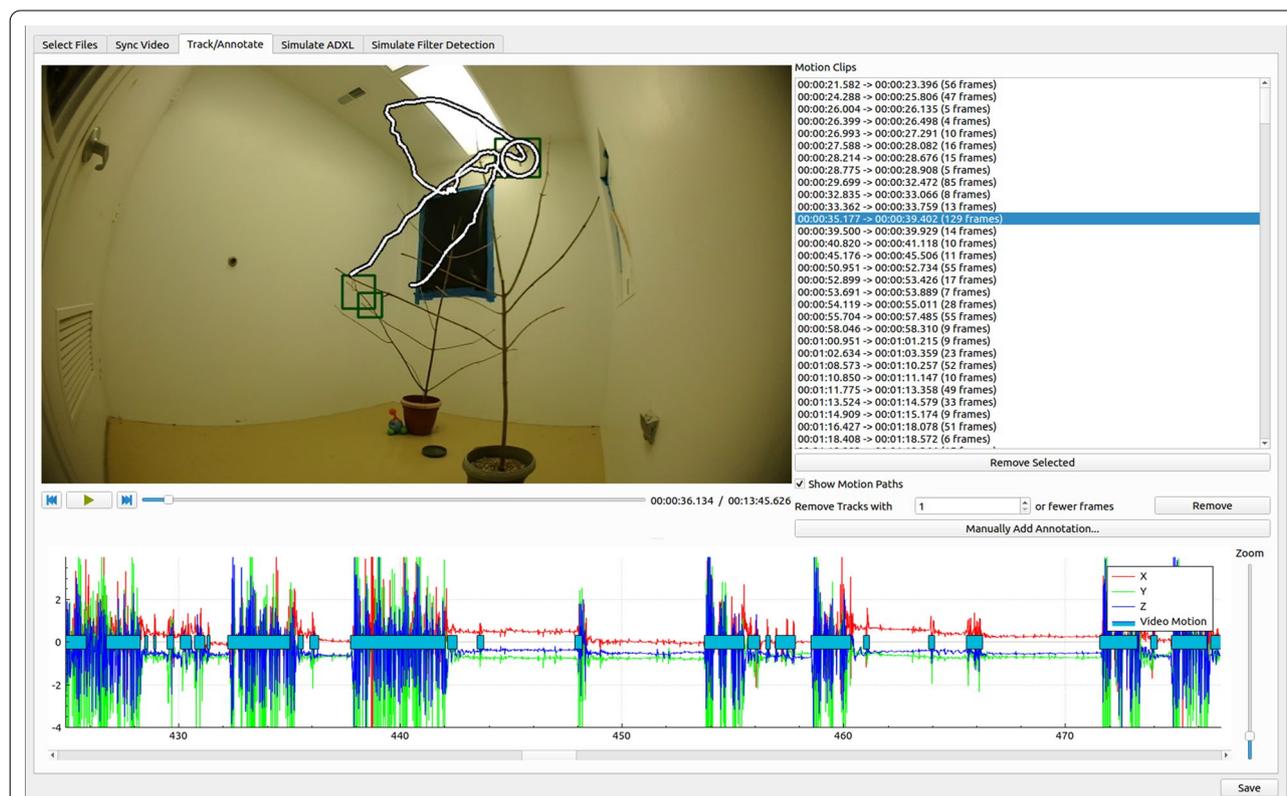


Fig. 5 Motion paths (white line) are drawn on-screen to show animal movements. These are also marked in the data as short blue rectangles. The estimated position of the animal is automatically marked with a circle. For our particular logger, the data plotter displays acceleration (in units of g) in the X, Y, and Z axes

thresholds to determine whether to transition to an “Active” or “Inactive” state (Fig. 4) [43].

When configured to use hardware-based activity detection in its lowest-power setting, known as “Wakeup Mode”, the ADXL362 polls the accelerometer sensor at 6 Hz until the Active state is reached, at which point it transitions to full-speed sampling until it returns to the Inactive state [40]. In cases where an animal is inactive for long time periods, the reduced sample rate will result in further energy savings, albeit at the risk of potentially missing short events.

We incorporated a simulation model of the ADXL362 activity detector into QValiData, exposing parameters that would otherwise not be adjustable during a real experiment. The simulation component of QValiData was designed to stream recorded data in a similar fashion to how activity detectors would ordinarily receive real sensor data, but without any restrictions on real-time sampling rates. In addition to implementing our own simulation models, we have developed a programming interface to simplify the implementation of additional simulation models for other bio-logger types.

Evaluation of methodology

We performed a series of validation experiments following the described methodology to demonstrate its effectiveness at assisting in the selection of activity detection parameters. These tests were performed with the Dark-Eyed Junco, using the ADXL362-based validation logger as our data source. For evaluating our methodology, we chose to compare the performance of various parameter and activity detection modes in detecting significant body movements, such as flying, hopping, and posture changes. Since the ADXL362 in wakeup mode samples at 6 Hz during its inactive state and may not reliably detect very short movements regardless of parameters, movements shorter than 0.5 s (15 video frames) in duration were excluded from performance comparisons.

Parameter selection

For our methodology evaluation purposes, we primarily focused on three parameters in the ADXL362 activity detector: Active Threshold, Inactive Threshold, and Inactive Time. The Active Threshold determines the minimum change in acceleration from a reference value that must be encountered on any of the three axes to transition to Active (high-sample rate) mode. The Inactive Threshold and Inactive Time, respectively, determine the maximum acceleration change in any axis to be considered inactive, and the number of consecutive samples which must read below the Inactive Threshold on all axes to transition back to Inactive mode (Fig. 6). Since the sample rate is 100 Hz when active, the units for Inactive

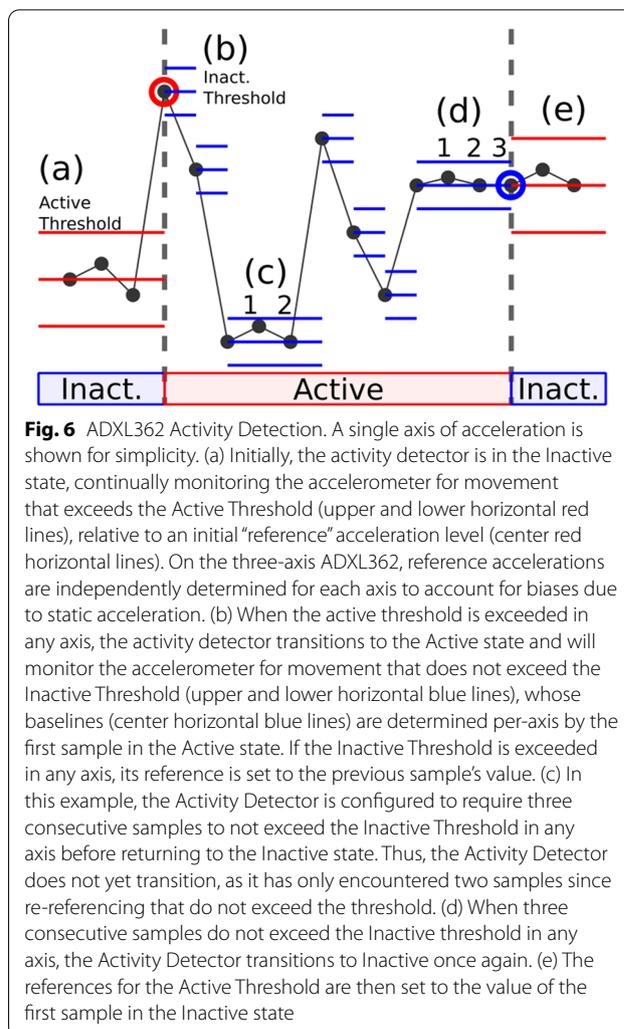


Fig. 6 ADXL362 Activity Detection. A single axis of acceleration is shown for simplicity. (a) Initially, the activity detector is in the Inactive state, continually monitoring the accelerometer for movement that exceeds the Active Threshold (upper and lower horizontal red lines), relative to an initial “reference” acceleration level (center red horizontal lines). On the three-axis ADXL362, reference accelerations are independently determined for each axis to account for biases due to static acceleration. (b) When the active threshold is exceeded in any axis, the activity detector transitions to the Active state and will monitor the accelerometer for movement that does not exceed the Inactive Threshold (upper and lower horizontal blue lines), whose baselines (center horizontal blue lines) are determined per-axis by the first sample in the Active state. If the Inactive Threshold is exceeded in any axis, its reference is set to the previous sample’s value. (c) In this example, the Activity Detector is configured to require three consecutive samples to not exceed the Inactive Threshold in any axis before returning to the Inactive state. Thus, the Activity Detector does not yet transition, as it has only encountered two samples since re-referencing that do not exceed the threshold. (d) When three consecutive samples do not exceed the Inactive threshold in any axis, the Activity Detector transitions to Inactive once again. (e) The references for the Active Threshold are then set to the value of the first sample in the Inactive state

Time correspond to hundredths of a second (0.01 s). The activity detector also features an Active Time setting; however, we chose to leave this at its minimum of 1 sample (0.01 s) since it is not used in wakeup mode [40].

We first set the Active and Inactive Thresholds to their maximum values, and the Inactive Time to its minimum value. Then, the Active and Inactive Thresholds were simultaneously reduced in 0.05 g steps, in order to determine the minimum acceleration change needed to detect all annotated events. We considered an annotated event to be detected when the simulation marked several sections, regardless of length, as “active” consistently throughout its duration. This indicated that the activity detector was able to detect a wide range of motion within an annotated event, rather than only the single most active part. The thresholds were reduced until all annotated events were covered by “active” samples in such fashion, or until no further increase in event coverage could be made without introducing excessive

Table 2 Parameters chosen through simulation-based validation, per trial

	Active thresh. (g)	Inactive thresh. (g)	Inactive time (s)
Trial 1	0.25	0.125	0.22
Trial 2	0.35	0.175	0.30
Trial 3	0.3	0.15	0.21
Trial 4	0.35	0.175	0.11
Trial 5	0.4	0.2	0.20

Each trial consists of 20–30 min of logger and video recordings for 4 individuals

Table 3 Final parameters used in performance evaluation tests

	Active threshold (g)	Inactive threshold (g)	Inactive time (s)
Validated	0.35	0.175	0.21
Bit Tag	0.5	0.5	0.20

false-positive samples. Then, the Inactive Threshold was halved to provide hysteresis. Finally, Inactive Time was increased until all annotated events were marked as “active” for the majority of their duration, or until no further improvement was observed. This reduced the likelihood that brief interruptions in acceleration, such as gliding periods between wing flaps, would cause a corresponding interruption in the recording of a single motion event. After each parameter adjustment, a simulation was run to evaluate performance. The parameters for each trial resulting from applying this method are summarized in Table 2.

For each parameter, we calculated the median across the five trials, to obtain what we will term “Validated” parameters (Table 3). Prior to the development of our simulation-based methodology, we had estimated a set of activity detection parameters for the ADXL362-based logger, using unassisted video analysis without a simulation component. The thresholds were determined by observing video of the bird during periods of significant activity, noting the approximate accelerations experienced by the logger during these periods, and then choosing a value that was low enough to capture the majority of such events. The inactive time was chosen to be long enough to cover any brief interruptions in motion that were determined to be part of a single activity event. These parameters were incorporated into an ultra-low-power logger, known as the “Bit Tag”, that summarized data by counting the number of seconds that the accelerometer was in the “Active” state in a fixed time interval. This set of parameters will be used as a baseline for

comparison against the “Validated” parameters and are henceforth termed “Bit Tag” parameters (Table 3).

We then ran simulations using the parameters from Table 3, both with and without Wakeup Mode enabled, and validated their performance using motion events extracted from video.

Performance metrics

For each run, we collected statistics on both a “per-sample” and “per-event” basis, compiled into Sample Statistics and Event Statistics, respectively.

Sample Statistics were defined as a count of raw accelerometer samples that fell into one of several categories: correctly identified samples, false-positive samples (samples marked as “active” when no activity was detected on video), and false-negative samples (samples marked as “inactive” when activity was detected on video). Since sample statistics are concerned with the performance of activity detection across the cumulative length of recorded data, these would be of interest to loggers employing the asynchronous sampling method.

Event Statistics were defined as the total number of movement events, identified via video analysis, that were classified into the same categories, albeit with slightly different criteria. Movement events were counted based on “coverage”, i.e., the percentage of each movement event that was marked as “active” by the simulated logger. Events with at least 50 percent coverage were considered correctly identified, while those with less than 50 percent coverage were considered missed by the logger and classified as “false negative”. Any sections of data that do not contain visually identifiable motion but were marked as “active” by the activity detector were assumed to have been erroneously marked and were classified as “false positive”. Since these statistics are concerned with counts of event statistics and somewhat normalize for event length, they tend to emphasize the performance of the activity detector in individual events and may be indicative of performance with summarization methods that record event counts. Although other statistics are obtainable in the simulation, these were most relevant to our experiment goals.

Counts of samples and events were summed across all five trials to obtain a single statistic for each set of simulation conditions.

Results

100-Hz sampling mode

When simulations were run at the full 100 Hz data rate, the validated parameters yielded a somewhat higher activity identification rate, particularly for shorter events (Fig. 7) at the expense of a higher false-positive rate (Tables 4, 5). This is to be expected, since our

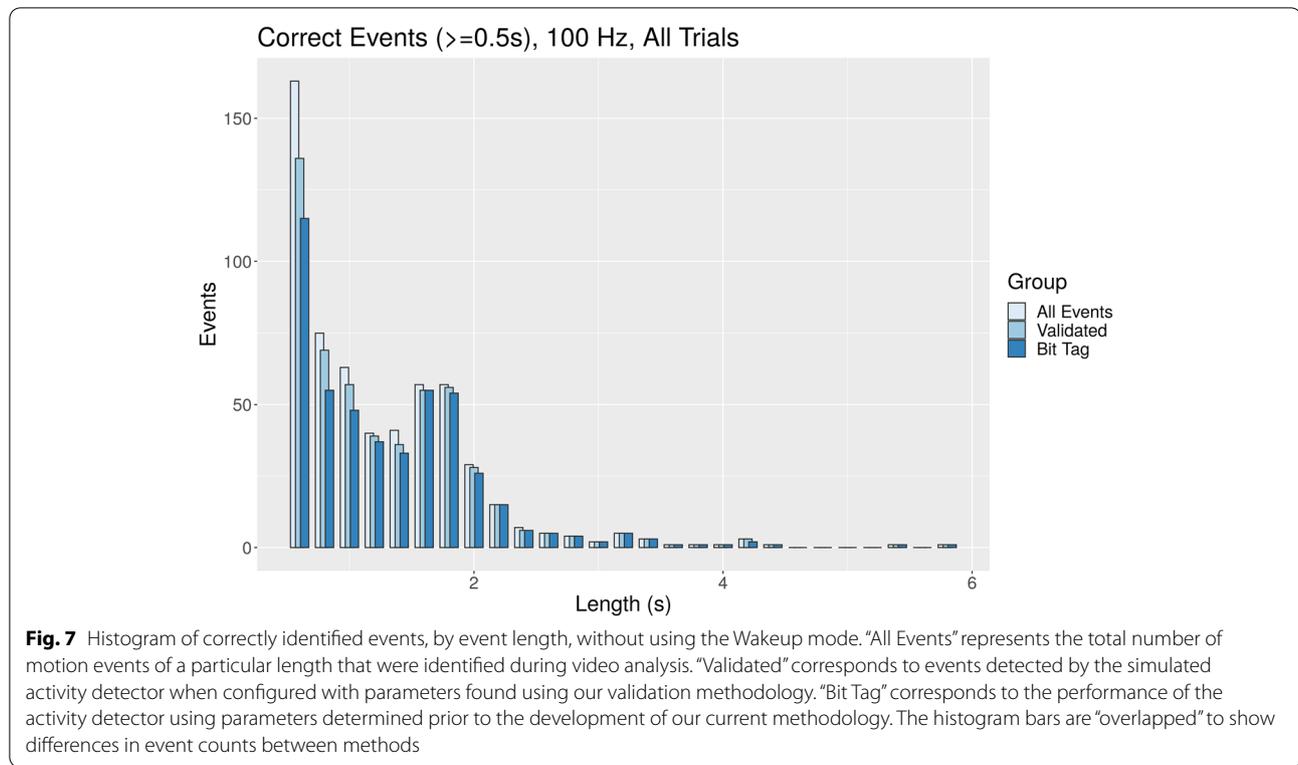


Table 4 Sample statistics, 100 Hz

	Correct samples	False-positive samples	False negative samples	Error rate (% samples)
Validated	550,382	79,710	6415	13.53
Bit Tag	583,569	41,556	11,382	8.32

Sample rate

Table 5 Event statistics, 100 Hz

	Correct events	False-positive events	False negative events	Error rate (% events)
Validated	525	961	50	65.82
Bit Tag	471	828	104	66.43

Sample rate

parameter selection procedure prioritized minimizing missed events, trading off event selectivity for sensitivity. This trend holds for both sample and event statistics. Although both sets of parameters experience a high event error rate due to false positives (Table 5), these false events are very short (Fig. 8). This is consistent with the substantially lower sample error rate (Table 4), implying

that these errors do not contribute substantially to the total length of recorded data.

Wakeup mode

When wakeup mode is enabled, the number of missed samples and events increase considerably (Tables 6, 7). This can be attributed to the slower reaction time of the activity detector, as the sample rate is reduced to 6 Hz in the inactive state. Most of the missed events are quite short (Fig. 9), with the majority of events longer than 1 s being successfully detected. In fact, with the reduced sample rate, a substantial portion of false-positive events and samples were eliminated, resulting in a decrease of both sample and event error rates (Tables 6, 7). As with the simulations performed without wakeup mode, the false-positive rate comprises the majority of error events (Fig. 10), but these events are usually less than 1 s in length and thus do not contribute substantially to the total sample accuracy.

Discussion

Our results show that, with the assistance of simulation-based validation, activity detection parameters can be selected to optimize for a particular data collection strategy (such as minimizing false negatives). In addition, these parameters can be directly and quantitatively compared to each other by simulating on the same data set,

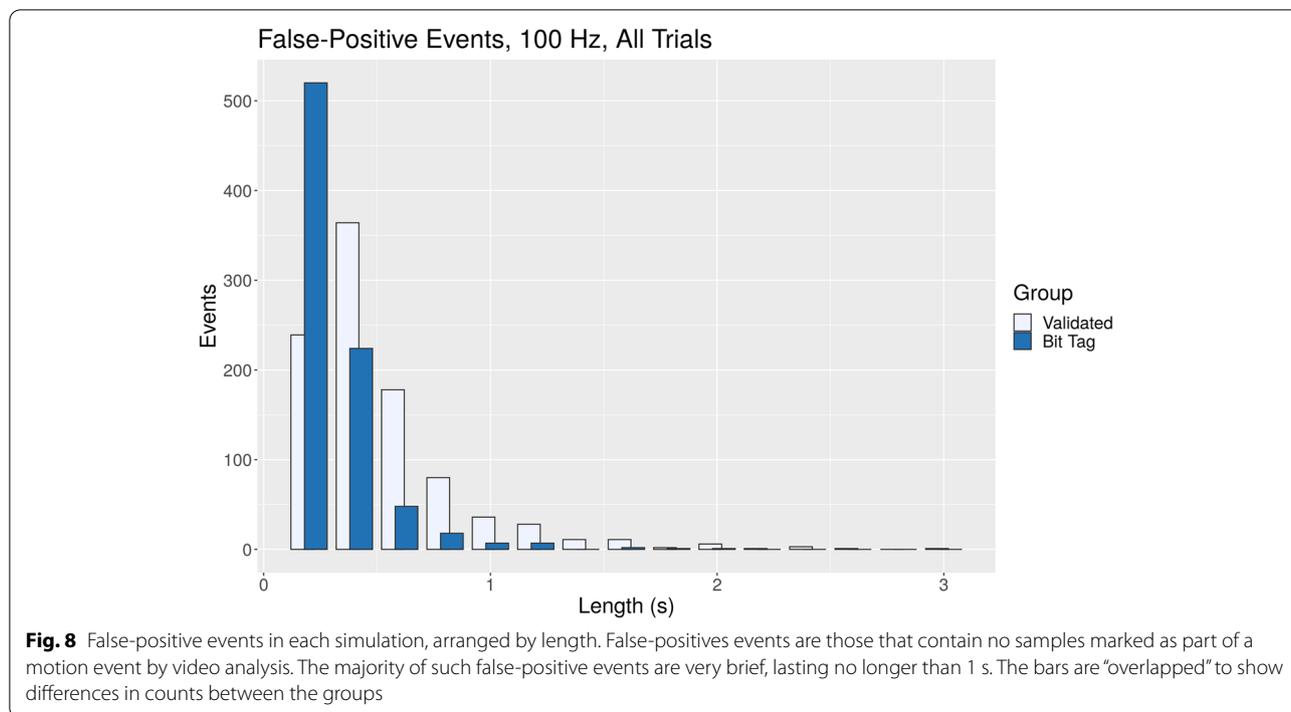


Table 6 Sample statistics, wakeup mode enabled

	Correct samples	False-positive samples	False negative samples	Error rate (% samples)
Validated	571,632	52,284	12,591	10.19
Bit Tag	593,209	25,307	17,991	6.80

Table 7 Event statistics, wakeup mode enabled

	Correct Events	False-positive events	False negative events	Error rate (% events)
Validated	444	437	131	56.13
Bit Tag	375	339	200	58.97

which would not have been possible in a non-simulation validation experiment and can be helpful for both logger and experiment design.

Effect of wakeup mode

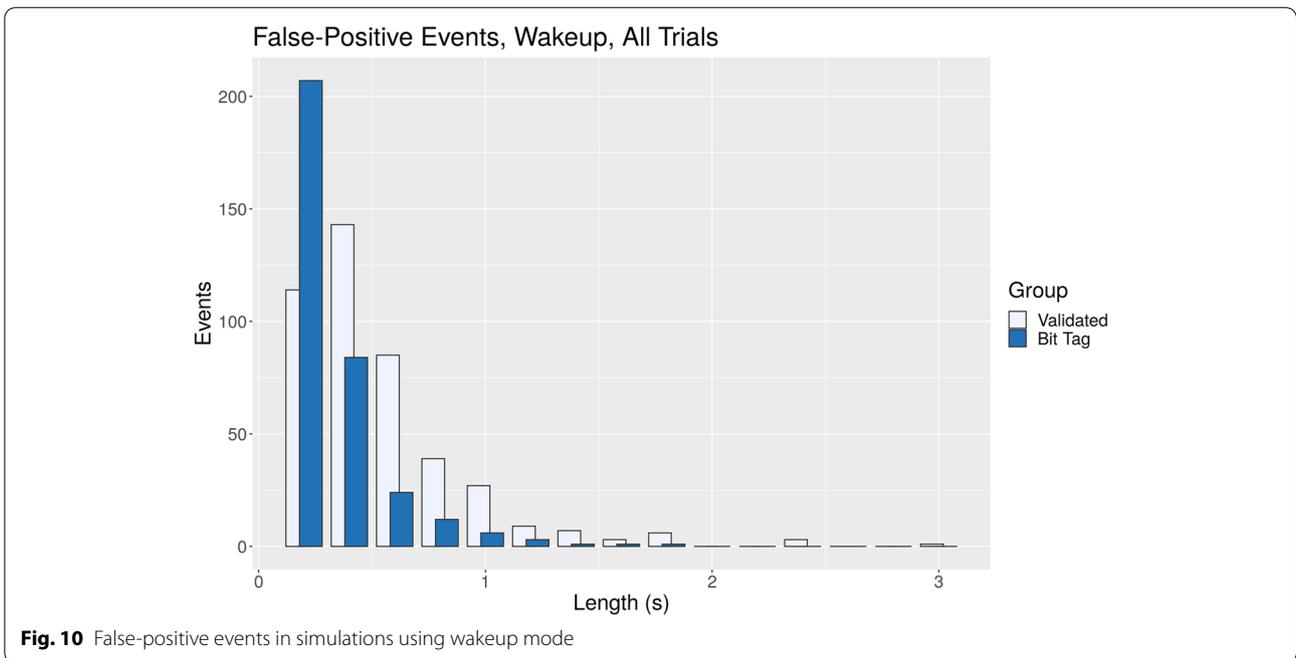
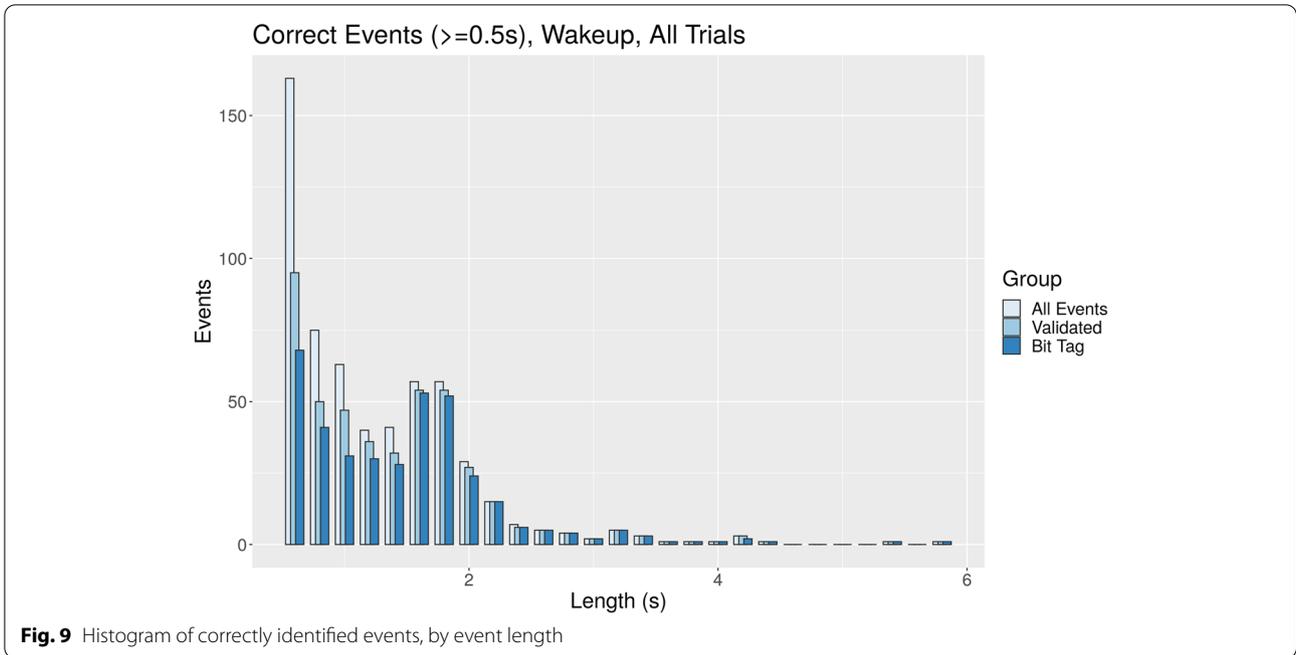
Our simulations revealed that a power-saving feature on our loggers can reduce the occurrence of false positives, with the tradeoff of missing short events. Many of these short events consisted of short twitches and body position changes and did not typically occur when the bird was in motion. Additionally, depending on logger system architecture, excessive triggering may have a substantial

impact on energy usage, especially if the logger consumes additional energy to transition from the inactive to the active state. Thus, the use of the wakeup feature could be beneficial to studies interested only in significant animal movement or locomotion without the added expense of mis-triggers due to short events.

If the recording of short events is desired, the logger can be configured to detect activity at the full sample rate regardless of activity level, which would increase the likelihood of catching such events. However, this may lead to a higher false-positive rate and significantly increases energy consumption during inactive periods, from 0.27 μ A (wakeup mode enabled) to 1.8 μ A (wakeup mode disabled, 100 Hz sample rate) [40], in addition to the energy consumption associated with waking up for additional events. Thus, the selection of a recording strategy may be motivated by energy considerations in addition to accuracy.

Occurrence and management of false-positive events

False-positive events comprised the majority of errors encountered in our validation experiments, and in fact caused the validated parameters to report higher error rates in some cases. As defined by our experiment procedure, false-positive events occur when the activity detector under test marks a portion of data as “active” when no movement was detected in video. Many of these events occurred as a result of small animal movements that were not detected in video analysis, or



possibly due to noise in accelerometer readings. As shown in both sets of validation experiments, these false-positive events are usually very brief and do not contribute much to the total recording time despite their high occurrence. If the reduction of these events is desired and the potential omission of legitimate short events is acceptable, loggers can be configured

to automatically reject events shorter than a predetermined length.

Validation experiment

Our methodology assumes that parameters being developed using animals involved in the validation experiment are applicable to animals in the final experiment.

Depending on animal behavior characteristics, the movements produced by animals in the controlled validation environment may not necessarily reflect their actual movements in the wild. Often, the validation environment is a small room illuminated with artificial lighting which, while favorable for video recording, may not be a suitable environment for recording some behaviors. Even if validation were to take place in a more natural setting, the proximity in time to human interaction (such as during logger attachment or animal observation) has the potential to affect the animal's behavior [5, 44, 45]. However, observer effects may be of less concern in studies where the kinematics of individual actions, rather than overall behavior, are of greater interest.

The loggers and parameters from our experiment were intended to be used primarily in experiments with captive juncos, and the activity of interest was the simple presence or absence of locomotion. Therefore, we believe this assumption would be true for our purposes. However, in other cases, it may be necessary to perform validation experiments in various different settings and to perform other preliminary experiments to characterize any potential differences in behavior between animals in the validation experiment and those in the final experiment.

Ideally, a validation experiment should draw upon data from several trials. Although simulation greatly reduces the number of trials necessary for a particular individual, it would be beneficial to incorporate validation trials from multiple individuals to control for differences in individual behavior. As we have observed from our experiment data, each animal produces unique activity patterns, varying in movement frequency, activity intensity, and other factors. Because of this, the occurrence of activity event misclassifications is unavoidable as activity detection is not always perfect. As a result, one must choose between recording all events of interest and possibly recording non-interesting events, or only recording events of interest and accepting the possibility of missed events. The decision will ultimately depend on the nature of the animal and bio-logger used in a particular experiment.

Our current data set is small, consisting of five trials with four individuals. As a result, some movements or behaviors may not be represented equally or in all trials, which may influence the parameters selected. Additionally, environmental factors such as objects placed in the test room can influence behavior. For instance, the absence of tree branches in the test room resulted in a complete absence of short-range flights in one trial, which raised its minimum acceleration threshold for flight detection and thus increased the median thresholds for validating other trials. Thus, future experiments

should collect data from as many unique individuals as possible, in settings that most resemble the animal's natural habitat.

In its current iteration, QValiData is only able to track large movements and may not be able to automatically annotate other kinds of behaviors with small movements, such as feeding or preening. This is a limitation of the present video tracking implementation, but should have little effect on validation if behaviors can still be identified by eye and manually annotated. Likewise, our methodology is only able to validate the detection of movements that can be reliably recorded by bio-logger sensors. For example, bio-loggers with accelerometer sensors may be unable to detect motions with very little changes in acceleration regardless of activity detection method. Therefore, proper sensor selection plays a crucial role in the effectiveness of validation for particular types of motion.

Development of QValiData

Prior to the development of QValiData, our workflow consisted of video magnification and motion tracking in Blender (version 2.79) [46] and simultaneous playback and data annotation in ELAN (version 5.1) [47]. Although Blender contains powerful motion tracking and video processing utilities, it was unable to consistently track flying birds, as wing flapping and erratic movements interfered with its motion capture system. This process alone would take approximately one hour for a 20- to 30-min video. ELAN allowed us to simultaneously play back video alongside data tracks and insert time-stamped annotations into the data. We found that pre-rendering videos with motion tracks was unnecessarily slow, requiring several hours of processing time, and it was difficult to track and manage files across different applications. Thus, we were motivated to develop a bespoke software tool that could manage experiment files, play back synchronized video and data, render motion-tracked video in real time, and run simulations of bio-logger data. Although the present iteration of QValiData still requires some human intervention for video annotation, it has virtually eliminated the need for video pre-rendering, which represented one of the most time-consuming portions of the workflow.

Simulation performance

One of the primary advantages of simulation is that using pre-recorded data allows the simulation to run faster than on real hardware, as it eliminates the need to wait for real-time sensor data. In our testing, a 20-min trial containing over 100,000 samples of accelerometer data could be simulated on a desktop computer in less than one second. Another advantage is the repeatability of tests, since different activity detection methods can be

tested on a consistent set of movements and directly compared with one another, facilitating quantitative comparisons and iterative refinement.

Expanding QValiData: frequency-domain example

Future validation experiments should also explore this methodology's ability to predict activity detection parameters for other bio-logger types and animal species, as well as for other movements of interest. For instance, special considerations may need to be made for sensors with lower sample rates or resolution, or for animals whose movement types are more difficult to distinguish from one another [20]. Although the methodology is in theory adaptable to such changes, real-world data would be beneficial in assessing its capabilities under different conditions.

One possible area of development is the implementation of activity detectors based on frequency-domain analysis. Although our present ADXL362 activity detector is sufficient for detecting a wide variety of energetic movements, it is only capable of detecting changes in acceleration. Thus, it may have trouble differentiating certain kinds of movement, such as wing flapping and fast posture changes. Although both wing flapping and fast posture changes produce similar short-term (within a wing-beat period) acceleration changes, wing flapping produces periodic changes of a relatively consistent frequency. However, the present activity detector is unable to distinguish between periodic and non-periodic acceleration changes.

Periodic acceleration change occurs in many kinds of animal locomotion, such as flying, walking, or swimming, and may be used to estimate energy expenditure, or to distinguish them from other movement types [48–51]. Particularly relevant to our study of the Dark-Eyed Junco is the detection of *zugunruhe*, or migratory restlessness, in captive individuals, which manifests as wing flapping, even when confined to cages [52]. An activity detector designed to detect periodic movements can make use of a digital filter to extract or exclude certain frequency components from movement data. For instance, a high-pass filter may be used to remove the constant acceleration due to gravity, while a band-pass filter centered around a bird's wing-beat frequency can be used to detect flight periods [49, 51].

In applications where energy limitations are less of a concern than storage, more advanced data analysis or classification methods that would ordinarily be done post hoc, may instead be implemented onboard the bio-logger itself. This would result in reduced sample memory usage at the cost of increased energy consumption due to added processing. For instance, movement classification involving decision trees or variability analysis may simply

store the classifier output instead of raw data. There exist bio-loggers with energy-harvesting devices that are constrained less by energy than by memory, creating a possible use case for this method [20]. Nevertheless, these more sophisticated classification methods must still be validated to ensure their correctness.

Inferring behaviors from data

Although the present validation methodology provides an approach to the question of how to determine the correctness of various data collection strategies, it does not currently address the question of how to directly infer behaviors from collected data. This topic has been addressed in other studies that have developed various behavior classification methods for sensor data [6, 19, 20, 26–29]. However, if such classification methods were to be implemented in a bio-logger simulation model, it would be possible, using our simulation methodology, to assess the accuracy of the behavior inferences made using such classifiers.

Combined video and sensor data analysis

Video annotation can be improved by incorporating sensor data into the motion tracking process. Currently, without human intervention, the video tracker must play through all frames to detect motion, which leads to long processing times. With the assistance of a simple filter that can detect periods of low or no activity in the sensor data that are certain to contain no interesting events, the video tracker will be able to skip these portions automatically. This filter will necessarily over-approximate the behavior of the activity detector to be validated, as it should be able to narrow the search space of sensor data without inadvertently omitting potentially interesting events.

Conclusion

The use of compressed recording methods requires careful characterization of an animal's movements and the bio-logger's response to such movements. This often requires several rounds of experimentation along with visual observation to ensure that the bio-logger is able to record activities of interest, but this leads to time-consuming trial-and-error experiments with animals that may not always provide consistent movement data. Our proposed methodology alleviates some of these issues by using simulated loggers, enabling activity detection methods to be quickly refined and evaluated with a reusable set of real data. Compressed sampling in bio-logging allows for the collection of activity data over increasingly long intervals and for smaller species, where energy and memory budgets are prohibitively constrained for conventional continuous recording methods. The use of

proper validation procedures ensures that the activity detection driving compressed recording operates correctly, increasing confidence in the use of these methods for long-term studies of animals in motion.

Acknowledgements

The authors would like to acknowledge Dustin Brewer, David Crandall, Dillon Gaughan, Bryce Himebaugh, and Zack Marvel for their support throughout the project.

Authors' contributions

All authors contributed to the development of the validation methodology as well as data analysis. JC developed QValiData and wrote the first draft of the manuscript with extensive input from GB and AF. GB led the development of the validation logger and provided insight on activity detection methods. AF oversaw data collection and advised on the animal factors of bio-logging experiments. All authors read and approved the final manuscript.

Funding

This material is based upon work supported by the National Science Foundation under Grant No. 1644717. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Availability of data and materials

QvaliData: Project Home Page: <https://github.com/jchen-cs/QValiData>. Archived Version: <https://zenodo.org/record/5061662>. Operating Systems: Ubuntu, macOS. Programming Language: C++-. Compiling from source requires OpenCV 4, Qt 5, and iir1-dev. License: GPLv3. A video tutorial can be found at: <https://zenodo.org/record/4580125>. The datasets supporting the conclusions of this article are available in the Zenodo repository, <http://doi.org/10.5281/zenodo.5061055> [53].

Declarations

Ethics approval and consent to participate

The experiments described in the current study were approved by the Indiana University Institutional Animal Care and Use Committee (protocol # 15-025).

Consent for publication.

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹School of Informatics, Computing and Engineering, Indiana University, Bloomington, IN 47405, USA. ²Environmental Resilience Institute, Department of Biology, Indiana University, Bloomington, IN 47405, USA.

Received: 23 October 2020 Accepted: 14 July 2021

Published online: 13 September 2021

References

- Holland RA, Wikelski M, Wilcove DS. How and why do insects migrate? *Science*. 2006;313(5788):794–6.
- Milner-Gulland EJ, Fryxell JM, Sinclair ARE. Migration impacts on communities and ecosystems: empirical evidence and theoretical insights. *Animal Migration: A Synthesis*: Oxford Scholarship Online; 2011. p. 131–43. <https://doi.org/10.1093/acprof:oso/9780199568994.003.0009>.
- Brown DD, Kays R, Wikelski M, Wilson R, Klimley AP. Observing the unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry*. 2013;1(1):20.
- Hammond TT, Springthorpe D, Walsh RE, Berg-Kirkpatrick T. Using accelerometers to remotely and automatically characterize behavior in small animals. *J Exp Biol*. 2016;219(11):1618.
- Wilson RP, Grant WS, Duffy DC. Recording devices on free-ranging marine animals: does measurement affect foraging performance? *Ecology*. 1986;67(4):1091–3.
- Sakamoto KQ, Sato K, Ishizuka M, Watanuki Y, Takahashi A, Daunt F, et al. Can ethograms be automatically generated using body acceleration data from free-ranging birds? *PLoS ONE*. 2009;4(4):e5379.
- Liechti F, Witvliet W, Weber R, Bächler E. First evidence of a 200-day non-stop flight in a bird. *Nat Commun*. 2013;4(1):2554.
- Bäckman J, Andersson A, Alerstam T, Pedersen L, Sjöberg S, Thorup K, et al. Activity and migratory flights of individual free-flying songbirds throughout the annual cycle: method and first case study. *J Avian Biol*. 2016;48:309.
- Backman J, Andersson A, Pedersen L, Sjöberg S, Tottrup AP, Alerstam T. Actogram analysis of free-flying migratory birds: new perspectives based on acceleration logging. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol*. 2017;203(6–7):543–64.
- Nuijten RJM, Gerrits T, Shamoun-Baranes J, Nolet BA. Less is more: On-board lossy compression of accelerometer data increases biologging capacity. *J Anim Ecol*. 2020;89(1):237–47.
- Bridge ES, Thorup K, Bowlin MS, Chilson PB, Diehl RH, Fléron RW, et al. Technology on the move: recent and forthcoming innovations for tracking migratory birds. *Bioscience*. 2011;61(9):689–98.
- Tomotani BM, Bil W, van der Jeugd HP, Pieters RPM, Muijres FT. Carrying a logger reduces escape flight speed in a passerine bird, but relative logger mass may be a misleading measure of this flight performance detriment. *Methods Ecol Evol*. 2019;10(1):70–9.
- Chivers LS, Hatch SA, Elliott KH. Accelerometry reveals an impact of short-term tagging on seabird activity budgets. *Condor*. 2016;118(1):159–68.
- Sjöberg S, Pedersen L, Malmiga G, Alerstam T, Hansson B, Hasselquist D, et al. Barometer logging reveals new dimensions of individual songbird migration. *J Avian Biol*. 2018;49:e01821.
- Shiple JR, Kapoor J, Dreelin RA, Winkler DW, Lecomte N. An open-source sensor-logger for recording vertical movement in free-living organisms. *Methods Ecol Evol*. 2017;9(3):465–71.
- Nishiumi N, Matsuo A, Kawabe R, Payne N, Huvneers C, Watanabe YY, et al. A miniaturized threshold-triggered acceleration data-logger for recording burst movements of aquatic animals. *J Exp Biol*. 2018. <https://doi.org/10.1242/jeb.172346>.
- Hedenström A, Norevik G, Warfvinge K, Andersson A, Bäckman J, Åkesson S. Annual 10-month aerial life phase in the common Swift *Apus apus*. *Curr Biol*. 2016;26(22):3066–70.
- Williams HJ, Shepard ELC, Duriez O, Lambertucci SA. Can accelerometry be used to distinguish between flight types in soaring birds? *Anim Biotelemetry*. 2015;3(1):45.
- Bom RA, Bouten W, Piersma T, Oosterbeek K, van Gils JA. Optimizing acceleration-based ethograms: the use of variable-time versus fixed-time segmentation. *Mov Ecol*. 2014;2(1):6.
- Shamoun-Baranes J, Bom R, van Loon EE, Ens BJ, Oosterbeek K, Bouten W. From sensor data to animal behaviour: an oystercatcher example. *PLoS ONE*. 2012;7(5):37997.
- Kawabata Y, Noda T, Nakashima Y, Nanami A, Sato T, Takebe T, et al. Use of a gyroscope/accelerometer data logger to identify alternative feeding behaviours in fish. *J Exp Biol*. 2014;217(18):3204.
- Watanabe YY, Takahashi A. Linking animal-borne video to accelerometers reveals prey capture variability. *Proc Natl Acad Sci U S A*. 2013;110(6):2199–204.
- Tanoue H, Komatsu T, Tsujino T, Suzuki I, Watanabe M, Goto H, et al. Feeding events of Japanese lates *Lates japonicus* detected by a high-speed video camera and three-axis micro-acceleration data-logger. *Fish Sci*. 2012;78(3):533–8.
- Green JA, Halsey LG, Wilson RP, Frappell PB. Estimating energy expenditure of animals using the accelerometry technique: activity, inactivity and comparison with the heart-rate technique. *J Exp Biol*. 2009;212(5):745–6.
- Elliott KH, Le Vaillant M, Kato A, Speakman JR, Ropert-Coudert Y. Accelerometry predicts daily energy expenditure in a bird with high activity levels. *Biol Lett*. 2013;9(1):20120919.
- McClune DW, Marks NJ, Wilson RP, Houghton JDR, Montgomery IW, McGowan NE, et al. Tri-axial accelerometers quantify behaviour in the Eurasian badger (*Meles meles*): towards an automated interpretation of field data. *Animal Biotelemetry*. 2014;2(1):5.

27. Tsuda Y, Kawabe R, Tanaka H, Mitsunaga Y, Hiraishi T, Yamamoto K, et al. Monitoring the spawning behaviour of chum salmon with an acceleration data logger. *Ecol Freshw Fish*. 2006;15(3):264–74.
28. Collins PM, Green JA, Warwick-Evans V, Dodd S, Shaw PJA, Arnould JPY, et al. Interpreting behaviors from accelerometry: a method combining simplicity and objectivity. *Ecol Evol*. 2015;5(20):4642–54.
29. Fehlmann G, O'Riain MJ, Hopkins PW, O'Sullivan J, Holton MD, Shepard ELC, et al. Identification of behaviours from accelerometer data in a wild social primate. *Animal Biotelemetry*. 2017. <https://doi.org/10.1186/s40317-017-0121-3>.
30. Ringgenberg N, Bergeron R, Devillers N. Validation of accelerometers to automatically record sow postures and stepping behaviour. *Appl Anim Behav Sci*. 2010;128(1–4):37–44.
31. Uhrikova Z, Sprdlik O, Hoskovcova M, Komarek A, Ulmanova O, Hlavac V, et al. Validation of a new tool for automatic assessment of tremor frequency from video recordings. *J Neurosci Methods*. 2011;198(1):110–3.
32. Renart A, Machens CK. Variability in neural activity and behavior. *Curr Opin Neurobiol*. 2014;25:211–20.
33. Nolan Jr. V, Ketterson ED, Cristol DA, Rogers CM, Clotfelter ED, Titus RC, et al. Dark-eyed Junco (*Junco hyemalis*). Birds of the World Cornell Lab of Ornithology. 2020. <https://doi.org/10.2173/bow.daejun.01>.
34. Olympus. Olympus Stylus TG-Tracker. https://www.olympusamerica.com/files/oima_cckb/tg_tracker_ENU.pdf. Accessed 23 Jun 2020.
35. Berg RTVd. Designing a small and low-energy wildlife tag for parakeets within an urban environment capable of tracking and online activity recognition. 2019. <http://essay.utwente.nl/79501/>. Accessed 3 Jun 2020.
36. OpenCV Team. Open Source Computer Vision Library. 4.1.0-pre ed2019. <https://opencv.org>.
37. OpenCV Team. How to Use Background Subtraction Methods 2018. https://docs.opencv.org/3.4/d1/dc5/tutorial_background_subtr.html. Accessed 23 Jun 2020.
38. Mallick S. Object Tracking using OpenCV (C++/Python) 2017. <https://www.learnopencv.com/object-tracking-using-opencv-cpp-python/>. Accessed 23 Jun 2020.
39. Rosebrock A. OpenCV Object Tracking 2018. <https://www.pyimagesearch.com/2018/07/30/opencv-object-tracking/>. Accessed 23 Jun 2020.
40. Analog Devices. Micropower, 3-Axis, $\pm 2\text{ g}/\pm 4\text{ g}/\pm 8\text{ g}$ Digital Output MEMS Accelerometer. 2019. <https://www.analog.com/media/en/technical-documentation/data-sheets/ADXL362.pdf>. Accessed 23 Jun 2020.
41. Shepard ELC, Wilson RP, Halsey LG, Quintana F, Gómez Laich A, Gleiss AC, et al. Derivation of body motion via appropriate smoothing of acceleration data. *Aquat Biol*. 2008;4:235–41.
42. STMicroelectronics. iNEMO inertial module: always-on 3D accelerometer and 3D gyroscope. <https://www.st.com/resource/en/datasheet/ism6dsox.pdf>. Accessed 23 Jun 2020.
43. Analog Devices. ADXL362. ADXL362InterruptAnimation_201302.pptx, editor. PowerPoint. ez.analog.com. 2013. https://ez.analog.com/cfs-file/__key/telligent-evolution-components-attachments/00-334-01-00-00-08-10-17/ADXL362InterruptAnimation_5F00_201302.pptx. Accessed 10 Aug 2020.
44. Rasmussen D. Observer influence on range use of macaca arctoides after 14 years of observation? *Laboratory Primate Newsletter*. 1991;30(3):6.
45. Jack KM, Lenz BB, Healan E, Rudman S, Schoof VAM, Fedigan L. The effects of observer presence on the behavior of *Cebus capucinus* in Costa Rica. *Am J Primatol*. 2008;70(5):490–4.
46. Blender Foundation. Blender. 2.79 ed2017. <https://www.blender.org>.
47. Nijmegen: Max Planck Institute for Psycholinguistics TLA. ELAN. 5.1 ed2018. <https://archive.mpi.nl/tla/elan>.
48. Gleiss AC, Wilson RP, Shepard ELC. Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure. *Methods Ecol Evol*. 2011;2(1):23–33.
49. Halsey LG, Portugal SJ, Smith JA, Murn CP, Wilson RP. Recording raptor behavior on the wing via accelerometry. *J Field Ornithol*. 2009;80(2):171–7.
50. Chung WY, Purwar A, Sharma A. Frequency domain approach for activity classification using accelerometer. *Conf Proc IEEE Eng Med Biol Soc*. 2008;2008:1120–3.
51. Van Walsum TA, Perna A, Bishop CM, Murn CP, Collins PM, Wilson RP, et al. Exploring the relationship between flapping behaviour and accelerometer signal during ascending flight, and a new approach to calibration. *Ibis*. 2019;162(1):13–26.
52. Ramenofsky M, Agatsuma R, Barga M, Cameron R, Harm J, Landys M, et al, editors. *Migratory Behavior: New Insights from Captive Studies*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2003. https://doi.org/10.1007/978-3-662-05957-9_6.
53. Chen, J., Brown, G. & Fudickar, A. QValidata Experiment Data. Zenodo. (2020). <https://doi.org/10.5281/zenodo.5061055>

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

