





SYSTEMS ARTICLE

Animal-borne wireless network: Remote imaging of community ecology

Shinkyu Park¹  | Konrad H. Aschenbach³ | Manjur Ahmed¹ | William L. Scott²  |
Naomi E. Leonard²  | Kyler Abernathy³ | Greg Marshall⁴ | Mike Shepard³ |
Nuno C. Martins¹ 

¹Department of Electrical and Computer Engineering, Institute for Systems Research, University of Maryland, College Park, Maryland

²Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, New Jersey

³Exploration Technology Lab, National Geographic Society, Washington, District of Columbia

⁴Marshall Innovation, Alexandria, Virginia

Correspondence

Shinkyu Park, Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544.
Email: skpark0828@gmail.com

Funding information

NSF Division of Electrical, Communications and Cyber Systems, Grant/Award Numbers: 1135726, 1135724, 1135719; Air Force Office of Scientific Research, Grant/Award Number: FA95501510367; National Geographic Society

Abstract

This article describes the design, construction, and field-testing of a standalone networked animal-borne monitoring system conceived to study community ecology remotely. The system consists of an assemblage of identical battery-powered sensing devices with wireless communication capabilities that are each collar-mounted on a study animal and together form a mobile ad hoc network. The sensing modalities of each device include high-definition video, inertial accelerometry, and location resolved via a global positioning system module. Our system is conceived to use information exchange across the network to enable the devices to jointly decide without supervision when and how to use each sensing modality. The ultimate goal is to extend battery life while making sure that important events are appropriately documented. This requires judicious use of highly informative but power-hungry sensing modalities, such as video, because battery capacity is constrained by stringent weight and dimension restrictions. We have proposed algorithms to regulate sensing rates, data transmission among devices, and triggering for video recording based on location and animal group movements and configuration. We have also developed the hardware and firmware of our devices to reliably execute these algorithms in the exacting conditions of real-life deployments. We describe validation of the performance and reliability of our system using deployment results for a mission in Gorongosa National Park (Mozambique) to monitor two species in their natural habitat: the waterbuck and the African buffalo. We present movement data and snapshots of animal point-of-view videos collected by 14 fully operational devices collared on 10 waterbucks and 4 buffaloes.

KEYWORDS

environmental monitoring, sensor networks

1 | INTRODUCTION

The study of social behavior of animal groups—as in predation, evasion, foraging, and migration—involves establishing hypotheses that explain how and why individuals in animal groups interact, and how the interactions lead to observed group phenomena (Ballerini

et al., 2008; Couzin & Krause, 2003; Sumpter, 2006; Vicsek & Zafeiris, 2012). To validate hypotheses, extensive data on animal group behavior need to be collected and analyzed.

Logging animal behavioral data by human observers is the most direct method, which requires ample time and effort, and is often hindered by spatiotemporal restrictions. In addition, human presence

can influence the studied behavior (Altmann, 1974). Numerous efforts to develop and deploy animal-borne systems that collect behavioral data, over a range of dimensions (Dyo et al., 2010; Juang et al., 2002; Krause et al., 2013; Marshall, 1998; Zhang, Sadler, Lyon, & Martonosi, 2004) sought to overcome these limitations.

Notably, there is growing interest in exploiting animal-borne imaging units to obtain animal point-of-view video recordings. In conjunction with geolocation data, they provide valuable information on an animal's interactions with the surrounding environment, other members of its species, and other species. However, because of battery capacity limitations, animal-borne systems can record video and other data continuously only for short deployments, during which significant events may not occur. Hence, techniques that enable selective data collection are crucial to maximize the deployment lifetime and, consequently, improve efficacy of these systems.

In this article, we describe the design, construction, and testing of an animal-borne monitoring system that, in contrast with existing technologies, consists of an ad hoc wireless network of identical battery-powered sensing devices that can execute coordinated strategies to enhance the relevance of the data logged while extending battery life. Each device, which is collar-mounted on a study animal, can record high-definition video, inertial accelerometry, and location provided by a global positioning system (GPS) module. Algorithms executed at every device use the information shared throughout the network to determine when and how each sensing modality is used. The controlled parameters include the rate at which data are registered and when to record video, which may be determined by triggers based on distance, location, or group motion and configuration. We designed the collar to reliably self-release each device from the animal once the battery is fully discharged, after which the deployment is considered concluded for the device. Once a device's deployment is complete, it activates an onboard radio-frequency beacon, which researchers use for localization, retrieval, and subsequent data download.

1.1 | Main technical challenges

Our goal is to develop a system to enable deployments that last weeks, not days, subject to stringent reliability and size constraints. These goals rule out the use or customization of proprietary off-the-shelf consumer products, such as cell phones, tablets, or action cameras. We concluded that an entirely new class of devices had to be developed, for which algorithms, hardware, and firmware had to be codesigned to address the following challenges:

1. *Limited battery capacity*—The capacity of the batteries that power every device is limited by the strict weight and size restrictions that animal-borne monitoring instruments must satisfy. Consequently, strategic management of power-hungry components, such as the camera and communication modules, is critical because, as we describe above, recharging the battery requires redeployment of the device.

2. *Communication constraints*—Due to the unpredictable (or undiscovered) mobility of free-ranging animals, it can be difficult to establish reliable communication links among the devices. As a result, communication strategies must be designed that enable reliable data sharing.

3. *Video recording*—Because of onerous power and memory requirements, video recording should only be activated sparingly, and preferably only when significant events occur. Our strategy is to use triggers to initiate video capture based on location, group movement, and configuration, which requires each device to sustain estimates of the positions and movements of other pertinent members of the group. The design of estimation algorithms that meet the aforementioned requirement, subject to the battery and communication constraints mentioned above, is a significant challenge.

The capacity of the battery is limited primarily by restrictions on the size and weight of the device, which must not exceed 3% of the weight of the animal it is mounted on. Our choice to equip each device with a 10.2 Ah battery pack¹ enabled us to package it and all the electronics in a compact housing that meets the weight limit. The maximum current draw of the device is approximately 320 mA, including video recording. Hence, each device at full power would operate for no more than 21 hr, which highlights the importance of addressing the challenges above for enhancing the efficiency of data collection and extending the deployment duration.

1.2 | Paper organization and outline of main contributions

This article has seven sections and three appendices. After the introduction given in Section 1, in Section 2 we outline a three-stage process we adopted to develop our system. In Section 3, we provide a concise survey of related literature reporting on animal-borne monitoring systems and their use for ecology studies.

Sections 4, 5, and 6 discuss our main contributions to address the challenges listed in Section 1.1, as outlined below:

- *Section 4* describes the design of three classes of model-based algorithms. Section 4.1 focuses on a sampling rate control algorithm that regulates the sensing and communication rates of the devices, Section 4.2 proposes event-driven strategies to reduce the rate of transmission of data for remote estimation, and Section 4.3 is centered on the design of distributed estimators tasked with sustaining omniscience of key state variables subject to network limitations.
- *Section 5* describes the hardware and the associated firmware. The hardware integration of the sensors, communication module, memory, and processing unit is discussed in Section 5.1, while

¹Due to its discharge curve characteristics and the fact that certain components, such as the camera module, require a minimum of 3.5 V to operate, the usable capacity of the battery is estimated at 6.8 Ah.

Section 5.2 outlines the architecture of the firmware that executes our algorithms, manages the sensors, communication unit, power usage, data flow among components, and data storage.

- *Section 6* summarizes a deployment of our system in Gorongosa National Park in Mozambique to monitor the behavior of two species in their natural habitat. A description of the system configuration used for the deployment is given in Section 6.1, while Sections 6.2 and 6.3 discuss lessons learned from the deployment data, including an analysis of the performance of the system and algorithms.

Section 7 is a summary of the paper that also discusses future directions and other possible uses of our system.

2 | OVERVIEW OF SYSTEM DEVELOPMENT

To create a system that addresses all three challenges listed in Section 1.1, we have adopted the three-phase development cycle depicted in Figure 1. Using lessons learned from a field deployment to refine the system's data collection performance and improve the quality of the data collected is the main benefit of such a systematic approach.

Below we briefly explain how we have carried out each of the three phases, while the details are provided in Sections 4, 5, and 6. A photograph of the resulting system in Figure 2 displays how the electronics is packaged in each device, which is also shown in more detail in Figure 10. A photo, selected video frames, and additional screenshots in Figure 2 also show a collared study animal and illustrate the type of imagery and tracking data obtained by our system.

2.1 | Phase 1 (model-based algorithm design)

In the first phase, we design model-based algorithms that regulate how the system adaptively chooses sensing and communication rates, and how key variables of animal group movements are estimated under communication and power constraints. In particular, we have designed sampling rate control, remote estimation, and distributed estimation algorithms, which are based on behavioral models of the animals we wish to monitor. The models used in this phase can be acquired from the existing body of animal ecology literature (see, for instance, Codling, Plank, & Benhamou, 2008; Fleming et al., 2014; McClintock et al., 2012), or through an analysis of deployment data, when the data are available. In adopting these algorithms, it is imperative to determine whether they can be executed in real-time and if the required sensing and communication modalities are realizable.

2.2 | Phase 2 (hardware/firmware design)

In the second phase, we build devices that can realize the functionalities that underlie the algorithms designed in the previous phase and other tasks required for stable and responsive system operation. In our design, we chose components that can support the sensing and communication modalities required by the algorithms.

For our purposes, to sense animal motions/movements and to record animal point-of-view video data, our animal-borne devices are equipped with a GPS, an inertial measurement unit (IMU), and an onboard camera. It also has a slot for a micro-SD card to store incoming data, and a wireless communication unit to connect to other devices and exchange data. In addition, each device has a microcontroller unit (MCU) to run the firmware, which is designed to implement all the essential functionalities for effective power management, autonomous data processing, and selective recording of animal point-of-view videos. That is to say that it executes our algorithms, and manages all the sensors, components, and data flows. In addition, it provides an interface between low-level electronic components and high-level modules that realize the algorithms.

2.3 | Phase 3 (field deployment)

Once a new iteration of the system is ready, in the third phase, we conduct a field deployment, which is also used for gathering performance data. In this paper, we describe our second deployment, out of two, which was carried out at Gorongosa National Park in Mozambique. Our future plans include using the data collected from the deployment to investigate and model certain behavioral characteristics of animal groups, which will then be used to refine our algorithms in a next development cycle (refer to Section 7 for details of future plans).

Lessons learned from our first deployment, which is not reported here, include the discovery that some camera modules were malfunctioning because of a design flaw that set the voltage of certain supply lines to levels below the minimum required for stable operation. This led us to redesign parts of our device.

3 | RELATED WORK

There is a multitude of research studies in the literature that involve the development of animal-borne systems to collect animal behavioral data. Here we briefly review some of the most relevant publications focusing on network-enabled and/or imagery monitoring systems. For further reading, we refer the reader to the review papers (Cooke et al., 2004; Krause et al., 2013; Moll, Millspaugh, Beringer, Sartwell, & He, 2007;

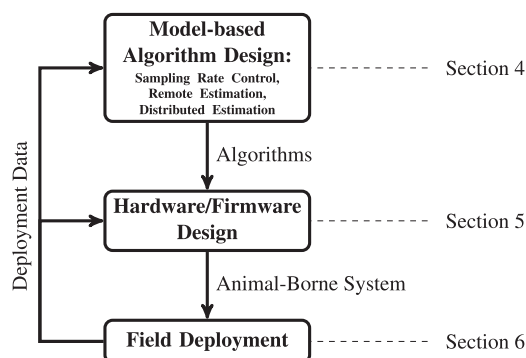


FIGURE 1 Three-phase system development cycle

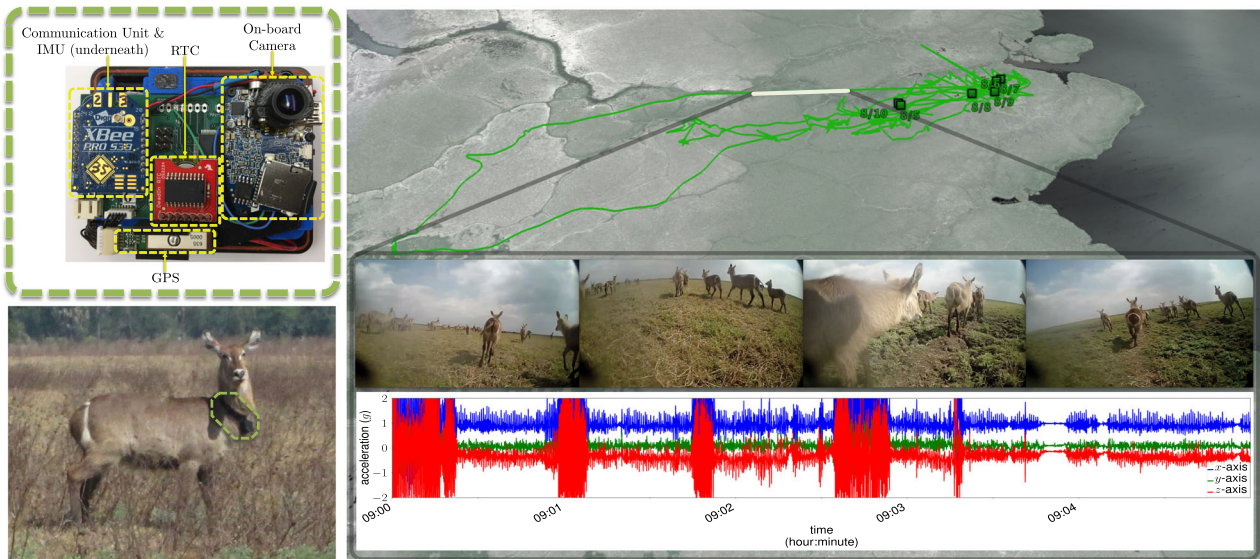


FIGURE 2 Overview of the animal-borne ecological monitoring system. Top left: Internal components of the animal-borne device. Bottom left: Waterbuck wearing the device on a collar. Right: Example of typical data collected by the device, comprising global positioning system tracks (top), frames from onboard video (middle), and triaxial accelerometer data (bottom) [Color figure can be viewed at wileyonlinelibrary.com]

Robert-Coudert & Wilson, 2005; Rutz & Hays, 2009; Tomkiewicz, Fuller, Kie, & Bates, 2010).

Location-based radiotelemetry technologies are widely adopted methods in studying community ecology. Implementation of such technologies using GPS has allowed researchers to remotely track animal locations and to understand how instrumented animals interact with group members and with surrounding environment (Cagnacci, Boitani, Powell, & Boyce, 2010). For instance, the ZebraNet project focused on designing a tracking system for wild zebras (Juang et al., 2002; Liu, Sadler, Zhang, & Martonosi, 2004; Zhang et al., 2004) in which the system adopts radiotelemetry to remotely collect GPS data at a stationary base camp. The value of data collected by GPS telemetry can be enhanced by incorporating satellite imagery, which provides useful information for studying animal–landscape interactions (Handcock et al., 2009).

A radio-frequency identification (RFID) technology that can sense animal social encounters serves as an energy-efficient data collection method for animal interaction studies (Krause, Wilson, & Croft, 2011). The work of Dyo et al. (2010) studies the behavior of European badgers using a network of RFID-based nodes, which register proximity data between stationary nodes and mobile nodes (RFID tags) collared on badgers. Encounternet (Mennill et al., 2012)—a new radiotelemetry technology—enables ultra-low-power detection of the presence of instrumented animals, equipped with a transmitter, near a preinstalled receiver. The performance of Encounternet was verified through tracking of long-tailed manakins in their natural habitats. Interanimal telemetry to register encounters

among members in a marine mammal group is also implemented in an aquatic environment (Holland, Meyer, & Dagorn, 2009). The work of Markham and Wilkinson (2008) investigates an animal-tracking system using both GPS and radiotelemetry-based proximity sensing.

The Crittercam from the National Geographic Society (Marshall, 1998) is a pioneering animal-borne imaging system for wildlife studies. In addition to sensors that can register animal motions and environmental variables, such as water temperature and pressure, the system has the capability of recording continuous images. This new powerful means of data collection has led to novel insights in the ecology and biology communities through system deployment on emperor penguins (Ponganis, Van Dam, Marshall, Knowler, & Levenson, 2000), American alligators (Nifong et al., 2014), large sharks (Heithaus, Marshall, Buhleier, & Dill, 2001), and marine mammals (Williams et al., 2000). Independently, Rutz, Bluff, Weir, and Kacelnik (2007) presented a miniature animal-borne video camera system (along with VHF radio tags for position tracking) for discovery and quantification of unexplored behavior of birds. The work of (Davis et al. (1999) presents an animal-borne video system and a data logger for studying hunting behavior of a marine mammal. Also an underwater imagery system was successfully deployed to study group behavior of penguins during underwater foraging (Takahashi et al., 2004). Although remote imaging systems provide a new dimension in animal behavioral data, they suffer from power and memory limitations and require the analysis of massive video data (Moll et al., 2007).

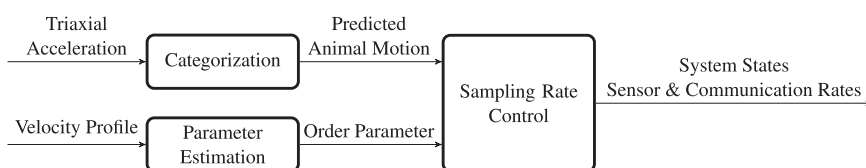


FIGURE 3 An illustration of the sampling rate control

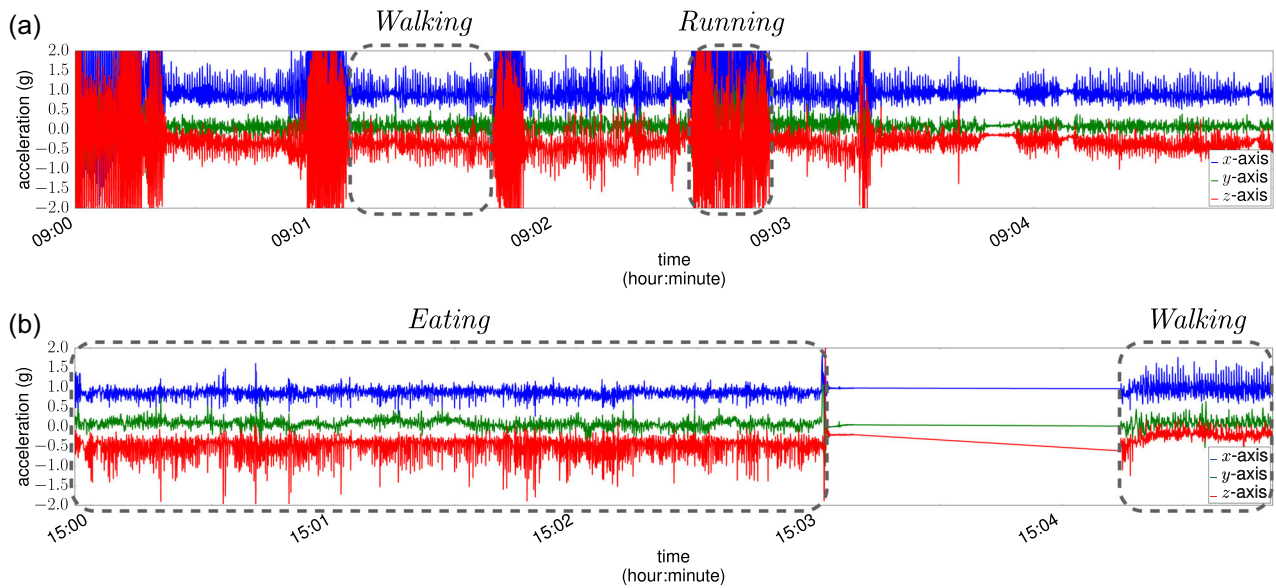


FIGURE 4 An example of triaxial acceleration data (collected during a deployment) that illustrates walking, eating, and running motions of an instrumented animal (waterbuck): (a) depicts the acceleration data collected while the animal is alternating between running and walking motions, and (b) depicts the acceleration data collected while the animal is predominantly eating. Note that in (b), no acceleration data were collected during the period between the eating and walking motions since the instrumented animal was stationary [Color figure can be viewed at wileyonlinelibrary.com]

3.1 | Comparison with existing methods

In contrast with existing work, as we discussed in Section 1, in our approach each device uses not only onboard sensing and computation, but it also utilizes information disseminated by an ad hoc wireless

network to enhance the relevance of the data collected. In addition to having to conceive a new hardware platform to make this possible, we had to introduce novel real-time algorithms that can be dependably executed subject to power, computational, and communication reliability limitations.

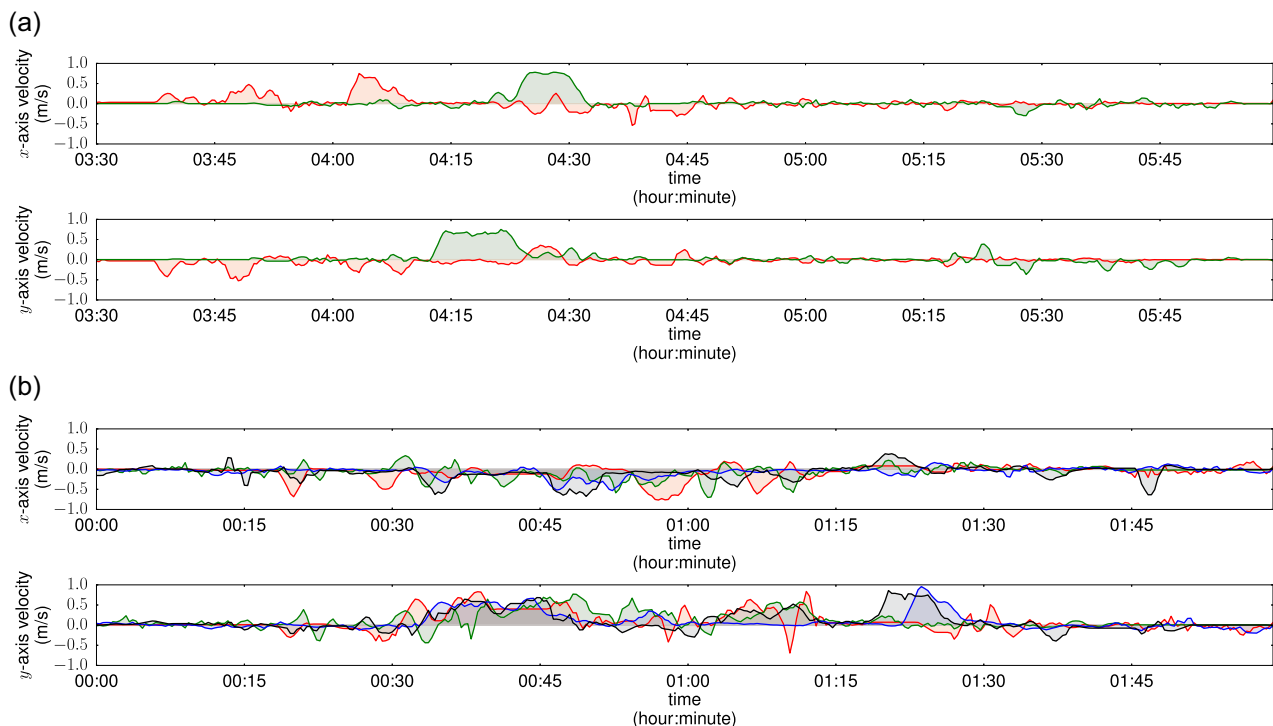


FIGURE 5 An example of the velocity profiles of two different species of animals obtained from global positioning system data: (a) depicts the velocity of two waterbucks and (b) depicts the velocity of four buffaloes [Color figure can be viewed at wileyonlinelibrary.com]

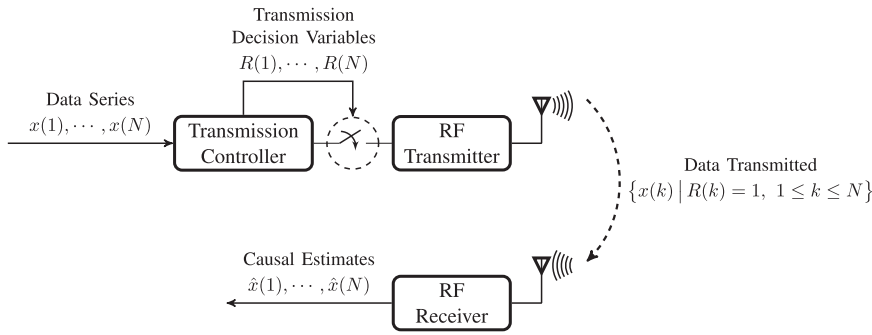


FIGURE 6 A diagram of the remote estimation algorithm. The values of transmission decision variables $R(1), \dots, R(N)$ and causal estimates $\hat{x}(1), \dots, \hat{x}(N)$ are determined according to (4)

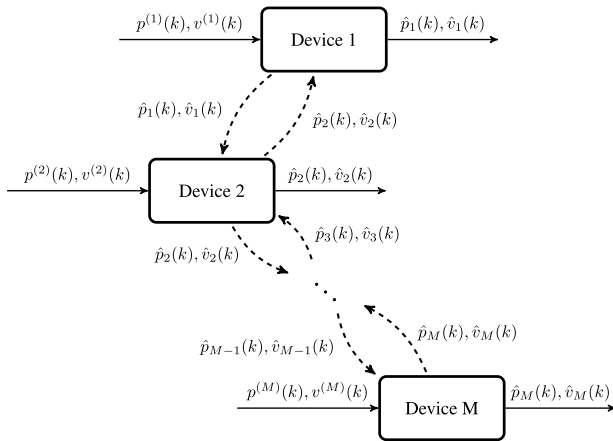
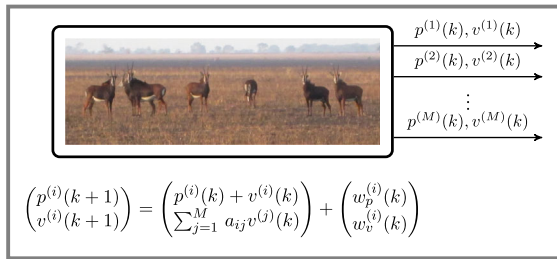


FIGURE 7 An illustration of the distributed estimation algorithm [Color figure can be viewed at wileyonlinelibrary.com]

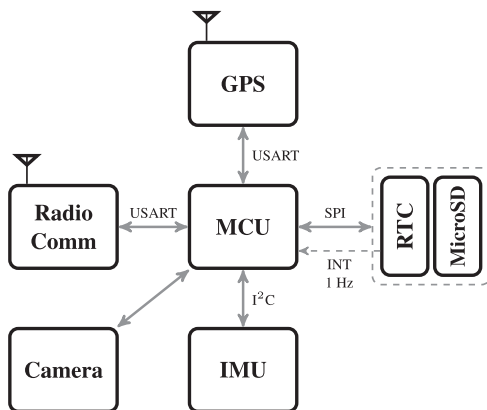


FIGURE 8 A diagram showing the main hardware components of the animal-borne device and their configuration in the system integration

4 | MODEL-BASED ALGORITHM DESIGN

To achieve the design goals listed in Section 1.1, we sought algorithms to support the following functionalities:

- (1) *Sampling rate control*: We developed algorithms to regulate the sensing and communication rates of each animal-borne device based on both individual and collective movements. The algorithms were conceived to promote adequate power-efficiency of sensing and information sharing.
- (2) *Remote estimation*: We proposed a two-block remote estimation scheme that any two devices can use to share their measurements, subject to power constraints. Here, an event-driven policy determines when a sending device ought to transmit to a receiving device, which, in turn, runs a remote estimator that optimally interpolates the data in-between transmissions.
- (3) *Distributed estimation*: We developed an algorithm that reliably fuses the partial movement data that are gathered and shared across the network by each device to enable the omniscience, at all devices, of the overall movement and locations of the animals in the group being monitored.

We now proceed to Sections 4.1, 4.2, and 4.3, where we describe the above-mentioned algorithms in detail. Later, in Section 5 we outline how the algorithms are implemented in the firmware and how they operate in concert (see Figure 9).

4.1 | Sampling rate control

We employ a sampling rate control algorithm to regulate the sensing and communication rates of each animal-borne device to match its needs while lessening power consumption.

We consider that there are four sampling modes, and that each represents an application-dependent sensing and communication rate. The sampling modes, ordered by decreasing rates, are denoted as *High*, *Alert*, *Ready*, and *Sleep*. Typically, the sensors and communication module are disabled in *Sleep* mode. Assigning appropriate rates to each mode is crucial to establish the correct trade-off between power consumption and performance.

As we explain in Section 4.1.3, the sampling mode is a function of a *predicted animal motion state* and the so-called *order parameter*,

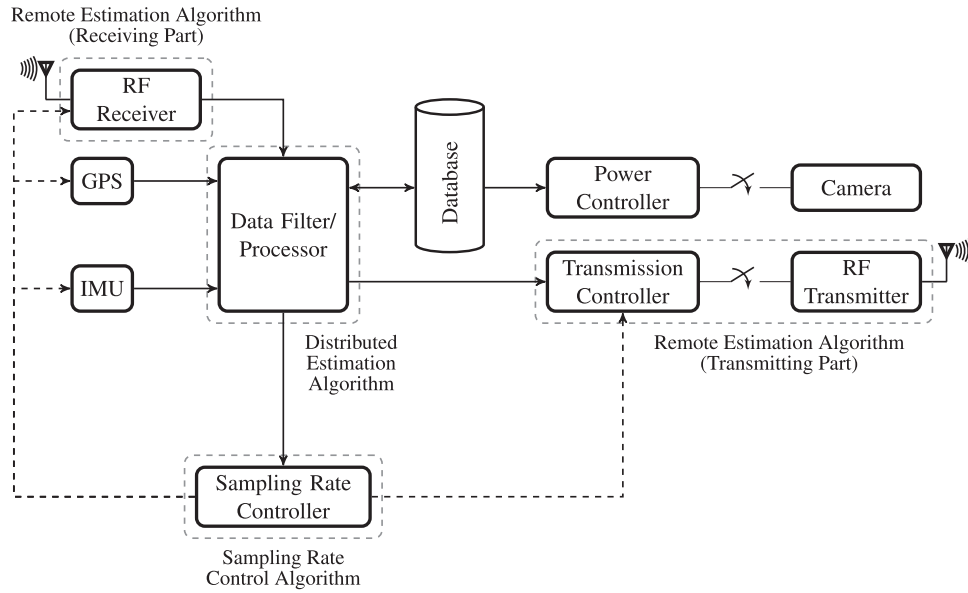


FIGURE 9 The firmware architecture of the animal-borne ecological monitoring system. The solid lines describe the direction of data flow and the dotted lines show which components in the firmware are controlled by the sampling rate controller. The dotted rectangles indicate a firmware component or a group of components that implement each algorithm described in Section 4

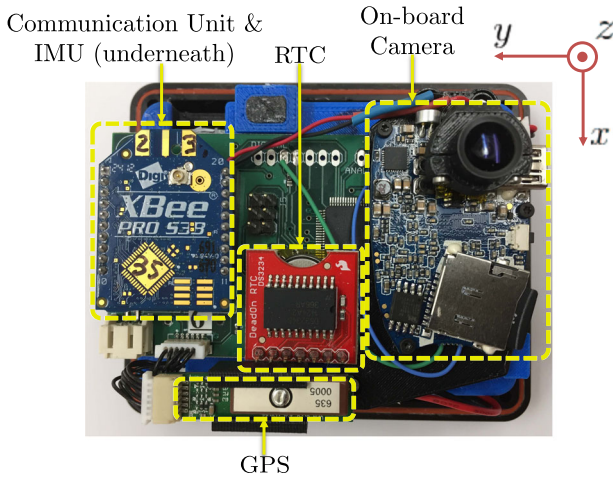


FIGURE 10 The hardware configuration of the animal-borne device [Color figure can be viewed at wileyonlinelibrary.com]

which describes the degree of alignment of the velocity profile for the animal group (see Figure 3). In Sections 4.1.1 and 4.1.2, we describe how these are obtained from the triaxial acceleration and velocity profile, respectively.

4.1.1 | Determining animal motion states

We categorize animal motion in four possible states denoted as *Resting*, *Walking*, *Eating*, and *Running*. Each state, for a given species, has a well-defined triaxial acceleration signature, which we seek to characterize to build a set-based animal motion predictor.

Before we proceed, we describe the notation used to specify our set-based animal state predictor:

- $a(k) = (a_x(k), a_y(k), a_z(k))$ —A triaxial acceleration data point at the k th discrete time step (see Figure 10 for the definition of the tri-axes on the animal-borne device).
- $\sigma(k) = (\sigma_x(k), \sigma_y(k), \sigma_z(k))$ —The standard deviation along each axis of the triaxial acceleration data points $\{a(j)\}_{j=k-m+1}^k$ corresponding to a time interval of preselected length m .

Based on $\sigma_x(k)$ and $\sigma_z(k)$, we estimate the animal motion state according to the following rule²:

$$\hat{m}(\sigma(k)) = \begin{cases} \text{Resting} & \text{if } \sigma_x(k) < \tau_x \text{ and } \sigma_z(k) < \tau_z, \\ \text{Walking} & \text{if } \tau_x \leq \sigma_x(k) < \tau'_x \text{ and } \tau_z \leq \sigma_z(k) < \tau'_z \text{ and } \sigma_x(k) + \tau_{xz} \geq \sigma_z(k), \\ \text{Eating} & \text{if } \tau_x \leq \sigma_x(k) < \tau'_x \text{ and } \tau_z \leq \sigma_z(k) < \tau'_z \text{ and } \sigma_x(k) + \tau_{xz} < \sigma_z(k), \\ \text{Running} & \text{if } \sigma_x(k) \geq \tau'_x \text{ and } \sigma_z(k) \geq \tau'_z, \\ \hat{m}(\sigma(k-1)) & \text{otherwise.} \end{cases} \quad (1)$$

More specifically, in (1), $\hat{m}(\sigma)$ denotes an estimated animal motion state given the standard deviation data σ , and τ_x , τ_z , τ'_x , τ'_z , and τ_{xz} are constants that, as we describe below, are determined using labeled acceleration data collected from a test deployment. We remark that $\sigma_y(k)$ is not present in (1) because its variation across motion states is less pronounced than that of $\sigma_x(k)$ and $\sigma_z(k)$ and, hence, its use would not

²For simple implementation and real-time execution of the algorithm on a low-power MCU, we use the variation of triaxial acceleration to identify and categorize animal motions. The reader is referred to Gurarie, Andrews, and Laidre (2009), Halsey, Green, Wilson, and Frappell (2009), Nathan et al. (2012), Shepard, Wilson, Halsey, et al. (2008), Shepard, Wilson, Quintana, et al. (2008), and Wang et al. (2015) for more sophisticated implementations of other algorithms.



FIGURE 11 The photos of study animals: (a) Waterbuck and (b) African buffalo [Color figure can be viewed at wileyonlinelibrary.com]

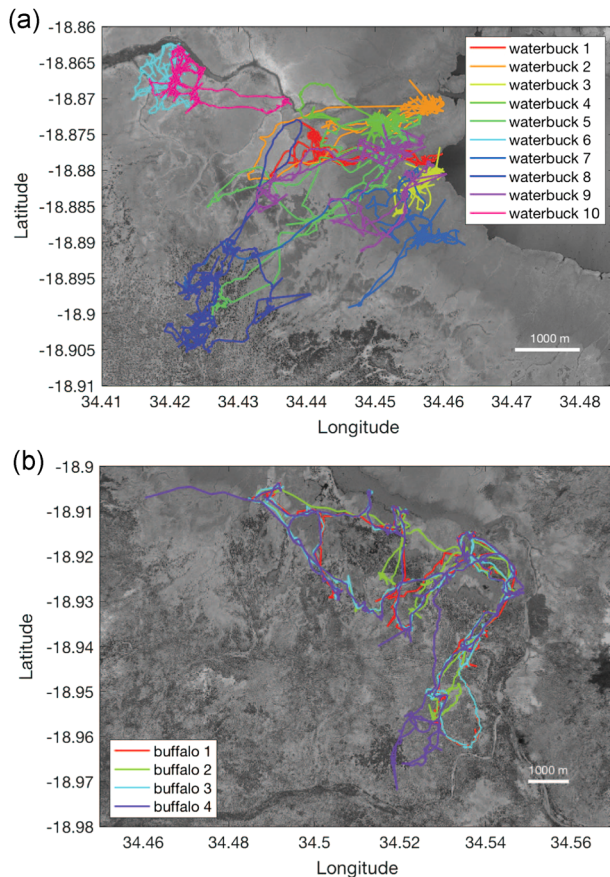


FIGURE 12 The global positioning system trajectory data of the study animals obtained during the system deployment: (a) shows the trajectories of 10 waterbucks and (b) shows the trajectories of four buffaloes [Color figure can be viewed at wileyonlinelibrary.com]

improve estimation fidelity significantly (see Figures 16a and 17a for illustrative examples).

Determining τ_x , τ_z , τ'_x , τ'_z , and τ_{xz}

Throughout system deployments, our system registers data on triaxial acceleration when video recording is activated. By cross-examining contemporaneous contiguous blocks of acceleration data and video recordings, we label each block with an animal motion

category (see Figure 4 for an example). Subsequently, we compute the standard deviation along each axis of the acceleration data within each block and apply supervised learning (Alpaydin, 2010) to find partitions defined by τ_x , τ_z , τ'_x , τ'_z , and τ_{xz} that we can use in (1) to predict each motion state from $\sigma(k)$. Notice that, since the devices are collared on the animal's neck, we can distinguish between *Walking* and *Eating*. More specifically, we can detect when an instrumented animal bends its neck downward and moves its head back-and-forth to eat. We detect this rhythmic motion using accelerometry along the z-axis relative to that of the x-axis. The details of the computation of all these constants are given in Appendix C.

4.1.2 | Determining the order parameter

We next explain a parametric method to assess collective behavior using velocity data. An example of planar velocity GPS data recorded for two different animal species is depicted in Figure 5. We use the so-called order parameter (Vicsek & Zafeiris, 2012) defined below to quantify the coherence of the movements of the animals in a group³ with M animals:

$$\psi(k) = \frac{1}{\bar{v}(k)} \left\| \sum_{i=1}^M v_i(k) \right\|_2. \quad (2)$$

Here, $v_i(k)$ is the planar (two-dimensional) velocity vector for animal i at time k , and $\bar{v}(k)$ is a normalizing factor defined as $\bar{v}(k) = \sum_{i=1}^M \|v_i(k)\|_2$. Note that $\psi(k)$ ranges between 0 and 1 for which the maximum is attained when the velocity vectors are aligned indicating the maximum degree of group motion coherence.

4.1.3 | Determining the sampling mode

Based on the motion categorization (1) and order parameter $\psi(k)$, we define system states and assign operating rates of the sensors and

³Here, we choose the parameter (2) to measure the alignment in animal group movements. For measuring other collective behavior, one may adopt other parameters such as behavioral correlation (Cavagna et al., 2010).

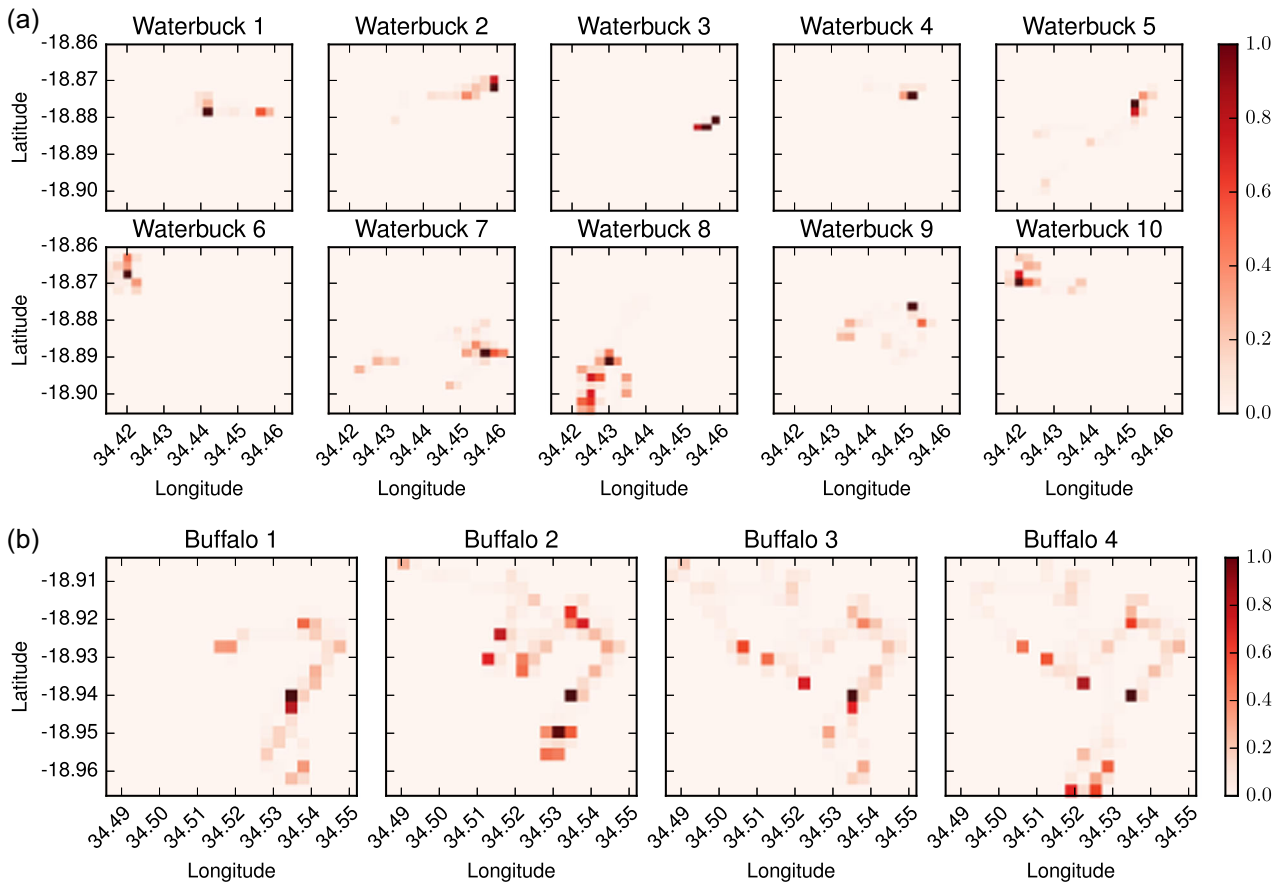


FIGURE 13 Heat maps illustrating the spatial distributions of the GPS data of each study animal depicted in Figure 12: (a) shows the spatial distributions of the GPS data of the waterbucks and (b) shows the spatial distributions of the GPS data of the buffaloes. GPS, global positioning system [Color figure can be viewed at wileyonlinelibrary.com]

communication unit. Below we provide an example of system states defined based on (1) and (2):

$$\text{SystemState} = \begin{cases} \text{Sleep} & \text{if } \hat{m}(\sigma(k)) \in \{\text{Resting}\}, \\ \text{Ready} & \text{if } \hat{m}(\sigma(k)) \in \{\text{Walking}, \text{Eating}, \text{Running}\}, \\ \text{Alert} & \text{if } \hat{m}(\sigma(k)) \in \{\text{Walking}, \text{Running}\} \text{ and } M > 1, \\ \text{High} & \text{if } \hat{m}(\sigma(k)) \in \{\text{Walking}, \text{Running}\} \text{ and } \psi(k) \geq \tau_\psi, \end{cases} \quad (3)$$

where M indicates the number of neighboring instrumented animals and τ_ψ is a preselected constant. Table 2 in Section 6.1 specifies operating rates of sensors and a communication unit associated with each state.

A choice of the parameters τ_x , τ_z , τ'_x , τ'_z , τ_{xz} for (1) and the threshold τ_ψ for (3) in the system deployment is described in Section 6.1. Using data collected from the deployment, we evaluate and depict the predicted animal motion (1) and the order parameter $\psi(k)$ in Section 6.2.

4.2 | Selective data transmission for remote estimation

To compute the order parameter (2) and to define the system states (3) using the sampling rate control algorithm, each device needs to estimate

animal movement parameters of instrumented animals within the distance-defined neighborhood. This requires data exchange between animal-borne devices. The communication unit is one of the major power-drawing hardware components in our system (see Table 1); hence frequent data transmission between devices results in rapid battery drain. For instance, to compute the order parameter $\psi(k)$ used in (3) based on the velocity profiles of neighboring instrumented animals, continuous tracking of neighbors is needed, which requires repeated transmission of GPS data between pairs of animal-borne devices. To prevent this rapid battery drain, we adopt a model-based algorithm that enables selective transmission of data from a transmitting device and remote estimation of original data at a receiving counterpart.

The essential operation of the algorithm can be described as follows. Based on a stochastic model of a data series (GPS trajectory data, for instance) to be transmitted, the transmitter makes a prediction of a receiver's causal estimate of each data point in the series. If the prediction error exceeds a preselected threshold then the transmitter sends the data to the receiver; otherwise, the transmitter does not send the data and reserves battery power (see Figure 6 for an illustration). To describe details of the algorithm, we adopt the following notation:

- $\{x(k)\}_{k=1}^N \subset \mathbb{X}$ – A series of N data points to be transmitted where each data point $x(k)$, which takes a value in a set \mathbb{X} , is available to the transmitter at the k th time step in a discrete time framework.
- $\{\hat{x}(k)\}_{k=1}^N \subset \mathbb{X}$ – Causal estimates of $\{x(k)\}_{k=1}^N$ computed by the receiver.
- $\{R(k)\}_{k=1}^N$ – Variables describing a sequence of decisions on whether to transmit the data $x(k)$ at each time k for which $R(k) = 1$ if $x(k)$ is transmitted, and $R(k) = 0$ otherwise. Also we define the *last transmission time* τ_k before time k as follows: Given $R(1), \dots, R(k-1)$,

$$\tau_k = \max\{1 \leq j \leq k-1 \mid R(j) = 1\}.$$

We follow a convention that $\tau_k = 0$ if $R(j) = 0$ holds for all j in $\{1, \dots, k-1\}$.

- $\{\mathcal{T}_k\}_{k=1}^N, \{\mathcal{E}_k\}_{k=1}^N$ – Functional forms of transmission policies $(\tau_k, x(\tau_k), x(k)) \mapsto \mathcal{T}_k(\tau_k, x(\tau_k), x(k)) \in \{0, 1\}$ and estimation rules $(\tau_k, x(\tau_k)) \mapsto \mathcal{E}_k(\tau_k, x(\tau_k)) \in \mathbb{X}$ for the transmitter and receiver, respectively, for which the following hold⁴:

$$R(k) = \mathcal{T}_k(\tau_k, x(\tau_k), x(k)), \quad (4a)$$

$$\hat{x}(k) = \mathcal{E}_k(\tau_k, x(\tau_k)). \quad (4b)$$

We assume that the underlying statistical model of $\{x(k)\}_{k=1}^N$ satisfies the Markov property. Examples of such models are a (stochastic) constant velocity model and a self-propelled particle model described below.

4.2.1 | Constant velocity model (Li and Jilkov, 2003)

The state $x(k) = (p^T(k) \ v^T(k))^T \in \mathbb{R}^2 \times \mathbb{R}^2$ evolves according to a state-space equation given by

$$p(k+1) = p(k) + v(k) + w_p(k), \quad (5a)$$

$$v(k+1) = v(k) + w_v(k), \quad (5b)$$

where $p(k)$ and $v(k)$ denote the position and velocity of a study animal, respectively, and $w_p(k)$ and $w_v(k)$ are random processes representing modeling error.

4.2.2 | Self-propelled particle model (McClintock et al., 2012)

The state $x(k) = (p^T(k) \ \theta(k))^T \in \mathbb{R}^2 \times \mathbb{T}$ evolves according to a state-space equation given by⁵

$$p(k+1) = p(k) + v(k) \begin{pmatrix} \cos(\theta(k) + \phi(k)) \\ \sin(\theta(k) + \phi(k)) \end{pmatrix}, \quad (6a)$$

$$\theta(k+1) = \theta(k) + \phi(k), \quad (6b)$$

where $p(k)$ and $\theta(k)$ denote the position and orientation of a study animal in the two-dimensional plane, respectively. The random processes $v(k)$ and $\phi(k)$ represent the translational and angular velocities, respectively.

The goal is to design transmission policies $\{\mathcal{T}_k\}_{k=1}^N$ and estimation rules $\{\mathcal{E}_k\}_{k=1}^N$ that regulate the transmission of the data series $\{x(k)\}_{k=1}^N$ from the transmitter and the computation of estimates $\{\hat{x}(k)\}_{k=1}^N$ at the receiver according to (4), respectively. Finding $\{\mathcal{T}_k\}_{k=1}^N$ and $\{\mathcal{E}_k\}_{k=1}^N$ that achieve the optimal trade-off between the estimation distortion and data transmission cost can be formulated as an optimization problem. Below, we provide two examples of performance indices to be optimized for Models (5) and (6), respectively:

$$\sum_{k=1}^N \mathbb{E} \left[\underbrace{\|p(k) - \hat{p}(k)\|_2^2 + \|v(k) - \hat{v}(k)\|_2^2}_{\text{squared estimation distortion}} + \underbrace{c_k R(k)}_{\text{transmission cost}} \right], \quad (7)$$

subject to (4), (5),

$$\sum_{k=1}^N \mathbb{E} \left[\underbrace{\|p(k) - \hat{p}(k)\|_2^2 + 4(1 - \cos(\theta(k) - \hat{\theta}(k)))}_{\text{squared estimation distortion}} + \underbrace{c_k R(k)}_{\text{transmission cost}} \right], \quad (8)$$

subject to (4), (6), where $(\hat{p}(k), \hat{v}(k))$ is an estimate of $(p(k), v(k))$, $(\hat{p}(k), \hat{\theta}(k))$ is an estimate of $(p(k), \theta(k))$, and c_k is a positive constant denoting the cost incurred by each transmission of data. Computational procedures for finding optimal solutions to (7) and (8) are described in (Lipsa & Martins, 2011; Park & Martins, 2014, 2016). We remark that these procedures are not designed for real-time execution, but need to be carried out before system deployment. However, the evaluation of (4) is executable in real-time, and hence so is the remote estimation algorithm described here.

A choice of the data transmission cost c_k in the system deployment is described in Section 6.1. Based on the data collected, we assess the performance of the remote estimation algorithm using various choices of c_k in Section 6.2.

4.3 | Distributed estimation

When the animal-borne device detects coherent group movement, for example, $\psi(k) \geq \tau_\psi$, within its neighborhood, it would trigger video recording. To execute the triggering based on joint configurations of study animals, each device needs to estimate the locations and velocities of all instrumented animals in the system's communication network. For this purpose, we design a model-based distributed algorithm to allow all the devices to estimate these variables. The algorithm operates over ad hoc communication

⁴Notice that according to the definitions of \mathcal{T}_k and \mathcal{E}_k , the transmission policy \mathcal{T}_k depends on the last transmission time τ_k , the data $x(\tau_k)$ transmitted at time τ_k , and the data $x(k)$ at time k , and the estimation rule \mathcal{E}_k depends on the last transmission time τ_k and the data $x(\tau_k)$ transmitted at time τ_k .

⁵The set \mathbb{T} represents the circle group.



FIGURE 14 Screenshots of video data: (a) shows a 5-min video captured during 8/5/2015, 15:00–15:05, from the device collared on Waterbuck 1 while the waterbuck is eating. (b) shows a 5-min video captured during 8/6/2015, 9:00–9:05, from the device collared on Waterbuck 2. The waterbucks on the video are traveling away from their core areas. (c) shows a 5-min video captured during 8/13/2015, 5:00–5:05, from the device collared on Buffalo 2. The buffalo herd on the video is marching along a common trail one after another [Color figure can be viewed at wileyonlinelibrary.com]

networks, and naturally realizes a feedback mechanism in estimation over networks. In this sense, it is fundamentally different from broadcasting-based communication protocols presented, for instance, in Juang et al. (2002).

To describe the algorithm, we adopt the following notation:

- M —The number of instrumented animals residing in the system's communication network.
- $p^{(i)}(k), v^{(i)}(k), p(k), v(k)$ —The variables $p^{(i)}(k)$ and $v^{(i)}(k)$ represent the position and velocity of the i th instrumented animal at the k th time step in a discrete time framework, respectively. We denote by $p(k)$ and $v(k)$ the concatenations of all the variables $p^{(i)}(k)$ and $v^{(i)}(k)$, respectively.
- $\hat{p}_i(k), \hat{v}_i(k)$ —Estimate of $p(k), v(k)$ by the i th animal-borne device.
- $\mathbb{N}_i(k)$ —The set of devices that are within the communication range of device i (including i itself) at time k .
- e_i, I_2 —The i th standard basis vector in \mathbb{R}^M and the 2D identity matrix, respectively.
- \otimes —The Kronecker product.

Consider a collective movement model for M animals that describes the time evolution of the position $p^{(i)}(k)$ and velocity $v^{(i)}(k)$ of the i th animal according to the following state-space equation:

$$p^{(i)}(k+1) = p^{(i)}(k) + v^{(i)}(k) + w_p^{(i)}(k), \quad (9a)$$

$$v^{(i)}(k+1) = \sum_{j=1}^M a_{ij} v^{(j)}(k) + w_v^{(i)}(k), \quad (9b)$$

where $w_p^{(i)}(k)$ and $w_v^{(i)}(k)$ are zero-mean random processes representing modeling error, and $\{a_{ij}\}_{i,j=1}^M$ are nonnegative constants representing velocity couplings and satisfying $\sum_{j=1}^M a_{ij} = 1$ $a_{ij} > 0$ for all i in $\{1, \dots, M\}$. Note that Model (9) can be viewed as a variant of the Vicsek model (Vicsek, Czirók, Ben-Jacob, Cohen, & Shochet, 1995).

The recursions in (9) suggest that the velocity of each member in an animal group is determined as a convex combination of the velocities of other group members. When the animal movements are all decoupled, one could select the constants $\{a_{ij}\}_{i,j=1}^M$ as follows: For all i, j in $\{1, \dots, M\}$, it holds that

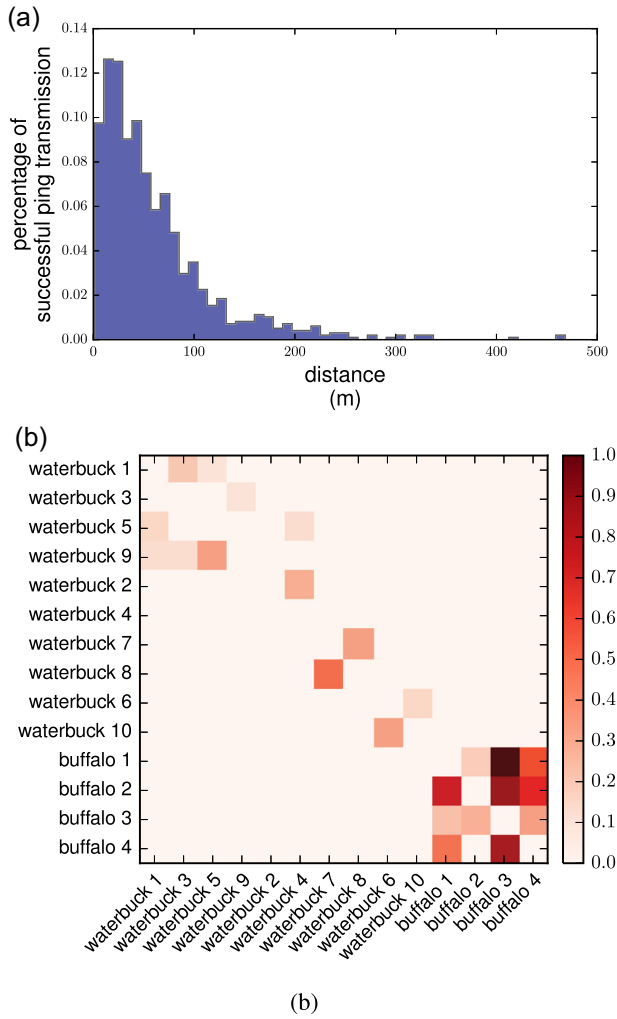


FIGURE 15 (a) Percentage of the success in ping message transmission with respect to the distance between associated pairs of (transmitting and receiving) devices. (b) Visualization of the distribution on message exchanges between every pair of the deployed devices [Color figure can be viewed at wileyonlinelibrary.com]

$$a_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

Suppose that device i can measure the position $p^{(i)}(k)$ and velocity $v^{(i)}(k)$ of its associated instrumented animal and it can communicate with neighboring devices denoted by $\mathbb{N}_i(k)$. The goal is to compute an estimate $\hat{p}_i(k)$, $\hat{v}_i(k)$ of $p(k)$, $v(k)$ at every device i . We adopt the following recursive update rule to compute $\hat{p}_i(k)$, $\hat{v}_i(k)$: For each l in $\{1, \dots, N_c\}$,

$$\hat{p}_i[l] = \sum_{j \in \mathbb{N}_i(k)} \frac{1}{|\mathbb{N}_i(k)|} (\hat{p}_j[l-1] + K_i^p (p^{(i)}(k) - C_i \hat{p}_j[l-1])), \quad (11a)$$

$$\hat{v}_i[l] = \sum_{j \in \mathbb{N}_i(k)} \frac{1}{|\mathbb{N}_i(k)|} (\hat{v}_j[l-1] + K_i^v (v^{(i)}(k) - C_i \hat{v}_j[l-1])), \quad (11b)$$

with the initial and final conditions, respectively, satisfying

$$\begin{aligned} \hat{p}_i[0] &= \hat{p}_i(k-1) + \hat{v}_i(k-1), & \hat{p}_i(k) &= \hat{p}_i[N_c], \\ \hat{v}_i[0] &= A \hat{v}_i(k-1), & \hat{v}_i(k) &= \hat{v}_i[N_c], \end{aligned}$$

where N_c is a positive integer denoting the number of intermediate updates by (11). The variable A is a $2M \times 2M$ matrix for which its i, j th block element equals $a_{ij} I_2 C_i = e_i^T \otimes I_2$, and K_i^p, K_i^v are gain parameters that need to be determined. See Figure 7 for an illustration of the algorithm.

In (11), we allow multiple intermediate updates between two discrete time steps $k-1$ and k . The estimate $\hat{p}_i(k)$, $\hat{v}_i(k)$ is equal to $\hat{p}_i[N_c]$, $\hat{v}_i[N_c]$ obtained after the N_c intermediate updates. In Section 6.2, using deployment data, we assess the effect of the number of updates N_c on the estimation performance.

Notice that each device i exchanges its estimate $\hat{p}_i[l]$, $\hat{v}_i[l]$ with its neighboring devices denoted by $\mathbb{N}_i(k)$, which can change with time k . This restricts data exchange between animal-borne devices and imposes a communication constraint on the estimation algorithm (11).⁶ In addition, each device's estimate depends both on its past estimates and the estimates of its neighbors, defined by neighborhood structure $\mathbb{N}_i(k)$. Hence, the estimation distortion at one device propagates and *feeds back* into those at other devices in the system's communication network.

In the literature on distributed estimation, numerous existing works focus on establishing convergence properties of (11) and finding the parameters K_i^p, K_i^v that achieve convergence. In Theorem 1 given in Appendix B, we provide a convergence result for (11). The reader is referred to (Park & Martins, 2017) and references therein for further convergence analyses on a wide class of models in distributed estimation.

5 | HARDWARE AND FIRMWARE DESIGN

In this section, we describe the hardware and firmware design choices for our animal-borne ecological monitoring system. We first describe the role of each hardware component and its configuration in the system integration. Then we explain a firmware architecture designed for controlling hardware components and implementing the algorithms described in Section 4.

5.1 | Hardware design

In our hardware platform, we have embedded a GPS and an IMU to collect movement and motion data, and a radio communication unit to sustain the system's communication network for data exchange. Each device has an MCU equipped with a micro-SD card for processing and storing data. A miniature camera is connected to the MCU which is used to capture animal point-of-view video data. The diagram in Figure 8 depicts the hardware platform and

⁶The data exchange between animal-borne devices is restricted by the set $\mathbb{N}_i(k)$ which is defined by the communication range of device i and, unlike the remote estimation algorithm described in Section 4.2, the distributed estimation algorithm itself does not regulate the transmission of data between animal-borne devices.

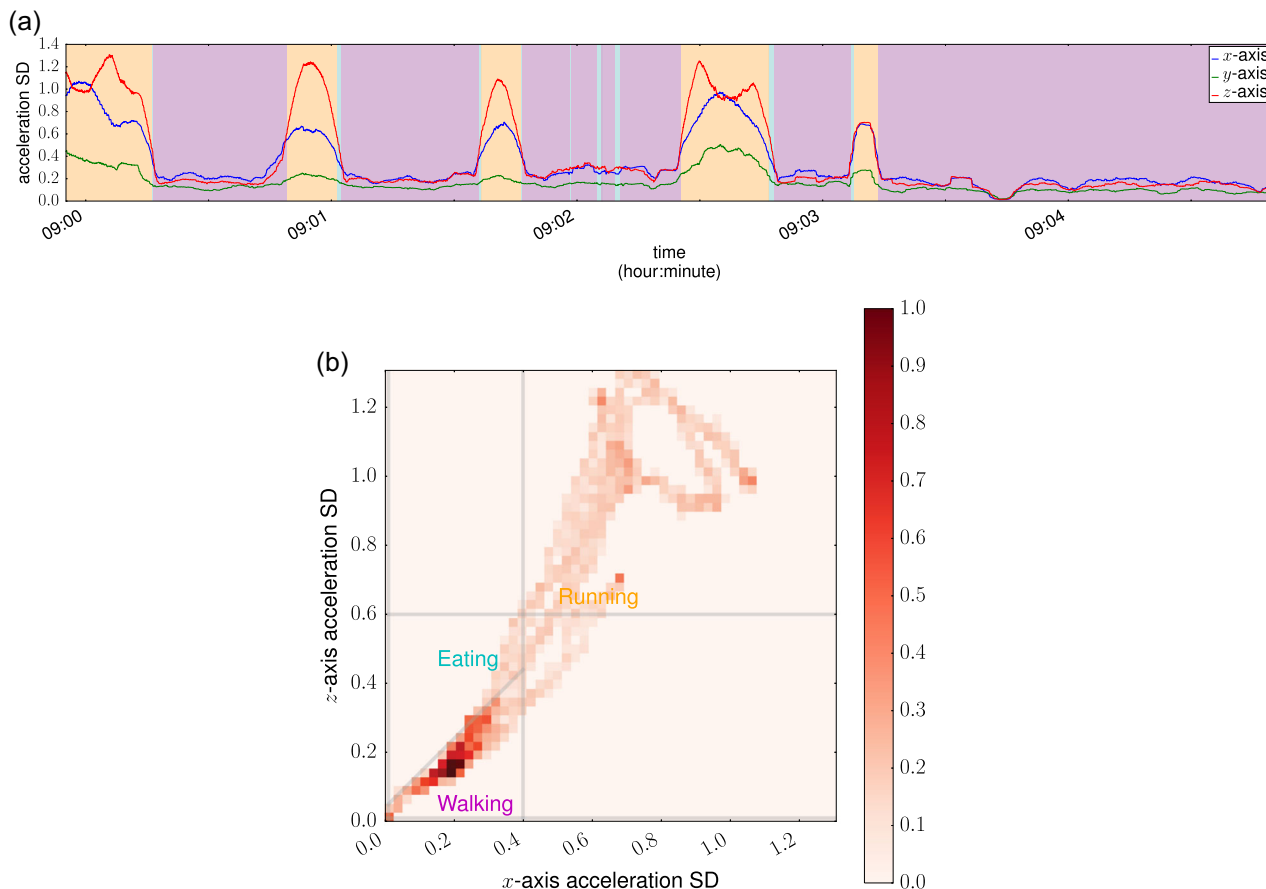


FIGURE 16 The SD and an associated heat map of acceleration data collected during 8/6/2015, 09:00–09:05, from the device collared on Waterbuck 2 (see Figure 4a for the plot of the acceleration data): (a) the SD of the data and (b) its associated heat map for which the two axes represent the SD of the acceleration data along the x- and z-axis. For the motion prediction rule (1), we use the parameters $\tau_x = \tau_z = 0.01$, $\tau'_x = 0.4$, $\tau'_z = 0.6$, $\tau_{xz} = 0.04$. According to the rule (1) (depicted by the gray lines in (b)) and resulting plot (b), the SD plot (a) is color-coded by time interval with respect to predicted animal motions {Walking, Eating, Running}, where the rule yields 88% prediction accuracy. SD, standard deviation [Color figure can be viewed at wileyonlinelibrary.com]

configuration of each component in the platform. Below we briefly describe the key specifications and configurations of hardware components in the system design. Table 1 lists details of each component including its power consumption.

Atmega2560V-8AU is a low-power eight-bit MCU operating at 8 MHz. It has 256 kB of FLASH, 4 kB of EEPROM, and 8 kB of internal SRAM. Firmware is stored in FLASH, and static/dynamic variables used are stored in SRAM. EEPROM is nonvolatile memory which can be used to store system parameters such as sensor calibration parameters. The external micro-SD card provides sufficient memory space to log sensor readings and routine system operation messages. Our firmware implements a buffer of size 512 bytes that temporarily stores data before it is written on the micro-SD card. This prevents continuous writing of data on the micro-SD card which would result in enormous power drain. The MCU consumes 15 mA at 3.3 V which is power-efficient for general computing purposes. However, this would add up to 1.8 Ah for a 5-day deployment mission which is a significant power draw for our purpose. To maintain low power consumption, we put the MCU in a power-down mode whenever it is idle, and use an external real-time

clock (RTC) unit to wake up the MCU every 1 s using the interrupt mechanism. The MCU will then execute preprogrammed routines and may go back to a power-down mode. Note that compared to that of the MCU, the power consumption of the RTC unit is small (see Table 1).

XBee-PRO 900 XSC S3B is used as a radio communication unit (a transceiver). The unit is designed to provide ad hoc networking and peer-to-peer communication capabilities, and it provides a CSMA/CA mechanism to sense (communication) channel use and to avoid packet collisions in data transmission. Also the unit has approximately 1.5 kB of buffer to temporarily store data coming through its receiver before the data are forwarded to the MCU via USART.

GP-735 is a GPS receiver that is based on u-blox's seventh generation chip. The GPS unit delivers the latitude and longitude coordinates, speed over ground, and course over ground information of instrumented animals at 1 Hz. In addition, it provides the Universal Coordinated Time which is used to synchronize internal clocks of animal-borne devices.

The IMU consists of an accelerometer (ADXL345), a digital compass (HMC5883L), and a gyroscope (ITG-3200). The

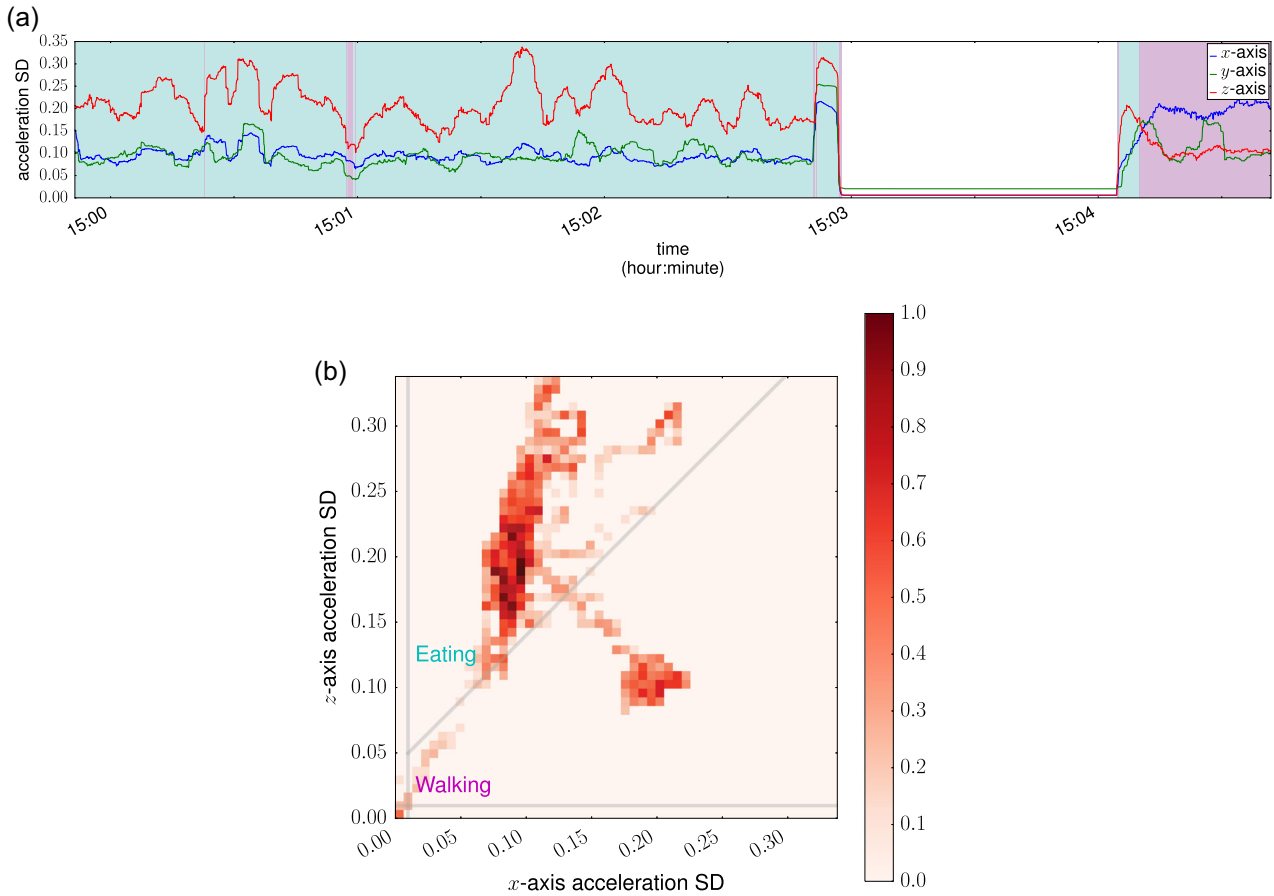


FIGURE 17 The SD and an associated heat map of acceleration data collected during 8/5/2015, 15:00–15:05, from the device collared on Waterbuck 1 (see Figure 4b for the plot of the acceleration data): (a) the SD of the data; and (b) its associated heat map for which the two axes represent the SD of the acceleration data along the x- and z-axis. For the motion prediction rule (1), we use the parameters $\tau_x = \tau_z = 0.01$, $\tau'_x = 0.4$, $\tau'_z = 0.6$, $\tau_{xz} = 0.04$ (same as in Figure 16). According to the rule (1) (depicted by the gray lines in (b)) and resulting plot (b), the SD plot (a) is color-coded by time interval with respect to predicted animal motions {Walking, Eating}, where the rule yields 97% prediction accuracy. SD, standard deviation [Color figure can be viewed at wileyonlinelibrary.com]

accelerometer can detect and record the event of its acceleration reading crossing a preselected threshold on its own. This mechanism is useful in activating the MCU from a power-down mode in the event of sudden change in animal movement.

TABLE 1 Model and power consumption of hardware components

H/W component	Model	Power consumption
MCU	Atmega2560V-8AU	15 mA
GPS	GP-735	37 mA
Accelerometer	ADXL345	0.1 mA
Magnetometer	HMC5883L	0.1 mA
Gyroscope	ITG-3200	7 mA
Radio communication	XBee-PRO 900 XSC S3B	30 mA (RX)/215 mA (TX)
RTC	DS3234	0.4 mA
Micro-SD	N/A	100 mA (RD/WR)
Camera	N/A	250 mA

Abbreviation: GPS, global positioning system; MCU, microcontroller unit; RTC, real-time clock.

5.2 | Firmware design

We design the firmware based on a functional architecture as depicted in Figure 9 which is used to implement algorithms and to control hardware components. To address the main technical challenges described in Section 2, we implement the firmware that has the following functionalities:

- Collect movement and motion data concurrently when instrumented animals are in motion.
- Periodically transmit ping messages to discover neighboring devices.
- Allow the hardware components—especially the sensors and communication unit—to operate at various rates.
- Locally store a short history of sensor data received from neighboring devices.

We briefly explain the main components of our firmware architecture in connection with the above functionalities and the algorithms described in Section 4.

TABLE 2 System parameters for the deployment in Gorongosa National Park

State	GPS rate	Accelerometer and digital compass rates (Hz)	Gyroscope rate	Ping rate	Data transmission rate	Power consumption (mA)
Sleep	1 fix/600 s	50	Off	No ping	No transmission	6
Ready	1 fix/400 s	50	Off	1 ping/400 s	No transmission	20
Alert	1 fix/20 s	50	50 Hz	1 ping/200 s	≤1 transmission/20 s	≤70
High	1 fix/20 s	50	50 Hz	1 ping/200 s	≥1 transmission/20 s	≥70

Abbreviation: GPS, global positioning system.

The data filter/processor processes incoming data from the IMU, GPS, and RF receiver. In particular, this module implements the distributed estimation algorithm described in Section 4.3, which uses data from the GPS and RF receiver to estimate the movement of instrumented animals within the system's communication network. Processed data by the data filter/processor are then stored in the database. We have assigned a fixed amount of memory space for the database and implemented circular linked lists for data storage. The data stored in the database are accessible by the power controller, which triggers video recording of the onboard camera upon the detection of collective animal behavior.

The transmission controller, RF transmitter, and RF receiver realize the remote estimation algorithm described in Section 4.2. The transmission controller identifies an underlying statistical model of the data series to be transmitted and executes data transmission via the RF Transmitter based on the transmission policies (4a). Upon data reception, the RF Receiver of a receiving device then estimates the values of the original data series based on the estimation rules (4b) and received data. In addition, the transmission controller and RF transmitter send out ping messages at a preselected rate for neighborhood discovery; upon the reception of a ping message, the transmitter controller registers the sender's ID in the list of neighboring devices.

The sampling rate controller (SRC) defines multiple system operating states and implements the sampling rate control algorithm described in Section 4.1. Each state is assigned the sensing rates of the GPS and IMU, and the communication rate, and the algorithm uses triaxial acceleration data and neighboring devices' movement data to infer individual motions and collective animal movements. This information is used to enable state transition of the system according to (3) and adaptively change the operating rates of system components.

6 | FIELD DEPLOYMENT EXPERIENCE AND SYSTEM PERFORMANCE MEASURE

In this section, we present deployment results and discuss the system performance assessed using data collected. We deployed our system in collaboration with staff at Gorongosa National Park in Mozambique. During the deployment, each animal-borne device was attached to a collar which was securely fastened around the animal's neck. See Figures 10 and 11 for photos of the device used

in the deployment and our study animals, respectively. The primary purpose of the Mozambique deployment was to test the system we designed and to collect animal behavioral data. Deployment data will be used to build behavioral models of study animals, and these models will in turn be used to improve our system design in a next system development cycle.⁷ A video describing the system operation and deployment in Gorongosa National Park is included in the online version of this paper (see Appendix A).

6.1 | System configuration

The deployment was conducted in August 2015 where 15 devices were collared on two different species: large antelope called waterbucks and African buffaloes, where the capture and animal-handling for the collaring process were performed strictly by staff of Gorongosa National Park. Except for one device that failed to start, the other 14 devices were fully operational throughout the deployment period. The devices registered GPS and IMU data along with animal point-of-view videos during 5- to 10-day deployment periods.

We defined four states—Sleep, Ready, Alert, and High—for the SRC implementation as in (3) where the parameters of (1) and (3) are determined as $\tau_x = \tau_z = 0.01$, $\tau_{xz} = 10$, and $\tau_\psi = 0.5$. We assigned a large number to the parameter τ_{xz} to categorize an animal's eating motion as *Walking*, as, in this deployment, we were aiming at collecting any data (including video recordings) while the study animals are moving and/or motioning in a group. Table 2 lists the operating rates of the sensors and communication unit along with the device's projected power consumption at each state.

In Alert state, each device updates the locations and velocities of its neighboring instrumented animals every 20 s using data transmitted from its neighbors. Every 200 s the device receives ping messages from its neighbors, which contain the location and velocity data of transmitting devices, and initializes the remote estimation algorithm with the transmitting devices. Using the algorithm, between two consecutive ping message transmissions, the device computes the locations and velocities of its neighboring instrumented animals. For the implementation of the algorithm, we adopted the

⁷We present preliminary insights on the behavior of study animals that are immediately obtained from deployment data. In-depth investigation of the behavioral characteristics of study animals will be published in separate papers.

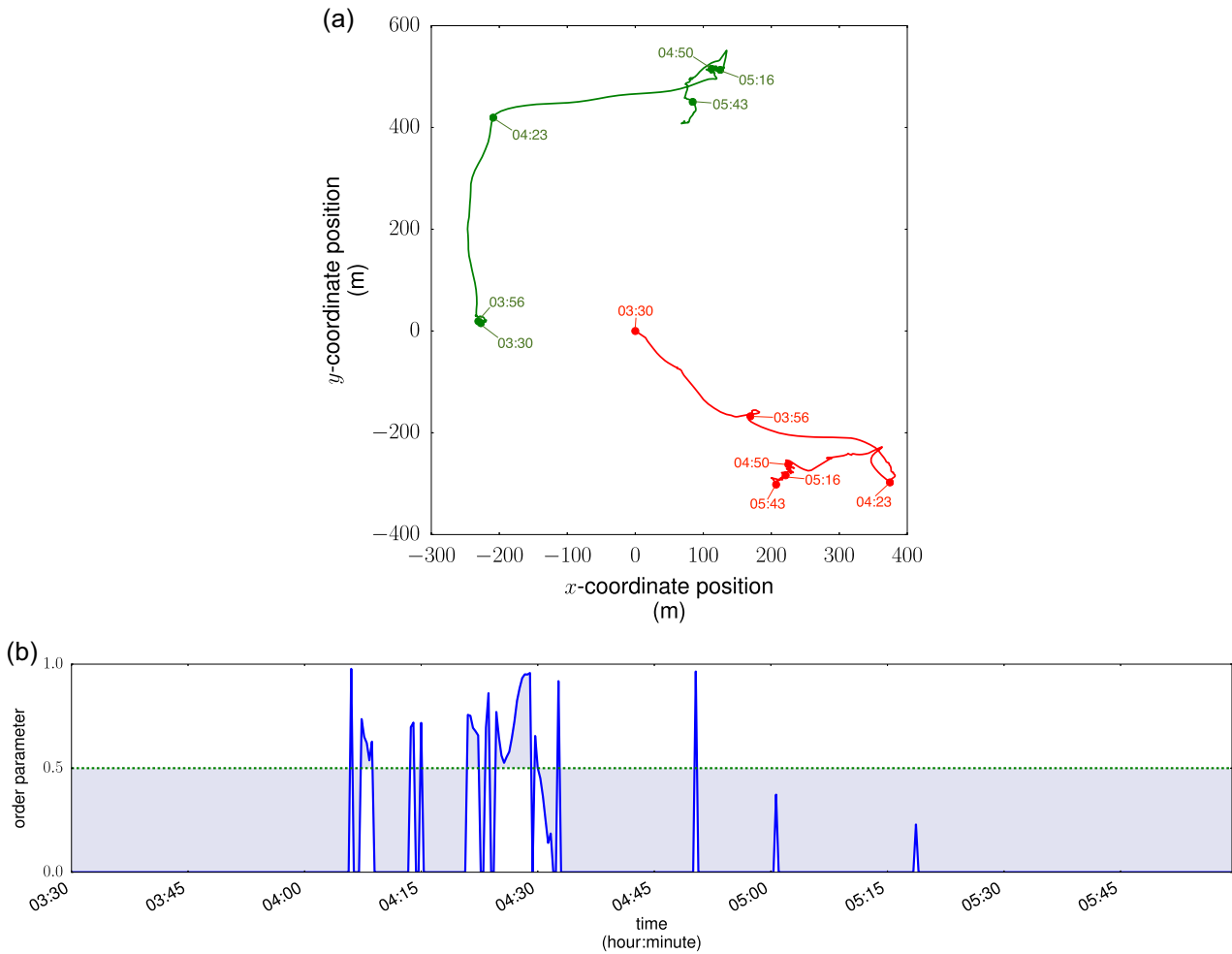


FIGURE 18 (a) Global positioning system trajectory data collected during 8/6/2015, 3:30–6:00, from the devices collared on Waterbucks 6, 10 and (b) the resulting order parameter [Color figure can be viewed at wileyonlinelibrary.com]

self-propelled particle model (6) and selected a constant transmission cost $c_k = 10$ for the performance index (8).

Upon the detection of the order parameter $\psi(k)$ defined in (2) crossing the threshold $\tau_\psi = 0.5$, the SRC transitions the system's state from Alert to High and, based on a predefined time and a 40 m proximity threshold, triggers video recording of the onboard camera. To reduce the risk of software-induced failure during the costly deployment reported in this article, we adopted the aforementioned simple proximity threshold, which relies only on the positions of each device relative to those of others in their immediate proximity. Consequently, to further reduce the risk of failure, we decided to not run our distributed estimation algorithm during the deployment because it was not needed as proximity-triggering does not rely on each device being able to resolve the group movements of all the other instrumented animals, as would other more complex triggering strategies that we intend to use in future work. Instead, we have carried out the performance evaluation of our distributed estimation algorithm, as reported in Section 6.2, after the deployment using real data downloaded from the devices.

6.2 | Analysis of deployment data and assessment of system performance

The entire GPS trajectory data of the waterbucks and African buffaloes we tracked are charted in Figure 12, and Figure 13 depicts the associated spatial distributions. A few representative screenshots of the video recorded are displayed in Figure 14. Inspection of these data suggests that, during the deployment, the waterbucks remain in a limited core area (indicated by the dark red areas in Figure 13a), with occasional excursions (possibly in a group), but generally also returning to this core. In contrast, the buffaloes seem to move more continuously and broadly across the landscape (see Figure 13b), and tend to use a common trail system (Couzin & Krause, 2003) and follow a certain path repeatedly in a group.

We noticed that the failure rate in transmissions between the devices increases significantly with distance as evidenced in Figure 15a, according to which only less than two out of one hundred ping attempts are successful for distances beyond one hundred meters. This analysis required cross-examining the logs of all ping transmissions and the GPS location data, both retrieved from the devices after the deployment.

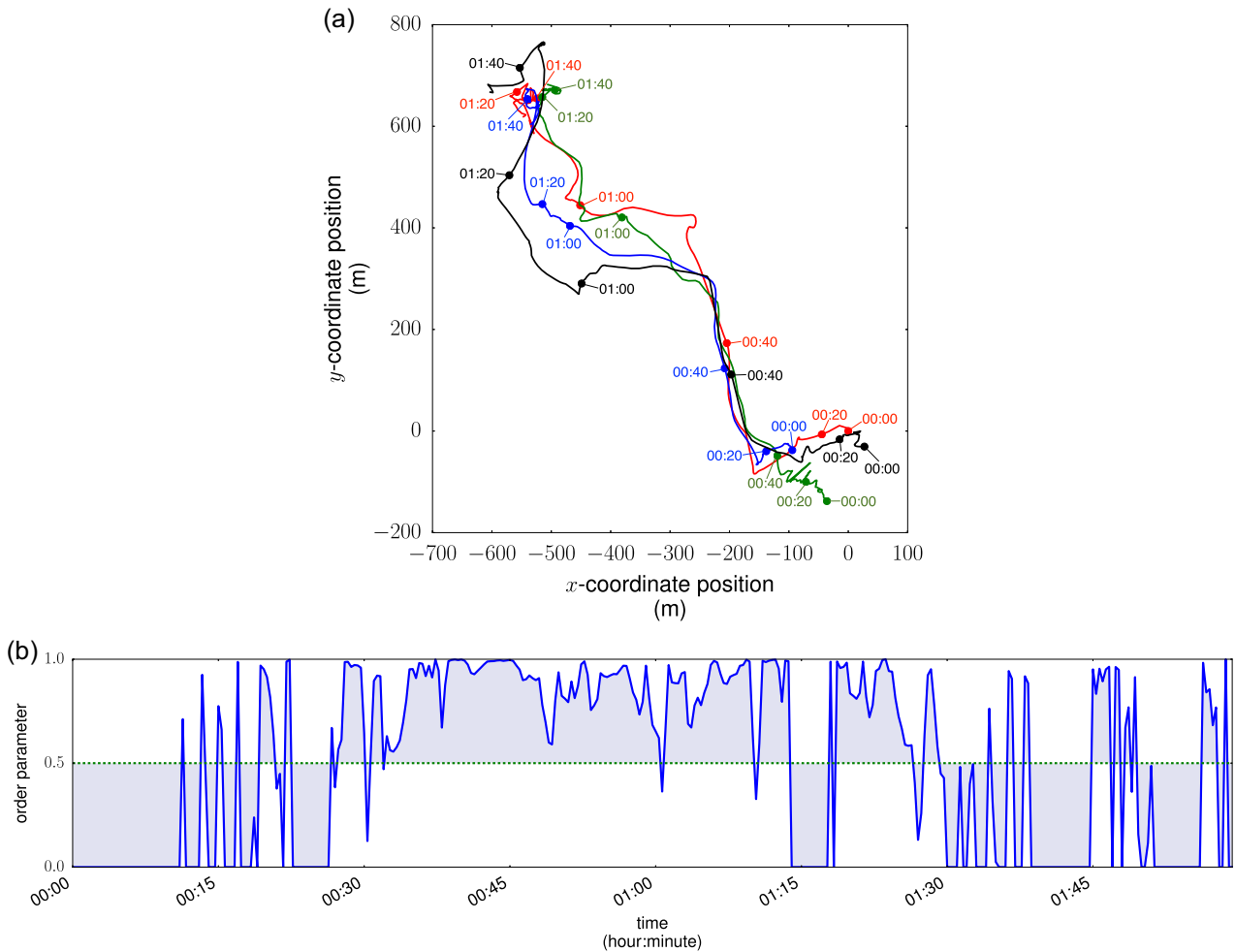


FIGURE 19 (a) Global positioning system trajectory data collected during 8/8/2015, 00:00–02:00, from the devices collared on Buffaloes 1, 2, 3, 4 and (b) the resulting order parameter [Color figure can be viewed at wileyonlinelibrary.com]

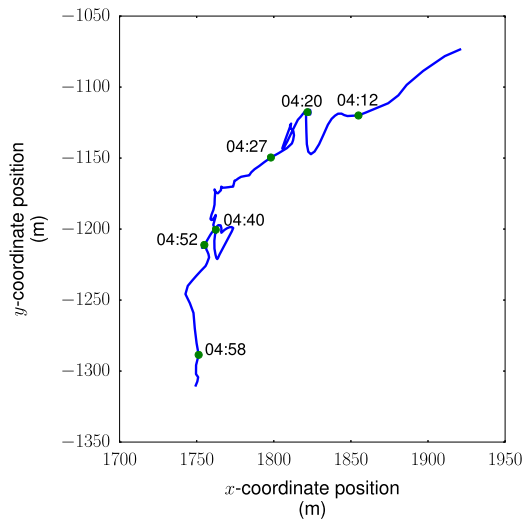


FIGURE 20 Global positioning system trajectory data collected during 8/6/2015, 04:10–05:00, from the device collared on Buffalo 1 [Color figure can be viewed at wileyonlinelibrary.com]

From this, we also found that, in regard to communication among devices mounted in the same species, as illustrated in Figure 15b, the devices collared on the buffaloes tend to exchange more messages than those collared on the waterbucks. In conjunction with the GPS data illustrated in Figure 13, these data support our observation that the instrumented buffaloes tend to travel in a group whereas the instrumented waterbucks tend to be spatially dispersed over a wide area.

Based on the data collected, we examine the performance of the three algorithms adopted in the system implementation. First of all, to verify the performance of the SRC algorithm (described in Section 4.1), we compute the standard deviation of the acceleration data along the x, y, and z-axis over every 5-s time window, and based on the prediction rule (1), we predict animal motions, which are depicted in Figures 16 and 17. Under the parameter choice described in the both figures, the prediction rule (1) achieves 88% and 97% prediction accuracy with the data depicted in Figures 16 and 17, respectively. Based on our analysis, we observed that the prediction performance depends on how frequently an instrumented animal changes its motions.

Figures 18 and 19 depict portions of GPS trajectory data of waterbucks and buffaloes, respectively, and the resulting order

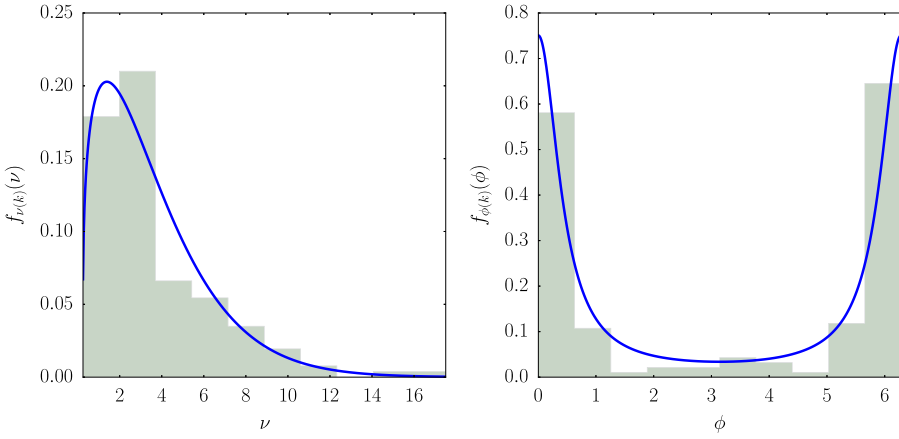


FIGURE 21 The comparisons between the probability density functions of $\nu(k)$ and $\phi(k)$ in (12) under the parameter choice (13) and the (normalized) histograms obtained from the global positioning system data depicted in Figure 20 [Color figure can be viewed at wileyonlinelibrary.com]

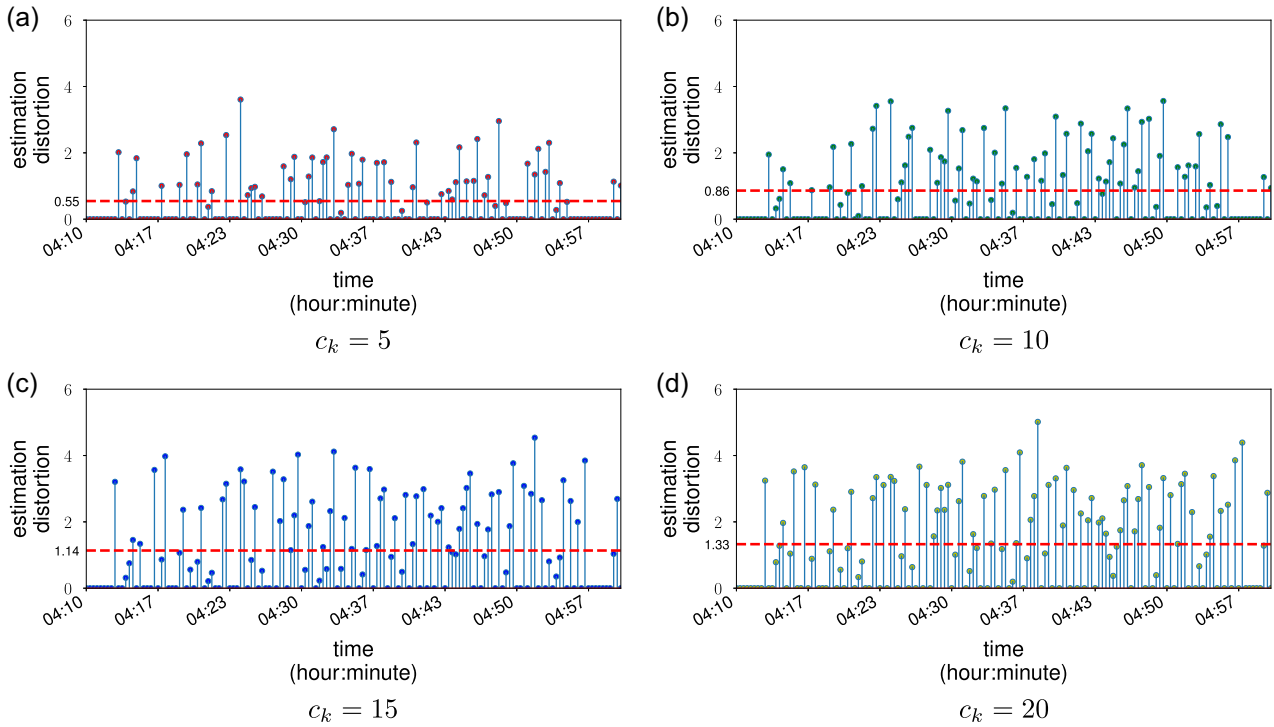


FIGURE 22 The estimation distortion (14) resulting from the optimal remote estimation scheme (4) under four different data transmission costs c_k , where the dotted red line depicts the average estimation distortion. The data transmission rate (15): (a) 88/150, (b) 72/150, (c) 66/150, and (d) 61/150 [Color figure can be viewed at wileyonlinelibrary.com]

parameter $\psi(k)$ computed using (2) and the GPS data. We set $\psi(k)$ to zero if no more than one animal is in motion. In comparison with Figure 18b, a larger portion of Figure 19b has $\psi(k)$ greater than 0.5 which implies that the trajectories depicted in Figure 19 are more coherent than those in Figure 18.

The acceleration data, depicted in Figures 16 and 17, show a clear contrast among three basic motions of waterbucks—*Walking*, *Eating*, and *Running*—where the labeling of the motions is performed using video data recorded along with the acceleration data. For this reason, this particular data set is chosen to evaluate the performance of the prediction (1). In addition, as shown in Figures 12 and 13, the core areas of the instrumented waterbucks are spatially distributed over a wide area, whereas the instrumented buffaloes tend to travel in a

group during the deployment. We select a subset of the GPS data for the waterbucks and buffaloes as shown, respectively, in Figures 18 and 19 to assess how the order parameter quantifies group movements of the instrumented animals.⁸

Next, we examine the performance of the remote estimation algorithm (described in Section 4.2) using the GPS data collected at 1 fix/20 s (alert state) from a device collared on one of the monitored buffaloes. The GPS data used to evaluate the algorithm are depicted

⁸As we have observed from the deployment data illustrated in Figures 13a and 15b, the instrumented waterbucks tend to be spatially distributed over a wide area and the resulting order parameter is expected to be low. To draw a clear contrast between the order parameters of the two species, we evaluate the order parameter of the two most closely located waterbucks—Waterbucks 6 and 10—and present it in Figure 18.

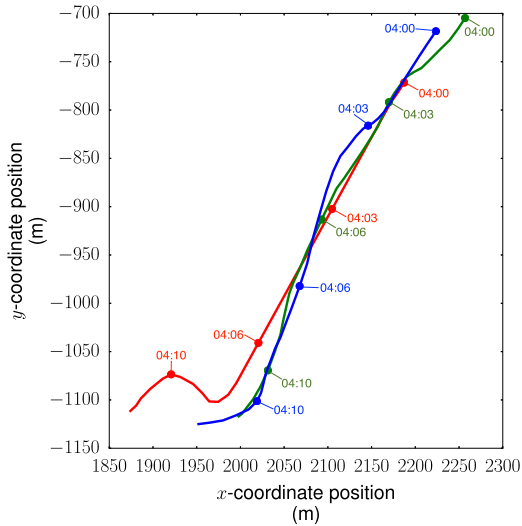


FIGURE 23 Global positioning system trajectory data collected during 8/6/2015, 04:00–04:12 from the devices collared on buffaloes 1, 2, 4 used to evaluate the distributed estimation algorithm [Color figure can be viewed at wileyonlinelibrary.com]

in Figure 20. To model the movement of the buffalo, we adopt the self-propelled particle model (6) for which the random processes $\nu(k)$ and $\phi(k)$ are chosen to be Weibull and Wrapped Cauchy, respectively, as in McClintock et al. (2012).

For simplicity, we assume that $\nu(k)$ and $\phi(k)$ are independent and identically distributed, and their respective probability density functions are given as follows:

$$f_{\nu(k)}(\nu) = \begin{cases} c_\nu \left(\frac{\nu - l_\nu}{s_\nu} \right)^{c_\nu - 1} \exp\left(-\left(\frac{\nu - l_\nu}{s_\nu}\right)^{c_\nu}\right) & \text{for } \nu \geq l_\nu, \\ 0 & \text{otherwise,} \end{cases} \quad (12a)$$

$$f_{\phi(k)}(\phi) = \frac{1}{2\pi} \frac{1 - c_\phi^2}{1 + c_\phi^2 - 2c_\phi \cos((\phi - l_\phi)/s_\phi)}. \quad (12b)$$

Using the GPS data, we compute the maximum likelihood estimates of the parameters of (12) as follows:

$$(c_\nu, l_\nu, s_\nu) = (1.0349, 0.8349, 2.8770), \quad (13a)$$

$$(c_\phi, l_\phi, s_\phi) = (0.6500, 0.0033, 0.9995). \quad (13b)$$

The plots in Figure 21 show the comparisons between the resulting probability density functions (12) and the histograms obtained from the GPS data of Figure 20.

By following a computational method described in (Park & Martins, 2016), we find the optimal remote estimation scheme (4) for the performance index (8) under several different choices of the data transmission cost c_k . Figure 22 depicts the resulting estimation distortion and the data transmission rate, respectively, defined by

$$[\|p(k) - \hat{p}(k)\|_2^2 + 4(1 - \cos(\theta(k) - \hat{\theta}(k)))]^{1/2}, \quad (14)$$

$$\text{data transmission rate} = \frac{\text{the total number of data transmissions}}{\text{the length of time horizon}}. \quad (15)$$

Notice that as c_k increases, the estimation distortion tends to increase while the total number of data transmissions decreases. A proper choice of c_k can be made using numerical simulations or evaluation of the deployment data set. For both cases, one may use Model (12) and simulated or deployment data set to evaluate (14) and (15) for multiple values of c_k , and select a suitable c_k .

Finally, we evaluate the performance of the distributed estimation algorithm (described in Section 4.3) in estimating the locations and velocities of three buffaloes that, according to Figure 23, follow what appear to be randomly perturbed straight trajectories⁹ that can be adequately modeled by (9). Using the GPS data collected from the devices mounted on these animals, we extract the location and velocity profiles, and then using (11), we compute the location and velocity estimates of all three buffaloes at every device. In the computation, we set the time interval between two consecutive discrete time steps as 20 s (which corresponds to the GPS rate in high state of the system), and we assume that the animal movements, described by (9), are all decoupled, that is, the velocity coupling constants $\{a_{ij}\}_{i,j=1}^M$ in (9) satisfy (10). Also, the neighborhood set $\mathbb{N}_i(k)$ in (11) is determined to be the set of devices that are within 100m from each device i (including device i itself); and we assign $K_i^p = K_i^v = 0.8 e_i \otimes I_2$ for the gain parameters in (11). Figure 24 depicts the resulting estimation distortion at all three devices under three different choices of the update rate N_c in (11) where the estimation distortion is defined as follows. Let $\hat{p}_i^{(j)}(k)$ and $\hat{v}_i^{(j)}(k)$ be the location and velocity estimates of the j th instrumented animal computed by device i , respectively. The distortion associated with the estimate $\hat{p}_i^{(j)}(k)$, $\hat{v}_i^{(j)}(k)$ is given by

$$[\|p^{(j)}(k) - \hat{p}_i^{(j)}(k)\|^2 + \|v^{(j)}(k) - \hat{v}_i^{(j)}(k)\|^2]^{1/2}. \quad (16)$$

Notice that as N_c increases, the resulting estimation distortion tends to decrease while the number of data transmissions required by (11) increases. Also, it is observed that the rise and fall of the estimation distortion over time consistently appears in all three figures in each row, which illustrates a feedback mechanism of the algorithm.

6.3 | Further observations and lessons learned for future deployments

Our discussion in Section 6.2 makes evident the potential that the data collected by our system during a deployment have to enable the study of the behavior of animals in their natural habitat. In this section, we propose additional uses for the measurements of certain sensors, such as when they can be employed to discern behaviors of interest. This type of characterization could be used to design

⁹We leave an extension and evaluation of the algorithm to other classes of animal group movements as future work.

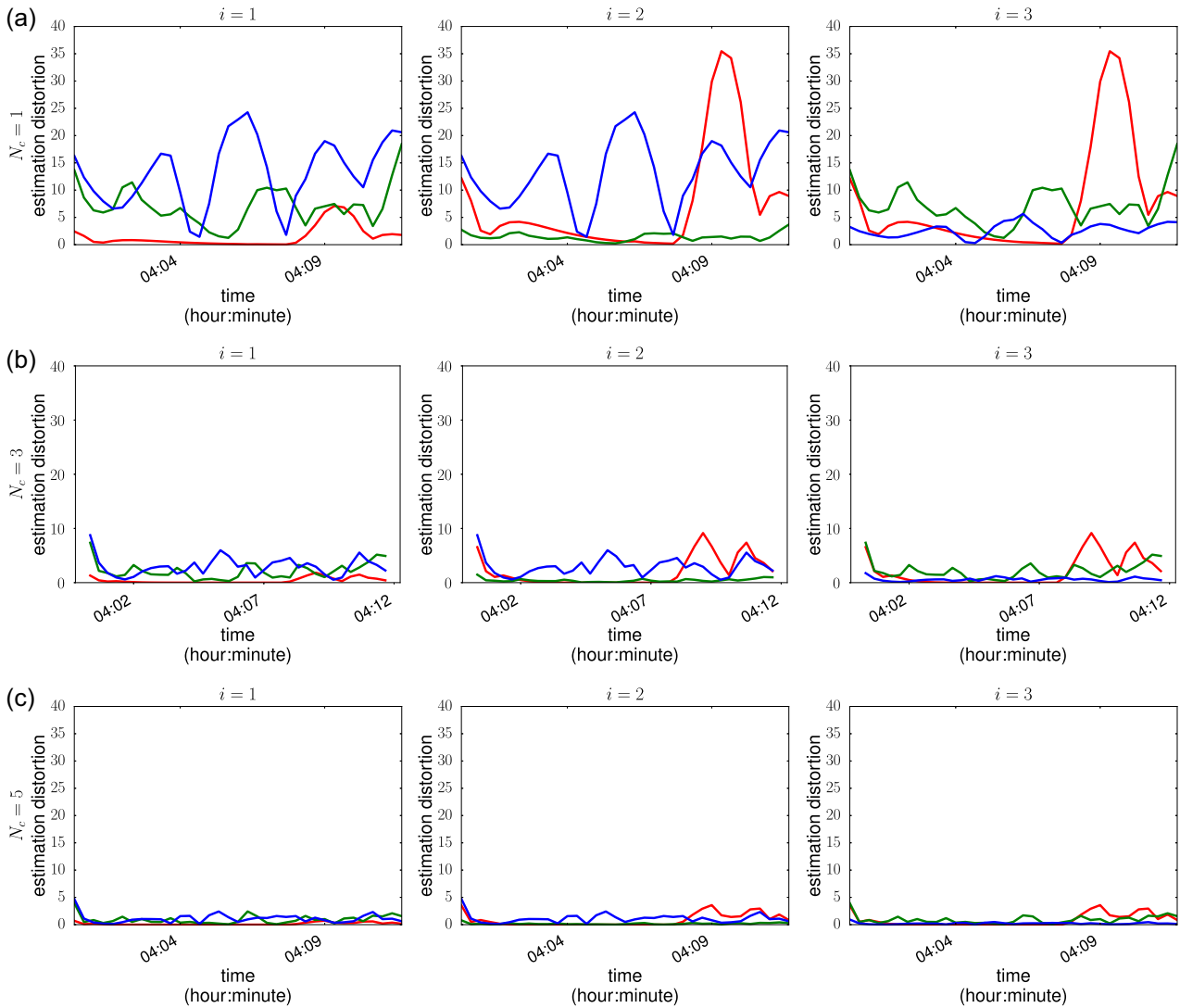


FIGURE 24 The estimation distortion (16) resulting from applying the distributed estimation algorithm to the GPS trajectory data depicted in Figure 23: The estimation distortion at each device $i = 1, 2, 3$ for (a) $N_c = 1$, (b) $N_c = 3$, and (c) $N_c = 5$. Each line in every graph is associated with a GPS trajectory in Figure 23 based on its color. GPS, global positioning system [Color figure can be viewed at wileyonlinelibrary.com]

triggering policies to capture video documenting these or other associated behaviors in future deployments.

To support our proposed methods, we also provide additional analyses based on the three sets of sensor data plotted in Figure 25, each corresponding to a video clip that is contemporaneously recorded by the same device and is illustrated in Figure 14 by a few selected frames. More specifically, we propose using sensor data to characterize the following behaviors:

- **Foraging behaviors:** Documenting what, when, and where each study animal eats is essential in any study of animal nutrition and foraging behavior. Hence, video recording such as the segment represented in Figure 14a is highly desirable because it clearly reveals what an instrumented animal eats, which otherwise would have to be determined by the analysis of fecal matter collected from the ground—a method that does not provide useful information such as which animal generated the sample or the

location and the precise time when the food was consumed. This suggests that the efficacy of future deployments seeking to specifically study foraging behavior could be significantly enhanced by employing onboard triggering mechanisms that activate video recording when an instrumented animal eats. Hence, the analysis represented in Figures 16 and 17 is specially relevant in this context because it verifies using deployment data that our motion predictor (1) can very effectively use onboard accelerometry measurements to discern *Walking*, *Eating*, and *Running*. This indicates that besides having the potential to be used as an onboard mechanism for triggering the recording of these behaviors, our motion predictor would also be very valuable for automating the postdeployment data processing needed to label GPS tracks documenting foraging over long periods of time. We also would like to emphasize that, in addition to the analysis in Figures 16 and 17, we validated the efficiency of our motion state predictor extensively using other data sets collected from our

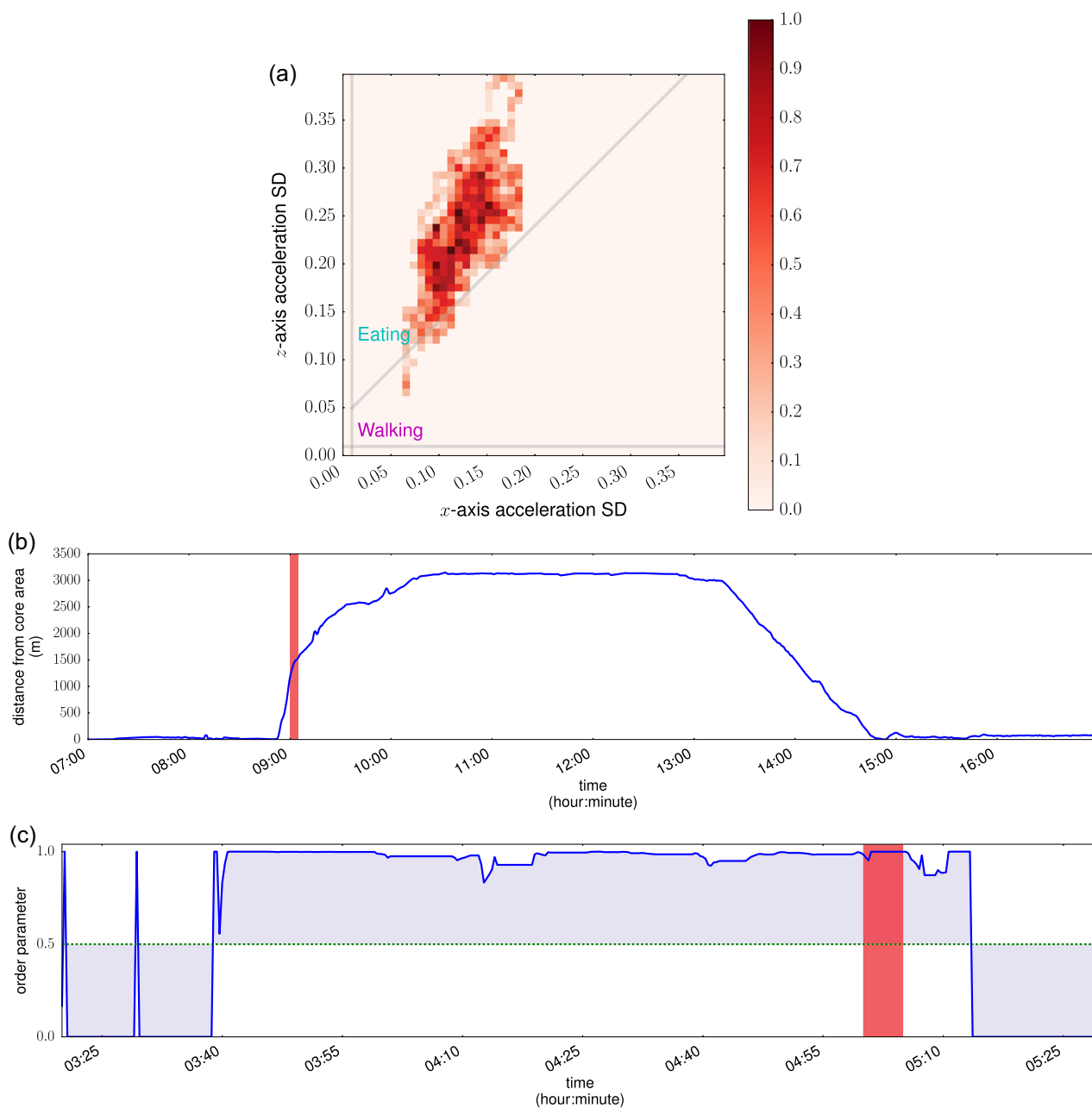


FIGURE 25 The data associated with the video frames presented in Figure 14: (a) shows the heat map of the SD of the accelerometry data logged contemporaneously with the video represented by the frames in Figure 14a; (b) shows the distance between the location of Waterbuck 2 and its core area in which the red band denotes the time period during which the video represented by the frames in Figure 14b was recorded; (c) shows the order parameter of Buffaloes 2, 3, 4 in which the red band denotes the time period during which the video represented by the frames in Figure 14c was recorded [Color figure can be viewed at wileyonlinelibrary.com]

deployment such as the one showing 98% accuracy in Figure 25a, which was logged contemporaneously with the video clip represented by the frames in Figure 14a.

- *Excursions from core areas:* Detecting excursions of land animals from their core areas is crucial to inform fauna conservation efforts, especially near large infrastructures or urban centers. We have observed from the deployment data documenting the deviation of waterbucks from their core areas, which can be obtained for each waterbuck using GPS tracks, that there were well-defined excursion events. An example illustrating this for Waterbuck 2 is shown in Figure

25b, which also indicates that simple distance thresholds could be quite effective in detecting an excursion event. Distance thresholds could also be used as onboard mechanisms to activate video recording when an excursion begins to determine whether there is a cause. This is illustrated by the video from which the frames in Figure 14b were extracted, which was recorded during the 5-min interval represented by the red band in Figure 25b.

- *Coherent group movements:* A comparison between Figures 18b and 19b showing the order parameter for two groups of study animals indicates that the movements of the former (waterbucks) are, on the

average, less coherent than the latter (buffaloes). This prompted us to further investigate the order parameter for groups of buffaloes for time periods during which we can validate based on video that they were indeed moving coherently. As is suggested by Figure 25c plotting the order parameter for Buffaloes 2, 3, and 4 in which the red band denotes the time interval during which the video represented by the frames of Figure 14c showing the buffaloes walking in line along a common trail was recorded, we expect the order parameter to be consistently near 1 when a group of buffaloes moves coherently. This indicates that the order parameter (2) may be useful both as an onboard mechanism for triggering the recording of coherent movement of buffalo herds and as a statistic for automating the postdeployment data processing needed to label GPS tracks according to movement coherence.

7 | SUMMARY, FUTURE PLANS, AND OTHER APPLICATIONS

This article details the conception and refinement of a system formed by identical networked battery-operated monitoring units meant to be mounted on mobile subjects to record their group motion, video, and other relevant data remotely. It also details our approach to codesigning algorithms, hardware, and firmware to enhance the value of the data collected while extending the time each unit can operate without having to recharge the batteries.

The article also describes a deployment mission at Mozambique's Gorongosa National Park to monitor two different species of free-ranging animals. Each animal in this study carried a collar-mounted monitoring unit capable of registering position, motion, and animal point-of-view video data, which collectively provide valuable information to community ecology studies. The data collected during the deployment were also used to evaluate the data acquisition performance of the system and to assess the performance of the algorithms. Furthermore, based on the analysis of the deployment data, the article describes the types of animal behaviors that our system can discern using accelerometry and location data. It also discusses triggering policies for video recording to document such behaviors.

Future plans: We are examining the deployment data to find a systematic way to determine parameters used in Models (5), (6), (9) and algorithms (1), (3), (11). For instance, we will identify a set of representative parameters of (6) for which we can precompute the remote estimation scheme associated with these parameters, and we plan to assess under what choice of the velocity couplings, Model (9) best describes group movement of the instrumented buffaloes. At the same time, we are working toward establishing a new class of parametrized animal behavioral models and modifying the associated algorithms to improve the data collection performance of our system. In an environment with interacting predator and prey species, dynamic models of pursuit and evasion in animal groups (such as Scott & Leonard, 2014) can be leveraged to refine predictions of the animals' motions, provided both species are collared with our devices.

Currently, we use heuristic mechanisms to trigger video recording that implement simple, and yet effective, ways to capture informative events. Based on the analysis given in Section 6.3 and improved collective animal behavioral models, we are developing new event-triggered schemes to further enhance the selectivity of video recording. For this purpose, we will employ our distributed estimation algorithm to promote omniscience of the key variables needed to implement triggers to discern group movements and configurations that characterize relevant events, such as foraging, predation, and collective defensive strategies to ward off attacks. We are also planning to carry out larger-scale deployments, which will allow further evaluation and refinement of our system.

Other applications: Our system could be customized for other applications, including disaster relief and firefighting. In these contexts, each monitoring unit, which would be carried by a responder, could support the following functionalities:

- Omniscience of the responders' positions and movements could be used to set up automatic alarms and provide real-time information to expert systems, which could issue recommendations to help coordinate the mobility of an overall team to mitigate risks and improve efficiency.
- The system could also provide valuable data on each mission to support analysis and future planning.

8 | ACKNOWLEDGMENTS

We are grateful to Robert Pringle (Princeton University) for his support on the deployment in Gorongosa National Park and to Jen Guyton (Princeton University) and Ryan Long (University of Idaho) for their assistance with fieldwork. The work of S.P., M.A., and N.C.M. was supported by NSF grant ECCS 1135726 and AFOSR grant FA95501510367. The work of W.L.S. and N.E.L. was supported by NSF grant ECCS 1135724. The work of K.H.A., K.A., G.M., and M.S. was supported by NSF grant ECCS 1135719 and the National Geographic Society.

ORCID

Shinkyu Park  <http://orcid.org/0000-0002-8643-404X>

William L. Scott  <http://orcid.org/0000-0002-0035-0671>

Naomi E. Leonard  <http://orcid.org/0000-0002-5328-3871>

Nuno C. Martins  <http://orcid.org/0000-0003-2083-8102>

REFERENCES

- Alpaydin, E. (2010). *Introduction to machine learning* (2nd ed.). Cambridge, MA, USA: The MIT Press.
- Altman, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, 49(3/4), 227–267.
- Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., & Zdravkovic, V. (2008). Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Sciences*, 105(4), 1232–1237.

- Cagnacci, F., Boitani, L., Powell, R. A., & Boyce, M. S. (2010). Animal ecology meets GPS-based radiotelemetry: A perfect storm of opportunities and challenges. *Philosophical Transactions of the Royal Society of London B*, 365(1550), 2157–2162.
- Cao, M., Morse, A. S., & Anderson, B. D. O. (2008). Reaching a consensus in a dynamically changing environment: A graphical approach. *SIAM Journal on Control and Optimization*, 47(2), 575–600.
- Cavagna, A., Cimarelli, A., Giardina, I., Parisi, G., Santagati, R., Stefanini, F., & Viale, M. (2010). Scale-free correlations in starling flocks. *Proceedings of the National Academy of Sciences*, 107(26), 11865–11870.
- Codling, E. A., Plank, M. J., & Benhamou, S. (2008). Random walk models in biology. *Journal of the Royal Society Interface*, 5(25), 813–834.
- Cooke, S. J., Hinch, S. G., Wikelski, M., Andrews, R. D., Kuchel, L. J., Wolcott, T. G., & Butler, P. J. (2004). Biotelemetry: A mechanistic approach to ecology. *Trends in Ecology and Evolution*, 19(6), 334–343.
- Couzin, I. D., & Krause, J. (2003). Self-organization and collective behavior in vertebrates. *Advances in the Study of Behavior*, 32, 1–75.
- Davis, R. W., Fuiman, L. A., Williams, T. M., Collier, S. O., Hagey, W. P., Kanatous, S. B., & Horning, M. (1999). Hunting behavior of a marine mammal beneath the antarctic fast ice. *Science*, 283(5404), 993–996.
- Dyo, V., Ellwood, S. A., Macdonald, D. W., Markham, A., Mascolo, C., Pásztor, B., & Yousef, K. (2010). Evolution and sustainability of a wildlife monitoring sensor network. *Proceedings of the 8th International Conference on Embedded Networked Sensor Systems*, 127–140, Zürich, Switzerland.
- Fleming, C. H., Calabrese, J. M., Mueller, T., Olson, K. A., Leimgruber, P., & Fagan, W. F. (2014). From fine-scale foraging to home ranges: A semivariance approach to identifying movement modes across spatiotemporal scales. *The American Naturalist*, 183(5), E154–E167.
- Gurarie, E., Andrews, R. D., & Laidre, K. L. (2009). A novel method for identifying behavioural changes in animal movement data. *Ecology Letters*, 12(5), 395–408.
- Halsey, L., Green, J., Wilson, R., & Frappell, P. (2009). Accelerometry to estimate energy expenditure during activity: Best practice with data loggers. *Physiological and Biochemical Zoology*, 82(4), 396–404.
- Handcock, R. N., Swain, D. L., Bishop-Hurley, G. J., Patison, K. P., Wark, T., Valencia, P., & O'Neill, C. J. (2009). Monitoring animal behaviour and environmental interactions using wireless sensor networks, GPS collars and satellite remote sensing. *Sensors*, 9(5), 3586.
- Heithaus, M. R., Marshall, G. J., Buhleier, B. M., & Dill, L. M. (2001). Employing Crittercam to study habitat use and behavior of large sharks. *Marine Ecology Progress Series*, 209, 307–310.
- Juang, P., Oki, H., Wang, Y., Martonosi, M., Peh, L. S., & Rubenstein, D. (2002). Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiences with ZebraNet. *Proceedings of the 10th International Conference on Architectural Support for Programming Languages and Operating Systems*, 96–107, San Jose, CA, USA.
- Holland, K. N., Meyer, C. G., & Dagorn, L. C. (2009). Inter-animal telemetry: Results from first deployment of acoustic 'business card' tags. *Endangered Species Research*, 10, 287–293.
- Krause, J., Krause, S., Arlinghaus, R., Psorakis, I., Roberts, S., & Rutz, C. (2013). Reality mining of animal social systems. *Trends in Ecology and Evolution*, 28(9), 541–551.
- Krause, J., Wilson, A. D., & Croft, D. P. (2011). New technology facilitates the study of social networks. *Trends in Ecology and Evolution*, 26, 5–6.
- Li, X. R., & Jilkov, V. P. (2003). Survey of maneuvering target tracking. Part I. dynamic models. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4), 1333–1364.
- Lipsa, G. M., & Martins, N. C. (2011). Remote state estimation with communication costs for first-order LTI systems. *IEEE Transactions on Automatic Control*, 56(9), 2013–2025.
- Liu, T., Sadler, C. M., Zhang, P., & Martonosi, M. (2004). Implementing software on resource-constrained mobile sensors: Experiences with Impala and ZebraNet. *Proceedings of the 2nd International Conference on Mobile Systems, Applications, and Services*, 256–269, Boston, MA, USA.
- Markham, A. C., & Wilkinson, A. J. (2008). *EcoLocate: A heterogeneous wireless network system for wildlife tracking* (pp. 293–298). Dordrecht: Springer Netherlands.
- Marshall, G. J. (1998). Crittercam: An animal-borne imaging and data logging system. *Marine Technology Society Journal*, 32(1), 11–17.
- McClintock, B. T., King, R., Thomas, L., Matthiopoulos, J., McConnell, B. J., & Morales, J. M. (2012). A general discrete-time modeling framework for animal movement using multistate random walks. *Ecological Monographs*, 82(3), 335–349.
- Mennill, D. J., Doucet, S. M., Ward, K.-A. A., Maynard, D. F., Otis, B., & Burt, J. M. (2012). A novel digital telemetry system for tracking wild animals: A field test for studying mate choice in a lekking tropical bird. *Methods in Ecology and Evolution*, 3(4), 663–672.
- Moll, R. J., Millspaugh, J. J., Beringer, J., Sartwell, J., & He, Z. (2007). A new 'view' of ecology and conservation through animal-borne video systems. *Trends in Ecology and Evolution*, 22(12), 660–668.
- 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.
- Nifong, J. C., Nifong, R. L., Silliman, B. R., Lowers, R. H., Guillette, L. J., Jr, Ferguson, J. M., & Marshall, G. (2014). Animal-borne imaging reveals novel insights into the foraging behaviors and diel activity of a large-bodied apex predator, the American alligator. *PLOS One*, 9(1), 1–11.
- Park, S., & Martins, N. C. (2014). Individually optimal solutions to a remote state estimation problem with communication costs. *IEEE 53th Conference on Decision and Control*, 4014–4019, Los Angeles, CA, USA.
- Park, S., & Martins, N. C. (2016). Optimal remote state estimation for self-propelled particle models. *IEEE 55th Conference on Decision and Control*, 327–333, Las Vegas, NV, USA.
- Park, S., & Martins, N. C. (2017). Design of distributed LTI observers for state omniscience. *IEEE Transactions on Automatic Control*, 62(2), 561–576.
- Ponganis, P., Van Dam, R., Marshall, G., Knowler, T., & Levenson, D. (2000). Sub-ice foraging behavior of emperor penguins. *Journal of Experimental Biology*, 203(21), 3275–3278.
- 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.
- Rutz, C., Bluff, L. A., Weir, A. A. S., & Kacelnik, A. (2007). Video cameras on wild birds. *Science*, 318(5851), 765–765.
- Rutz, C., & Hays, G. C. (2009). New frontiers in biologging science. *Biology Letters*, 5, 289–292.
- Scott, W. L., & Leonard, N. E. (2014). Dynamics of pursuit and evasion in a heterogeneous herd. *53rd IEEE Conference on Decision and Control*, 2920–2925, Los Angeles, CA, USA.
- Shepard, E. L. C., Wilson, R. P., Halsey, L. G., Quintana, F., Laich, A. G., Gleiss, A. C., & Norman, B. (2008). Derivation of body motion via appropriate smoothing of acceleration data. *Aquatic Biology*, 4(4), 235–241.
- Shepard, E. L. C., Wilson, R. P., Quintana, F., Laich, A. G., Liebsch, N., Albareda, D. A., & Macdonald, D. W. (2008). Identification of animal movement patterns using tri-axial accelerometry. *Endangered Species Research*, 10, 47–60.
- Sumpter, D. (2006). The principles of collective animal behaviour. *Philosophical Transactions of the Royal Society of London B*, 361(1465), 5–22.
- Takahashi, A., Sato, K., Naito, Y., Dunn, M., Trathan, P., & Croxall, J. (2004). Penguin-mounted cameras glimpse underwater group behaviour. *Proceedings of the Royal Society B: Biological Sciences*, 271(Suppl 5), S281–S282.
- Tomkiewicz, S. M., Fuller, M. R., Kie, J. G., & Bates, K. K. (2010). Global positioning system and associated technologies in animal behaviour

and ecological research. *Philosophical Transactions of the Royal Society B*, 365(1550), 2163–2176.

- Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I., & Shochet, O. (1995). Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75, 1226–1229.
- Vicsek, T., & Zafeiris, A. (2012). Collective motion. *Physics Reports*, 517 (3-4), 71–140.
- Wang, Y., Nickel, B., Rutishauser, M., Bryce, C. M., Williams, T. M., Elkaim, G., & Wilmers, C. C. (2015). Movement, resting, and attack behaviors of wild pumas are revealed by tri-axial accelerometer measurements. *Movement Ecology*, 3(1), 1–12.
- Williams, T. M., Davis, R. W., Fuiman, L. A., Francis, J., Le, B. J. B., Horning, M., & Croll, D. A. (2000). Sink or swim: Strategies for cost-efficient diving by marine mammals. *Science*, 288(5463), 133–136.
- Zhang, P., Sadler, C. M., Lyon, S. A., & Martonosí, M. (2004). Hardware design experiences in ZebraNet. *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems, SenSys 2004*, ACM, 227–238, Baltimore, MD, USA.

SUPPORTING INFORMATION

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

How to cite this article: Park S, Aschenbach KH, Ahmed M, et al. Animal-borne wireless network: Remote imaging of community ecology. *J Field Robotics*. 2019;1–25. <https://doi.org/10.1002/rob.21891>

APPENDIX A: INDEX TO MULTIMEDIA EXTENSIONS

The video is available as Supporting Information in the online version of this paper.

Extension	Media type	Description
1	Video	System operation and deployment in Gorongosa National Park

APPENDIX B: CONVERGENCE PROPERTIES OF THE DISTRIBUTED ESTIMATION ALGORITHM

To establish convergence results, we adopt graph-theoretic notions from (Cao, Morse, & Anderson, 2008).

Definition 1. Let $\mathcal{G}_1 = (\mathbb{V}, \mathbb{E}_1)$ and $\mathcal{G}_2 = (\mathbb{V}, \mathbb{E}_2)$ be directed graphs. A graph $\overline{\mathcal{G}} = (\mathbb{V}, \mathbb{E})$ is said to be a *composition* of \mathcal{G}_1 and \mathcal{G}_2 if (i, j) belongs to \mathbb{E} whenever there is a vertex $l \in \mathbb{V}$ for which $(i, l) \in \mathbb{E}_1$

and $(l, j) \in \mathbb{E}_2$ hold. For notational convenience, we let $\overline{\mathcal{G}} = \mathcal{G}_2 \circ \mathcal{G}_1$.

Definition 2. Let $\mathcal{G}(k) = (\mathbb{V}, \mathbb{E}(k))$ be a directed graph whose edge set $\mathbb{E}(k)$ depends on the time index $k \in \mathbb{N}$. We say that $\mathcal{G}(k)$ is *jointly strongly connected* over a sequence of time indices $\{k_1, \dots, k_m\}$ if the composition $\overline{\mathcal{G}} = \mathcal{G}(k_m) \circ \dots \circ \mathcal{G}(k_1)$ of $\mathcal{G}(k_1), \dots, \mathcal{G}(k_m)$ is strongly connected.

Remark 1. Suppose that there exists a positive integer T such that for every n in \mathbb{N} , a graph $\mathcal{G}(k)$ is jointly strongly connected over $\{nT, \dots, (n+1)T - 1\}$. According to Proposition 4 in (Cao et al., 2008), the composition $\overline{\mathcal{G}}(n) = \mathcal{G}((n+1)T - 1) \circ \dots \circ \mathcal{G}(nT)$ is a complete directed graph for each n in \mathbb{N} .

Below, we provide a theorem that establishes uniform exponential convergence on the distributed estimation (11). We consider the case where there are no noise terms in (9), that is, $w_p^{(i)}(k) = w_v^{(i)}(k) = 0$ for all i in $\{1, \dots, M\}$ and k in \mathbb{N} ; however we note that, as the convergence is uniform exponential, the statement of the theorem implies bounded estimation distortion when all the noise terms in (9) are bounded.

Theorem 1. Consider the distributed estimation (11) in which its underlying graph $\mathcal{G}(k) = (\mathbb{V}, \mathbb{E}(k))$ is defined by a vertex set $\mathbb{V} = \{1, \dots, M\}$ and a time-varying edge set $\mathbb{E}(k) \subseteq \mathbb{V} \times \mathbb{V}$ for which $(j, i) \in \mathbb{E}(k)$ if and only if $j \in \mathbb{N}_i(k)$. Suppose that there exists a positive integer T such that for every n in \mathbb{N} , the graph $\mathcal{G}(k)$ is jointly strongly connected over $\{nT, \dots, (n+1)T - 1\}$. There are parameter choices for K_i^p, K_i^v for which the estimate $\hat{p}_i(k) \hat{v}_i(k)$ computed by (11) uniformly exponentially converges to $p(k) v(k)$ for all i in $\{1, \dots, M\}$.

Proof. For simplicity, we provide a proof for the case where $N_c = 1$ and proceed with a choice of the parameters $K_i^p = K_i^v = 0.8 e_i \otimes I_2$. Notice that if $N_c = 1$, then (11) can be concisely rewritten as follows:

$$\hat{p}_i^-(k) = \hat{p}_i(k-1) + \hat{v}_i(k-1), \quad (\text{B1a})$$

$$\hat{v}_i^-(k) = A \hat{v}_i(k-1), \quad (\text{B1b})$$

$$\hat{p}_i(k) = \sum_{j \in \mathbb{N}_i(k)} \frac{1}{|\mathbb{N}_i(k)|} (\hat{p}_j^-(k) + K_i^p (p^{(i)}(k) - C_i \hat{p}_j^-(k))), \quad (\text{B2a})$$

$$\hat{v}_i(k) = \sum_{j \in \mathbb{N}_i(k)} \frac{1}{|\mathbb{N}_i(k)|} (\hat{v}_j^-(k) + K_i^v (v^{(i)}(k) - C_i \hat{v}_j^-(k))). \quad (\text{B2b})$$

To prove the statement, let us first investigate the convergence of the velocity estimate $\hat{v}_i(k)$. By defining $\tilde{v}_i(k) = v(k) - \hat{v}_i(k)$, we obtain

$$\tilde{v}_i(k) = (I_{2M} - 0.8 e_i e_i^T \otimes I_2) A \sum_{j \in \mathbb{N}_i(k)} \frac{1}{|\mathbb{N}_i(k)|} \tilde{v}_j(k-1). \quad (\text{B3})$$

For notational convenience, we define

$$\tilde{v}'_i(k) = \begin{cases} \sum_{j \in \mathbb{N}_i(k)} \frac{1}{|\mathbb{N}_i(k)|} \tilde{v}_j(k-1), & k \geq 1, \\ \tilde{v}_i(0), & k = 0, \end{cases} \quad (\text{B4})$$

and $\tilde{v}'(k) = (\tilde{v}'_1(k) \dots \tilde{v}'_M(k))^T$. Note that $\tilde{v}_i(k)$ can be expressed as $\tilde{v}_i(k) = (I_{2M} - 0.8 e_i e_i^T \otimes I_2) A \tilde{v}'_i(k)$ for $k \geq 1$. According to (B3), the estimation error $\tilde{v}'(k)$ evolves according to the following equation:

$$\tilde{v}'(k) = \Phi^v(k) \tilde{v}'(k-1), \quad (\text{B5})$$

where $\Phi^v(k)$ is a $2M^2 \times 2M^2$ matrix that is obtained from (B3) for which the relation (B5) holds. Note that $\Phi^v(k)$ is an element-wise nonnegative matrix and the sum of all elements in each row is equal to or less than 1.

Suppose that there is a constant $0 \leq \lambda < 1$ for which

$$\|\Phi^v((n+1)T-1) \dots \Phi^v(nT)\|_\infty \leq \lambda \quad (\text{B6})$$

holds for all positive integer n , where $\|\cdot\|_\infty$ is the (induced) infinity matrix norm. Using the fact that $\Phi^v(k)$ is uniformly bounded for all k in \mathbb{N} , in conjunction with (B3) and (B5), we conclude that $\|\tilde{v}'(k)\|_\infty \leq \alpha \lambda^{k/T} \|\tilde{v}'(0)\|_\infty$ holds for some $\alpha > 0$. This establishes the uniform exponential convergence of $\tilde{v}_i(k)$ to $v(k)$ for all i in $\{1, \dots, M\}$.

According to Remark 1, without loss of generality, we assume that the composition $\mathcal{T}(n) = \mathcal{G}((n+1)T-1) \circ \dots \circ \mathcal{G}(nT)$ is a complete directed graph for each n in \mathbb{N} . Under this assumption, in what follows, we prove that there is a constant $0 \leq \lambda < 1$ for which (B6) holds for all n in \mathbb{N} . We note that the matrix product $\Phi^v((n+1)T-1) \dots \Phi^v(nT)$ is an element-wise nonnegative matrix and the sum of all elements in each row is equal to or less than 1. Hence, the following inequality holds element-wise:

$$\Phi^v((n+1)T-1) \dots \Phi^v(nT) \mathbf{1} \leq \mathbf{1}, \quad (\text{B7})$$

where $\mathbf{1}$ is a $2M^2$ -dimensional vector with all its elements equal to 1. Since there is an edge from vertex j to vertex i for every i, j in the composition graph $\mathcal{T}(n)$, according to (B3) and (B4), the inequality in (B7) holds strictly element-wise. Also for fixed A , since there are only a finite number of realizations of $\Phi^v(k)$, we conclude that $\|\Phi^v((n+1)T-1) \dots \Phi^v(nT)\|_\infty \leq \lambda$ for some $0 \leq \lambda < 1$.

From the above relation, for any x in \mathbb{R}^{2M^2} , we can infer that

$$\begin{aligned} & \|\Phi^v((n+1)T-1) \dots \Phi^v(nT) x\|_\infty \\ &= \max_{i \in \{1, \dots, 2M^2\}} \left| \sum_{j=1}^{2M^2} [\Phi^v((n+1)T-1) \dots \Phi^v(nT)]_{ij} x_j \right| \\ &\leq \max_{i \in \{1, \dots, 2M^2\}} \sum_{j=1}^{2M^2} [\Phi^v((n+1)T-1) \dots \Phi^v(nT)]_{ij} |x_j| \\ &\leq \max_{i \in \{1, \dots, 2M^2\}} \sum_{j=1}^{2M^2} [\Phi^v((n+1)T-1) \dots \Phi^v(nT)]_{ij} \max_{j \in \{1, \dots, 2M^2\}} |x_j| \\ &< \lambda \|x\|_\infty. \end{aligned}$$

Since the choice of x is arbitrary, we conclude that $\|\Phi^v((n+1)T-1) \dots \Phi^v(nT)\|_\infty < \lambda$.

Now we investigate the convergence of the position estimate $\hat{p}_i(k)$. By defining $\tilde{p}_i(k) = p(k) - \hat{p}_i(k)$, we obtain

$$\tilde{p}_i(k) = (I_{2M} - 0.8 e_i e_i^T \otimes I_2) \sum_{j \in \mathbb{N}_i(k)} \frac{1}{|\mathbb{N}_i(k)|} [\tilde{p}_j(k-1) + \tilde{v}_j(k-1)]. \quad (\text{B8})$$

By similar arguments used in proving the convergence of $\tilde{v}(k)$, in conjunction with the exponential convergence of $\tilde{v}(k)$ to the origin, we can conclude that $\|\tilde{p}(k)\|_\infty \leq \beta_p \mu^{k/T} \|\tilde{p}(0)\|_\infty + \beta_v \lambda^{k/T} \|\tilde{v}(0)\|_\infty$ holds for some $\mu, \lambda, \beta_p, \beta_v > 0$, where $\tilde{p}(k) = (\tilde{p}_1(k) \dots \tilde{p}_M(k))^T$. In conjunction with the uniform exponential convergence of $\tilde{v}_i(k)$ for all i in $\{1, \dots, M\}$, we conclude that the estimate $\hat{p}_i(k)$ uniformly exponentially converges to $p(k)$ for all i in $\{1, \dots, M\}$. \square

APPENDIX C: MARGIN-MAXIMIZING SELECTION OF $\tau_x, \tau_z, \tau'_x, \tau'_z, \tau_{xz}$

By the construction of the motion prediction rule (1), each of the parameters $\tau_x, \tau_z, \tau'_x, \tau'_z, \tau_{xz}$ is used to distinguish exactly two animal motions for which we can apply supervised learning to compute each parameter. As a case in point, given $\tau_x, \tau_z, \tau'_x, \tau'_z$, the motion prediction rule (1) is simplified as follows:

$$\hat{m}(\sigma) = \begin{cases} \text{Walking} & \sigma_x + \tau_{xz} \geq \sigma_z, \\ \text{Eating} & \sigma_x + \tau_{xz} < \sigma_z. \end{cases} \quad (\text{C1})$$

Thus, given training data set $\{(\sigma^{(i)}, m^{(i)})\}_{i=1}^K$ satisfying $\tau_x \leq \sigma_x^{(i)} < \tau'_x, \tau_z \leq \sigma_z^{(i)} < \tau'_z$, and $m^{(i)} \in \{\text{Walking}, \text{Eating}\}$, the parameter τ_{xz} can be computed by solving the following optimization that maximizes the margin for the motion prediction error (Alpaydin, 2010, Chapter 10.9):

$$\begin{aligned} & \min_{a, b, \xi^{(i)}} a^2 + \lambda \sum_{i=1}^K \xi^{(i)}, \\ & \text{subject to } m^{(i)}(a(\sigma_z^{(i)} - \sigma_x^{(i)}) + b) \geq 1 - \xi^{(i)} \quad \forall i \in \{1, \dots, K\}, \\ & \xi^{(i)} \geq 0 \quad \forall i \in \{1, \dots, K\}, \end{aligned} \quad (\text{C2})$$

where for notational convenience we reassign $\text{Walking} = -1$ and $\text{Eating} = +1$. The variable a represents (the reciprocal of) the margin of the prediction rule (C1), the variable $\xi^{(i)}$ denotes a slack variable specifying the deviation of the training data from the margin, and λ is a constant that determines the trade-off between the margin (quantified as a^2) and the deviation from the margin (quantified as $\sum_{i=1}^K \xi^{(i)}$). We select $\lambda = 1$ and the resulting solution a, b is used to define $\tau_{xz} = -b/a$. As the rest of the parameters $\tau_x, \tau_z, \tau'_x, \tau'_z$ can be similarly computed, we omit details for brevity.