

Sensor and the Beast: Generalised Methods to Recognise Animal Behaviour and Quantify Energy Expenditure Using Inertial Sensors, and Applications

Présentée le 6 août 2020

à la Faculté des sciences et techniques de l'ingénieur
Laboratoire de mesure et d'analyse des mouvements
Programme doctoral en biotechnologie et génie biologique

pour l'obtention du grade de Docteur ès Sciences

par

Priritish CHAKRAVARTY

Acceptée sur proposition du jury

Prof. D. Pioletti, président du jury
Prof. K. Aminian, directeur de thèse
Prof. E. Shepard, rapporteuse
Dr Y. Ropert-Coudert, rapporteur
Prof. S. Sakar, rapporteur

Abstract

How an animal allocates time and energy to different activities has repercussions for its survival, fitness, and response to changing environmental conditions. Fine-scale information on animal behaviour and energy expenditure can help develop better informed, targeted strategies to conserve and manage wild animal populations in the face of increasing anthropogenic impact, and climate change. However, wild animals are notoriously difficult to observe, either because they avoid humans entirely or live in habitats that make following them impossible. Miniature data-logging devices attached to animals can address these difficulties by making it possible to remotely observe animals through the ‘eyes’ of recorded data. Inertial sensors such as accelerometers that are sensitive to movement and the lack of it have offered a powerful tool to remotely quantify both animal behaviour and energetics over the last two decades. However, even though many studies have been able to infer common, coarse-scale animal behaviours such as resting, locomotion, and feeding/foraging from accelerometer data, there is a lack of general methods that can be applied across species. Further, few studies have been able to resolve fine-scale behaviours accurately, possibly because of the complexity of nonstationary signals recorded during such behaviours. Finally, techniques currently used to quantify energy expenditure from accelerometer data rely on assumptions that may not always be valid. Magnetometers, another kind of inertial sensor often included in animal tags, have shown the potential to capture both static and dynamic components of movement, as well as the ability to detect some behaviours that accelerometers miss. Yet, no study has explored the potential of magnetometers to recognise common animal behaviours, and perform a one-to-one comparison between accelerometers and magnetometers to assess the relative strengths and weaknesses of these two sensors.

The main objectives of this thesis were to address these issues by developing general, robust methods to quantify behaviour and energy expenditure (EE) from inertial sensor data. The thesis first presents a biomechanically driven learning

approach for the recognition of common animal behaviours, separating them on the basis of posture, movement intensity, and movement periodicity. Using data collected on 10 wild meerkats (*Suricata suricatta*), I show the high accuracy and robustness of this model to recognise common meerkat behaviours (resting, vigilance, foraging, and running) not only with accelerometer data, but also, in a separate study, with magnetometer data. Complementary benefits of the two sensors were found as a result of this: the magnetometer had higher robustness to inter-individual variability in dynamic behaviours, while the accelerometer was better at estimating posture. Next, the thesis presents a new approach to resolve fine-scale behaviours. Here, complex behaviour is conceptualised as being composed of characteristic impulsive movements (microevents) producing brief shock signals in accelerometer data. Requisite signal processing techniques are developed to first detect ('seek') microevents in acceleration signals, and then robust machine learning ('learn') is used to separate different types of microevents. This 'seek and learn' approach effectively helps cross the hurdle of nonstationarity that had previously hindered progress in this direction. Using data collected on fine-scale foraging modes in meerkats – searching, one-armed digging, two-armed digging, and prey chewing, I show that this approach has high overall accuracy, and outperforms approaches based on more classical machine learning approaches.

Inspired from approaches adopted in human studies to estimate EE, this thesis proposes a new method for estimating EE based on metabolic equivalent of task (MET), where rate of EE for each activity is given in terms of multiples of resting metabolic rate. The MET-based method only requires durations of behaviour – obtained, for instance, from an accelerometer-based behaviour recognition model, and leverages measurements and empirical relationships from past studies on the same or multiple surrogate species. Going through behaviour duration rather than the raw acceleration signal helps it avoid assumptions about the precision of gravity compensation to estimate dynamic body acceleration. MET-based EE estimates obtained in kilo-Joules showed good agreement with previous measurements of EE in meerkats using the gold-standard doubly labelled water technique. Further, MET-based EE correlated strongly with vectorial dynamic body acceleration, indicating its potential for applicability on other species as well.

Finally, I come full circle by applying these ‘engineering’ developments to the ‘ecological’ question of how differences in individual traits, such as body mass, affect foraging strategy in free-ranging, wild, group-living meerkats. Meerkats transition very quickly from one behaviour to another, and investigating this question through direct observation of multiple individuals is arduous. Using the coarse- and fine-scale behaviour recognition model and EE-estimation technique developed in this thesis, I found that differences in body mass in females were strongly linked to differences in foraging strategy: lighter females searched almost exclusively for low-value surface prey while heavier females preferred to dig for higher-value below-ground prey. Inferences drawn from this foraging strategy and known prey distributions in the Kalahari suggest that lighter females may be ‘settling’ for low-value surface prey in order to avoid losing food to heavier females in foraging conflicts they are likely to lose.

The algorithms developed in this thesis highlight the fact that, with appropriate data processing techniques, low-power sensors like accelerometers can provide realistic estimates of energy expenditure in the standard units of kilo-Joules, and pick up fine differences in behavioural patterns that would be difficult to ascertain even when direct observation of wild animals is possible.

Keywords: accelerometer, magnetometer, behaviour recognition, meerkat, energy expenditure, machine learning, animal ecology, cross-validation, foraging.

Résumé

La façon dont un animal alloue du temps et de l'énergie à différentes activités a des répercussions pour sa survie, sa valeur adaptative darwinienne, et sa réponse aux conditions environnementales changeantes. L'information détaillée sur le comportement et dépense énergétique animale peut aider à développer des stratégies mieux renseignées et mieux ciblées pour conserver et gérer les populations animales sauvages en face de l'impact anthropique croissant et le changement climatique. Cependant, les animaux sauvages sont notoirement difficiles à observer, soit parce qu'ils évitent complètement les humains ou vivent dans des habitats qui les rendent impossibles à suivre. Les dispositifs d'enregistrement de données (« data-loggers ») miniatures attachés aux animaux peuvent résoudre ces difficultés en rendant possible l'observation à distance des animaux à travers les « yeux » des données enregistrées. Les capteurs inertiels tels que les accéléromètres qui sont sensibles au mouvement et son absence ont offert un outil puissant pour quantifier à distance le comportement animal ainsi que l'énergie animale au cours des deux dernières décennies. Néanmoins, même si de nombreuses études ont été en mesure de déduire des comportements animaux communs grossiers tel que le repos, la locomotion, et l'acte de se nourrir/recherche de nourriture, il y a un manque de méthodes générales pouvant être appliquées de manière plus universelle. En outre, peu d'études ont été en mesure de reconnaître avec précision des comportements détaillés, peut-être en raison de la complexité des signaux non stationnaires enregistrés lors de ces comportements. Enfin, les techniques actuellement utilisées pour quantifier la dépense énergétique à partir des données de l'accéléromètre reposent sur des hypothèses qui ne sont pas toujours valables. Le magnétomètre, un autre capteur inertiel souvent inclus dans des data-loggers, a démontré son potentiel pour capturer les composantes statiques et dynamiques du mouvement, ainsi que la capacité à identifier des comportements que l'accéléromètre ne détecte pas. Pourtant, aucune étude n'a exploré le potentiel du magnétomètre à reconnaître des comportements animaux communs, et effectuer une comparaison entre l'accéléromètre et le magnétomètre pour évaluer les forces et les faiblesses relatives de ces deux capteurs.

Les principaux objectifs de cette thèse étaient de résoudre ces problèmes en développant des méthodes générales et robustes pour quantifier le comportement et

la dépense énergétique (DE) à partir des données des capteurs inertiels. La thèse présente d'abord une approche d'apprentissage basée sur des considérations biomécaniques pour la reconnaissance des comportements communs des animaux, en les séparant sur la base de la posture, de l'intensité du mouvement et de la périodicité du mouvement. En utilisant des données récoltées sur 10 suricates sauvages (*Suricata suricatta*), je montre la haute précision et la robustesse de ce modèle pour reconnaître les comportements les plus communs des suricates (repos, vigilance, recherche de nourriture et course) non seulement avec les données de l'accéléromètre, mais aussi, dans un autre chapitre, avec des données de magnétomètre. Des avantages complémentaires des deux capteurs ont été trouvés à la suite de cela: le magnétomètre avait une plus grande robustesse à la variabilité inter-individuelle dans les comportements dynamiques, tandis que l'accéléromètre était meilleur à estimer la posture. Ensuite, la thèse présente une nouvelle approche pour résoudre les comportements à résolution plus fine. Ici, un comportement complexe est conceptualisé comme étant composé de mouvements impulsifs caractéristiques (micro-événements) produisant de brefs signaux de choc dans les données de l'accéléromètre. Des techniques de traitement des signaux nécessaires sont développées pour détecter (« seek ») les micro-événements dans les signaux d'accélération, puis un apprentissage automatique robuste (« learn ») est utilisé pour séparer les différents types de micro-événements. Cette approche « seek and learn » effectivement aide à surmonter l'obstacle de la non-stationnarité qui avait précédemment entravé les progrès dans cette direction. En utilisant des données collectées sur les modes de recherche de nourriture fine-résolution chez les suricates – recherche d'endroit pour creuser, creusage à un bras, creusage à deux bras, et mastication de proies – je montre que cette approche a une haute précision et surpasse les approches basées sur des approches d'apprentissage machine plus classiques.

Inspirée des approches adoptées dans les études humaines pour estimer la DE, cette thèse propose une nouvelle méthode d'estimation de la DE basée sur l'équivalent métabolique de la tâche (EMT), où le taux de DE pour chaque activité est donné en termes de multiples du taux métabolique au repos. La méthode basée sur l'EMT ne nécessite que des durées de comportement – obtenues, par exemple, à partir d'un modèle de reconnaissance du comportement basé sur un accéléromètre, et exploite les mesures et les relations empiriques des études antérieures sur la même ou d'autres espèces similaires. Le fait de passer par la durée du comportement plutôt que par le signal d'accélération brut aide à éviter à faire des hypothèses sur la

précision de la compensation de gravité pour estimer l'accélération dynamique du corps. Les estimations de la DE basées sur l'EMT obtenues en kilo-Joules ont montré un bon accord avec les mesures précédentes de la DE chez les suricates en utilisant la technique de référence de l'eau doublement marquée. De plus, la DE basée sur l'EMT était fortement corrélée avec l'accélération dynamique vectorielle du corps, indiquant également son potentiel d'applicabilité sur d'autres espèces.

Enfin, je boucle la boucle en appliquant ces développements « ingénierie » à la question « écologique » de comment les différences dans les traits individuels, tels que la masse corporelle, affectent la stratégie de recherche de nourriture des suricates sauvages vivant en groupe. Les suricates passent très rapidement d'un comportement à un autre, et étudier cette question par l'observation directe de plusieurs individus est ardu. En utilisant le modèle de reconnaissance du comportement à échelle grossière et fine, et la technique d'estimation DE développée dans cette thèse, j'ai trouvé que les différences de masse corporelle chez les femelles étaient fortement liées aux différences de stratégie de recherche de nourriture: les femelles plus légères recherchaient presque exclusivement des proies de surface de faible valeur énergétique tandis que les femelles plus lourdes préféraient creuser pour trouver des proies souterraines de plus grande valeur. Les inférences tirées de cette stratégie de recherche de nourriture et des distributions de proies connues dans le Kalahari suggèrent que les femelles plus légères « se contentent » de proies de surface de faible valeur afin d'éviter de perdre de la nourriture aux femelles plus lourdes dans les conflits de recherche de nourriture qu'elles sont susceptibles de perdre.

Les algorithmes développés dans cette thèse mettent en évidence le fait qu'avec des techniques de traitement des données appropriées, des capteurs de basse consommation tels que des accéléromètres peuvent fournir des estimations réalistes de la dépense énergétique en unités standard de kilo-Joules, et détecter de fines différences de comportement qui seraient difficile à déterminer même lorsqu'une observation directe d'animaux sauvages est possible.

Mots clés: accéléromètre, magnétomètre, reconnaissance de comportement, suricate, dépense énergétique, apprentissage automatique, écologie animale, validation croisée, recherche de nourriture.

Acknowledgements

The practice of observing bodily sensations objectively, equanimously, and systematically – *Vipassanā* meditation – rediscovered by Gautama the Buddha, introduced to me by my compassionate brother Rohan, and taught so carefully by Shri S.N. Goenka and his assistant teachers, helped me become more disciplined, hard-working, clear-minded, objective, and decisive. It prised me out of the grips of eight months of depression in late December 2018, and gave me the tools to face the powerful negative forces in my life head on. It would have been impossible for me to reach the end of this PhD without it.

Prof. Kamiar Aminian, my thesis director, allowed me a place and financially-protected time to work towards my PhD in his laboratory. I thank him for the faith he reposed in me to work on this wonderful project. His enormous experience helped guide the thesis work towards higher quality and rigour. I also thank him for giving me ample opportunities to contribute to teaching at EPFL.

Prof. Arpat Ozgul, the collaborating professor for this project at the University of Zurich (UZH), adjusted to the interdisciplinary nature of our meetings with aplomb, and provided insights at key moments in the project, and for that I am grateful to him. He gave me the useful advice of not subtitling incoming information.

Dr. Gabriele Cozzi, postdoctoral associate at UZH, showed me the Kalahari. He quenched my endless questions with patience, humour, and his vast knowledge of the wild and of human social behaviour. I would always feel more enthused about the project after talking to him. I was amazed by how deeply he understood the ‘engineering’ part of the project, and helped uncover ‘blind spots’ in my manuscripts. Despite English not being his first language, he helped me improve my writing style by urging me to distil each sentence into a clear set of simple words expressing the main message and nothing but. He was my window into the world of ecologists, and always introduced me to them with a glowing opinion of my work. I am grateful for his encouragement and selflessness.

I thank the jury members – Prof. Emily Shepard, Dr. Yan Ropert-Coudert, Prof. Selman Sakar, and Prof. Dominique Pioletti – for reading the thesis in depth, for their comments and questions, and for their availability during the peculiar circumstances imposed by the Covid pandemic.

Most of my PhD salary was financed by Swiss National Science Foundation (SNSF) grant no. CR32I3_159743. I would like to thank the SNSF for supporting this research, and the entire team of researchers that applied for this grant. I thank Dr. Nicolas Perony for his role in preparing the grant application, and Dr. Fabien Massé for putting him in contact with Prof. Aminian for collaboration.

Prof. Marta Manser at UZH and Prof. Tim Clutton-Brock at Cambridge University, directors of the Kalahari Meerkat Project (KMP), allowed us access to the Kalahari Meerkat Project, where all the data presented in this thesis were collected. The long-term research on meerkats was supported by a European Research Council Advanced Grant (No. 294494) to Prof. Clutton-Brock and by grants from UZH to Prof. Manser. I thank all the field managers, collaborators, and assistants for facilitating field work and helping with data collection, and in particular Dr. David Gaynor, Tim Vink, Ana Morales González, Héctor Ruiz Villar, Natasha Harrison, and Zoe Allin. Dr. Nino Maag helped me a lot during my first field visit, and answered all of my meerkat-related questions comprehensively and with great patience and clarity. I appreciate how he made me feel respected even when I asked very basic questions. He contributed to data collection for my project on his personal time, trained field assistants, and helped me understand the different meerkat behaviours. Selin Ersoy painstakingly, cheerfully, and efficiently annotated meerkat videos, and the quality of this work was essential to obtaining good results.

I am grateful to the Kalahari meerkats. After having spent about 6 hours with the meerkat group *Baobab*, collecting video footage with no other human being in sight, I experienced a moment of a kind of oneness with them that is hard to describe. This moment helped me salvage the will to continue through the most trying times.

The Biologging Symposium at Konstanz in 2017 opened my eyes to the amazing world of animals. I am grateful that the community made me – an engineering PhD student – feel ‘included’. The positive response to my talk there gave me an external reference for my work quality, which helped me build confidence in my work and in my own abilities. The Symposium also gave me the opportunity to meet and correspond with top researchers in the field such as Dr. Yan Ropert-Coudert, Dr. Akiko Kato, and Prof. Dominique Berteaux.

The one-year long review of my first paper at *Methods in Ecology and Evolution* drastically improved my critical thinking and scientific writing. This went a long way in helping me produce good first iterations for the following manuscripts. For this, I am sincerely grateful to the Handling Editor, Prof. Robert O’Hara, an anonymous and generous Associate Editor, and three anonymous reviewers.

M. Pascal Morel’s ingenuity with hardware helped me collect good quality data, and I think it was essential for obtaining good results. I appreciated the sense of wonder that would come over him when he watched meerkat videos or talked about them. I thank his wife, Mme Christine Morel, for gifting me a stuffed meerkat toy that turned out to be very effective for presenting my project to visitors to the lab. I thank Dr. Anisoara Ionescu for her readiness to help, for always encouraging me, and for expressing a childlike wonder for the meerkats. I thank Mme Francine Eglese for her positivity, humour, optimism, and help with administrative matters.

Dr. Julien Favre gave me an MSc experience that reinforced my decision to do a PhD. I will always be grateful for his help in July 2018 – despite being busy with a new-born and a new lab to run, he found the time to help me out during the

darkest weeks of my PhD by infusing me with positivity, a sense of freedom, and joy. I have scarcely come across another human being who combines such high levels of analytical and emotional intelligence. I thank Dr. Hooman Dejnabadi for the compassion, humility, and thoroughness with which he helped me with a difficult signal processing problem. His technical brilliance, kindness, and clarity are a gift to this world, and I wish him all the success he deserves. I found his monk-like demeanour and outlook very endearing.

Dr. Benedikt Fasel introduced me to Michael Ende's quote about not looking at the whole road at once but only at the next step. Dr. Christopher Moufawad el Achkar helped me get started with Weka, and I am thankful for his good humour and unhindered laughter. I learnt a lot from Dr. Mathieu Falbriard's social intelligence and the control he exercised over his emotions. I am grateful to him for listening to me patiently and without judgement in the tough times. I thank all the past and present LMAM members I had an opportunity to interact with and learn from: Abolfazl Soltani, Lena Carcreff, Salil Apte, Gaëlle Prigent, Arash Atrsaei, Mahdi Rad, Dr. Mina Baniasad, Dr. Matteo Mancuso, Yasaman Izadmehr, Martin Savary, Joaquin de Pablo Cabeza, Dr. Seyed Abdolmajid Yousefsani, Risa Ishizaki, and Hirotaka Kobayashi.

I thank all the students who gave me an opportunity to contribute to the fulfilling task of teaching during the PhD. Maiki Maalberg for her brilliant results with magnetometer data, Laure Bruyère for her tenacity and for taking leopard tortoise data analysis off the ground, Hamza Chaoui, Marion Groperrin and Hugues Vinzant for their findings on meerkat head turns, Pierre-Alexandre Léziart for his excellent work on meerkat chewing detection, Mahmoud Zgolli for his study on the amount of data needed for behaviour recognition, Gauvain Rudler for his work on identifying table tennis services from inertial sensor data, and all the students who took *Sensors in Medical Instrumentation*. Prof. Manser, along with her MSc student Seija-Mari Filli at UZH, initiated the leopard tortoise project, and I am grateful for the opportunity to have collaborated with her.

I thank all members of the catering and cleaning staff at EPFL for their wonderful work in providing healthy food with a cheery smile, and for allowing me to work in a clean space. I thank the Rolex Learning Centre for access to their amazing pan-Switzerland library network. I thank the doctoral program administrative staff members Mme Sonja Bodmer and Mme Sandra Roux for keeping me informed about the doctoral school requirements and addressing all my queries efficiently and cheerfully. I thank EPFL's excellent Mediacom team for publicising the meerkat project within and outside the school, and Dr. Dietrich Reinhard for doing so within the Life Sciences department. I thank EPFL for state-of-the-art infrastructure, the support it provides to doctoral students, and a picturesque campus environment.

Table tennis helped me strike a balance with work in the first half of my PhD. I thank all members of the Lausanne Club de Tennis de Table for making me feel so

welcome at the club. I am grateful to Alexandra Tchalakian and Colin Rochat in particular for helping me realise my childhood dream of playing some sport at the national level. I thank my former team-mates of the *première ligue régionale* Erwan Busoni, Prof. Vincent Dion, and Micaël Obrenovic for fun matches. I am grateful to Radovan Zunjanin of Serbia for transmitting his amazing knowledge of table tennis, and helping me become a stronger player and braver person. I thank Tomasz Wisniewski of Poland for urging me to believe I could win against stronger players. I thank Cédric Junillon for introducing me to Racketlon, and giving me the opportunity to serve as table tennis trainer at the Dorigny gym.

I thank Thomas Simonet for patiently teaching me French and about French culture. I thank him, Erik Mailand, Dr. Andreas Schüler and Conny Schüler for giving me the opportunity to play music on stage. I thank Dr. Saurabh Tembhrune and Dr. Sanket Patel for the small and pleasant Indian get-togethers we would do from time to time. I thank Colas Droin for teaching me about personality types and introducing me to Carl Jung's work. I thank Charlotte Moerman for her kind empathy and tips on running.

Christa and Dominique Lehnerr, with their unconditional love, gentle demeanour, and non-judgemental acceptance, transformed me as a person. From them, I learnt the physical, emotional, and relational aspects of a healthy and harmonious lifestyle. I am grateful to have spent time with other members of their family as well – Emmanuel, Marie, Jean-David, Bruno, and Morgane L'Hoste-Lehnerr. The two and a half years I spent living in their home – and our seven-year friendship – were therapeutic and spiritually elevating.

Rohan, my younger brother, introduced me to *Vipassanā* meditation. Had it not been for his own spiritual search, I would have missed this treasure. I thank him for being my first student, and for bearing the brunt of the countless mistakes I made as an early teacher. His support in mid 2018 to early 2019 in setting up spiritual visits in India for me were just invaluable, and I owe my PhD to what I found during that time. Thank you, Rohan.

There was a bumpy red brick road outside my childhood home in Chandigarh. When I was about 3-6 years old, my mother and father, walking on this road on either side of me on our way home, would suddenly put an arm below each of my armpits, hoist me up in the air, and sprint forward exclaiming '*Changdollā! Changdollā!*' as I became airborne and ecstatic. They eventually slowed down, breathless, as my feet returned to the ground. In Bengali, they call this some-metres long home-made roller coaster '*changdollā*'. The feeling of freefall would be exhilarating, and would give me a huge smile that I'd sport for the rest of the day. Looking back, I realise that my parents actually *changdollā*-ed me all the way from the foothills of the Himalayas to those of the Alps. I bow down to their selfless, lifelong sacrifices that carried me here.

~:~:~

Contents

Abstract	i
Résumé	v
Acknowledgements	ix
List of figures	xvii
List of tables	xix
Chapter 1 Introduction	1
1.1 Behaviour and energy are central to animal ecology.....	3
1.2 Overview of remote observation of wild animals.....	4
1.3 Meerkats and study area	12
1.4 Thesis goals.....	14
1.5 Thesis outline.....	16
References	19
Chapter 2 State of the Art in Animal Behaviour Recognition, and Theoretical Analysis of Sensor Data Interpretation	37
Summary.....	39
2.1 State of the art.....	41
2.1.1 The use of accelerometers to infer animal behaviour	41
2.1.2 The use of magnetometers to infer animal behaviour.....	47
2.1.3 Conclusion.....	48
2.2 Theoretical analysis	49
2.2.1 On the separation of static and dynamic acceleration.....	49
2.2.2 Magnetometer: can a rotating magnetometer measure angular velocity?.....	58
2.4 Conclusion	60
References	62

Chapter 3 | Coarse-Scale Behaviour Recognition Using Accelerometers..... 73

Abstract.....	75
3.1 Introduction.....	77
3.2 Material and Methods	80
3.2.1 Biomechanically Driven Behaviour Recognition and Validation.....	80
3.2.2 Case Study: Kalahari Meerkats.....	84
3.3 Results.....	89
3.3.1 Observed Behaviour.....	89
3.3.2 Performance Evaluation	89
3.4 Discussion.....	93
Appendices	100
References	107

Chapter 4 | Coarse-Scale Behaviour Recognition Using Magnetometers... 111

Abstract.....	113
4.1 Background.....	115
4.2 Methods	117
4.2.1 Deriving biomechanical descriptors of movement using magnetometer data	117
4.2.2 Data collection and groundtruthing.....	120
4.2.3 Developing candidate features to quantify the biomechanics of movement.....	121
4.2.4 Feature selection.....	123
4.2.5 Behaviour recognition scheme and cross-validation.....	125
4.3 Results.....	126
4.3.1 Collected data.....	126
4.3.2 Features to quantify biomechanical descriptors from triaxial magnetometer data	127
4.3.3 Performance evaluation, and comparison with accelerometer-based behaviour recognition	127
4.4 Discussion.....	130
4.5 Conclusion	136
Appendices	137

Chapter 5 | Fine-Scale Behaviour Recognition Using Accelerometers 151

Abstract.....	153
5.1 Introduction.....	155
5.2 Material and Methods	158
5.2.1 Microevent identification	158
5.2.2 Case study: Foraging in Kalahari meerkats	162
5.3 Results.....	167
5.3.1 Collected data and detected microevents	167
5.3.2 Feature selection and data resampling	167
5.3.3 Performance evaluation and comparison	169
5.4 Discussion.....	171
5.5 Conclusion	176
Appendices	178
References	210

Chapter 6 | Behavioural Time-Energy Budgets and Body Mass-Dependent Foraging Strategies in Meerkats..... 215

Abstract.....	217
6.1 Introduction.....	219
6.2 Material and Methods	221
6.2.1 Data collection and preparation	221
6.2.2 Coarse- and fine-scale behaviour recognition model (CoFiBRec)	222
6.2.3 Deriving estimates of <i>EE</i>	222
6.2.4 Descriptive metrics and statistical analysis.....	226
6.3 Results.....	228
6.3.1 Collected data.....	228
6.3.2 Time budgets.....	228
6.3.3 <i>EE</i> computation, and energy budgets.....	232
6.3.4 Behavioural differences	232
6.4 Discussion.....	237
Appendices	241
References	258

Chapter 7 Conclusion	263
7.1 Main contributions, general discussion, and limitations.....	265
7.1.1 Biomechanically driven learning	265
7.1.2 Cross-validation (CV)	270
7.1.3 MET-based energy expenditure (EE) estimation.....	272
7.1.4 Meerkat foraging strategy	274
7.2 Broader discussion and critique.....	275
7.3 Future perspectives	278
References	281
Curriculum Vitae	283

List of figures

Figure 1.1. The diversity of tagged animals.....	7
Figure 1.2. Accelerometers: an important tool for understanding animal ecology.....	8
Figure 1.3. Meerkats (<i>Suricata suricatta</i>).....	13
Figure 2.1. Metro map of animal behaviour recognition techniques.....	44
Figure 2.2. The spring-mass-damper model for a uniaxial capacitive accelerometer.....	51
Figure 2.3. Translating and rotating accelerometer in the presence of gravity.....	55
Figure 3.1. Sensor axes orientation.....	85
Figure 3.2. Biomechanically informed behaviour recognition scheme.....	87
Figure 3.3. Five-second snapshots of the four behaviours of interest.....	90
Figure 3.4. Transparent classification.....	92
Figure 4.1. Using magnetometer data to distinguish between meerkat postures.....	119
Figure 4.2. Meerkat with collar, axes, and Earth's fields.....	121
Figure 4.3. Five-second snapshots of calibrated triaxial magnetometer data.....	128
Figure 4.4. Behaviour recognition scheme.....	129
Figure 4.5. Decision boundaries and feature distributions.....	132
Figure S4.1. Magnetometer calibration.....	144
Figure S4.2. Variation in calibrated magnetometer norm (lumped).....	145
Figure S4.3. Variation in calibrated magnetometer norm by recording session.....	146
Figure 5.1. Microevent detection and characterisation.....	161
Figure 5.2. Microevent classification scheme.....	164
Figure 5.3. Observed signals and detected microevents.....	168
Figure 6.1. Meerkats equipped with collars.....	223
Figure 6.2. Coarse- and fine-scale behaviour recognition (CoFiBRec) scheme.....	224
Figure 6.3. Triaxial acceleration and recognised behaviour.....	230
Figure 6.4. Time and energy budgets by season.....	231
Figure 6.5. Body mass-dependent foraging strategies in females.....	235
Figure 6.6. The evolution of energy expenditure with time of day.....	236

Figure S6.1. Comparing VeDBA and MET-based coding.....	247
Figure S6.2. Resting duration evolution with time of day	251
Figure S6.3. Vigilance duration evolution with time of day	252
Figure S6.4. Searching duration evolution with time of day	253
Figure S6.5. One-armed digging duration evolution with time of day	254
Figure S6.6. Two-armed digging duration evolution with time of day.....	255
Figure S6.7. Chewing duration evolution with time of day	256
Figure S6.8. Running duration evolution with time of day.....	257
Figure 7.1. Leopard tortoise tagged with an accelerometer	269

List of tables

Table 1.1. Biologged variables and applications.....	5
Table 3.1. Summary of collected data.....	91
Table 3.2. EQDIST cross-validation results.....	94
Table 3.3. STRAT cross-validation results	94
Table 3.4. LOIO cross-validation results	95
Table S3.1. Measures to quantify model performance.....	100
Table S3.2. Features retained from Nathan et al., 2012.	101
Table S3.3. Feature selection	102
Table S3.4 EQDIST cross-validation results.	103
Table S3.5. STRAT cross-validation results	103
Table S3.6. LOIO cross-validation results	104
Table S3.7. Aggregate conf. matrix obtained SVM-SVM-SVM for LOIO.....	105
Table S3.8. Individual-wise performance with SVM-SVM-SVM	106
Table S3.9. Individual-wise conf. mats. with SVM-SVM-SVM.....	106
Table 4.1. Feature development	124
Table 4.2. Summary of collected data.....	126
Table 4.3. STRAT cross-validation results	131
Table 4.4. LOIO cross-validation results	131
Table S4.1. Ranking intensity features for static vs dynamic	138
Table S4.2. Ranking intensity features for foraging vs running	139
Table S4.3. Ranking periodicity features for foraging vs running.....	139
Table S4.4. Agg. conf. mat. obtained during LOIO with magnetometer data	140
Table S4.5. Agg. conf. mat. obtained accelerometer data.....	140
Table S4.6. Agg. LOIO performance with magnetometer.....	140
Table S4.7. Individual-wise performance with magnetometer	141
Table S4.8. Comparison of accelerometer vs magnetometer for LOIO.....	141
Table S4.9. Characterising variation in magnetometer norm.....	143
Table 5.1. Summary of groundtruth data and detected microevents.....	170
Table 5.2. Results with LOIO	173
Table 5.3. Effect of microevent model parameters on performance	174

Table S5.1. Definitions for terms used in the feature table.....	179
Table S5.2. List of 200 features and computation details	179
Table S5.3. The eight best features	190
Table S5.4. The family of 38 features presented in Nathan et al., 2012.....	195
Table S5.5. Feature selection from Nathan et al. 2012 (window size 1.5 s).....	196
Table S5.6. Feature selection from this study’s list (window size 1.5 s).....	197
Table S5.7. Feature selection from Nathan et al. 2012 (window size 0.3 s).....	198
Table S5.8. Feature selection from this study’s list (window size 0.3 s).....	199
Table S5.9. Individual-wise conf. mats. for μEvId	201
Table S5.10. Individual-wise μEvId performance	201
Table S5.11. Results with balanced dataset during LOIO	202
Table S5.12. Results with imbalanced dataset during LOIO	203
Table S5.13. Effect of M_{env} on μEvId performance.....	204
Table S5.14. Effect of f_c on μEvId performance	205
Table S5.15. Effect of th_{amp} on μEvId performance.....	206
Table S5.16. Effect of th_{std} on μEvId performance.....	207
Table S5.17. Effect of M_{seg} on μEvId performance	208
Table S5.18. Effect of w on μEvId performance.....	209
Table 6.1. Summary of collected data.....	229
Table S6.1. CoFiBRec parameters	243
Table S6.2. Feature computation.....	244
Table S6.3. Energetic cost of running for mammals of different mass.....	245
Table S6.4. Behaviour-wise VeDBA	247
Table S6.5. Behaviour proportions derived from previous data	248
Table S6.6. Individual-wise time budgets.....	249
Table S6.7. Individual-wise energy budgets	249
Table S6.8. Individual-wise estimates of 24-hour energy expenditure.....	250

CHAPTER 1

Introduction

1.1 | Behaviour and energy are central to animal ecology

Time and energy are fundamental currencies in the study of animal ecology (McNamara & Houston 1986; Ydenberg et al. 1994; Blanckenhorn, Preziosi & Fairbairn 1995; Butler et al. 2004).

At any given time, animals must prioritise one activity over the others, since the movements required to perform multiple activities simultaneously are mutually incompatible (McFarland 1977) – this gives rise to observed behaviour. Behaviour mediates interactions between an individual and its environment (Sih, Ferrari & Harris 2011). The behaviour of individual animals is an important determinant of survival (e.g. Tinbergen 1963; Dwyer & Lawrence 2005; Korstjens, Lehmann & Dunbar 2010; Dadda, Koolhaas & Domenici 2010), mechanism of evolution (Davies, Krebs & West 2012; Brown et al. 2013), and driver of collective behaviour (e.g. Partridge 1982; Ballerini et al. 2008; Yates et al. 2009), all of which affect population responses to changing environmental conditions and extinction risks (Caro 1999; Reed 1999; Simpson et al. 2010). Behavioural data shed light on the mechanisms and consequences of animal decision-making (e.g. McFarland 1977; Dill 1987; Strandburg-Peshkin et al. 2015), and changes in behaviour are observed when environmental conditions change (e.g. Komers 1997; Tuomainen & Candolin 2011). Detailed behavioural data can help obtain a mechanistic and predictive understanding of ecological systems (Dell et al. 2014; Nathan et al. 2008), and will help inform efforts to manage and conserve wild populations (Caro 1998; Sutherland 1998; Curio 1996; Clemmons & Buchholz 1997) in the face of increased anthropogenic impact (e.g. Sih, Ferrari & Harris 2011; Haddad et al. 2015; Ayram et al. 2017) and rapid environmental change.

Energy has been called an important medium of biological exchange between an animal and its environment (Townsend & Calow 1981; Bennett 1986). Animals must expend energy to move, e.g. during locomotion (Gold 1973; Cavagna, Heglund & Taylor 1977), as well as for internal processes during rest, such as thermoregulation (Pinshow et al. 1976; Renecker & Hudson 1986). Different activities involve different energetic costs (Weathers et al. 1984; Wilson & Culik 1993), and how animals allocate energy to different activities has implications for life-history traits and fitness (Burton et al. 2011).

1.2 | Overview of remote observation of wild animals

Studying the behaviour and energetics of wild animals living in their natural environments, however, is challenging, and even direct observation is difficult since most species either avoid humans entirely or live in habitats that makes following them around impossible (Block 2005; Rutz & Hays 2009). Another pitfall of direct observation is the risk of human presence influencing the animal's behaviour (Carpenter 1934; Schneirla 1950). Efforts to address these issues gave birth to the field of biotelemetry/biologging (Cooke et al. 2004; Ropert-Coudert et al. 2009), wherein tags containing one or multiple sensors are implanted in or attached to animal bodies to measure the quantities of interest.

1.2.1 | Sensors in biologging

The first step towards remote 'observation' of animals was said to have been taken in the 1930's, where a depth gauge was attached to a harpooned fin whale (*Balaenoptera physalus*) to measure its diving depth (Scholander 1940; Kooyman 2004; Ropert-Coudert et al. 2009). Since then, spearheaded by early developments in the telemetry of marine animals (Hussey et al. 2015) and seabirds (Wilson et al. 2002), the suite of sensed variables has expanded vastly, e.g. pressure (Butler 1989), swimming speed (Wilson & Bain 1984; Le Boeuf et al. 1992), position (e.g. Weimerskirch, Salamolard & Jouventin 1992; Juell & Westerberg 1993; Weimerskirch et al. 1993; Rempel, Rodgers & Abraham 1995; Ryan et al. 2004; Clark et al. 2006), stomach temperature (Grémillet & Plös 1994; Wilson et al. 1995; Kato et al. 1996; Bost et al. 1997), images/video (e.g. Marshall et al. 2007; Thomson & Heithaus 2014), sound (e.g. Burgess et al. 1998), biopotentials (e.g. Cooke, Thorstad, & Hinch 2004). These have been used to help answer questions concerning, for instance, activity budgets (e.g. Bodkin, Monson & Esslinger 2007; Russell et al. 2015), feeding/foraging activity (e.g. Weimerskirch, Wilson & Lys 1997; Calambokidis et al. 2007), and movement patterns (e.g. Sims et al. 2006; Weimerskirch et al. 2007; Sims et al. 2008) (Table 1.1, taken from Cooke et al. 2004).

Desired variable	Sensor details	Applications
Sound, vibration	Electret and dynamic microphones, piezoceramic benders, piezo films	Measurement of feeding activity in chewing mammals; recording communication among animals such as birds and marine mammals
Pressure in body cavities, depth in water, barometric pressure	Micromachined Si pressure transducers	Assessment of opercular rate in fish; pericardial pressure; frequency and amplitude of tail beats in water; jet pressure of squid and blood pressure in numerous animals; depth of diving animals
Acceleration, locomotor rhythms, activity	Micromachined (multi-axis) accelerometers, Hg tip switches, body temperature (death or torpor indicator for endotherms)	Assessment of movement and swimming dynamics of aquatic animals in two or three dimensions; used primarily in fish and marine mammals
Salinity of milieu, ionic strength of body fluids	Conductivity cells	Reveals information on the environmental conditions faced by the animal; assessment of the ionoregulatory status during migration of fish between salt and fresh waters; assessment of ionic balance of mammals and reptiles in arid environments
Imagery	Miniaturized versions of commercially available image capture technology such as still cameras and video	Provides the 'animal's perspective', enabling the coupling of behaviour and physiology; animal borne video is frequently used on marine mammals
Blood flow rates, swimming or flying speed	Magnetically or optically sensed rotors, drag and strain gauges, heated thermistors, ultrasonic (piezoelectric) Doppler	Quantification of several flow and speed related variables such as cardiac output (which is highly correlated with metabolic rate)
Position (e.g. of body parts), orientation	Magnetic reed switches, Hall-effect sensors, variable resistors, optical interrupter modules, ultrasound, Hall-effect magnetic field angle encoders, fluxgate compasses	Document movement of appendages, such as flippers, wings and tails; mandibular movement indicating feeding activity; migration biology
Light	Photocells, phototransistors and photodiodes, integrated circuits (light to voltage, light to frequency, some with ability to measure in several wavelength bands)	Assessment of diel activity patterns, hibernation, pollination ecology; bioluminescence of marine organisms encountered by fish or marine mammals; geolocation
Temperature	Thermistors, transistor junctions, thermocouples, heat flux sensors	Detection of gross and tissue specific temperatures associated with thermal ecology; feeding activity in marine mammals and birds
Chemistry of body fluids or external milieu	Fluors or colorimetric compounds located at tips of optical fibers or on exterior of glass capsules and excited or measured by LEDs and phototransistors; ion-specific electrodes	Quantification of blood chemistry in free-ranging animals; currently few examples of applications to ecological research although much opportunity; includes remote collection of blood samples for later analysis
Biopotentials ^b (ECG, EMG, EEG)	Biocompatible electrodes, high input impedance amplifiers	Assessment of gross patterns of locomotor activity (EMG), estimate energetics (EMG, ECG), record neural activity (EEG, neural activity); relevant to many types of animal

Table 1.1. **Biologged variables and applications.** A wide variety of animal-borne sensors have been used to measure variables that help answer a variety of questions in animal ecology. Table taken from Cooke et al. 2004. ECG: electrocardiogram; EMG: electromyogram; EEG: electroencephalogram.

The field has since grown to encompass terrestrial species as well, with many studies using position data obtained using VHF (Very High Frequency) radio, or GPS or other satellite-based tracking, often combined with environmental information such as weather, habitat and sometimes the location of other animals (Kays et al. 2015), to address questions on resource selection (e.g. D'Eon & Serrouya 2005; Johnson & Gillingham 2008; Fieberg et al. 2010), animal movement (e.g.

DeCesare, Squires & Kolbe 2005; Nams 2005; Cagnacci et al. 2010), foraging behaviour (e.g. Lesmerises et al. 2015), and predation (e.g. Hebblewhite & Merrill 2007; Merrill et al. 2010; Krofel, Skrbinšek & Kos 2013) (Urbano et al. 2010). Figure 1.1 shows a few species of different sizes and modes of locomotion living in diverse habitats, all of which have been successfully tracked using miniature body-attached sensors.

1.2.2 | The use of accelerometers and magnetometers in biologging

Accelerometers

The use of accelerometers has by far exceeded the use of other inertial sensors (such as magnetometers). Since the first use of miniature accelerometers more than two decades ago for monitoring the behaviour of Adélie penguins (*Pygoscelis adeliae*) (Yoda et al. 1999), there has been an explosion in the number of studies using animal accelerometry, with accelerometers being successfully deployed on >120 species (Brown et al. 2013).

The rapid uptake of accelerometers for animal studies may have resulted from several factors: (i) accelerometers are noninvasive, and can be attached externally to the body of the animal (Brown et al. 2013), (ii) accelerometers can record at high sampling frequencies and over a large range of accelerations (Béliveau et al. 1999), (iii) portable accelerometers have been used since the 50's to characterise human activity (e.g. Saunders et al. 1953; Morris 1973; Redmond & Hegge 1985; Reddy et al. 1991; Bouten et al. 1997) and energy expenditure (e.g. Wong et al. 1981; Montoye et al. 1983), and (iv) significant advances in solid-state technology led, and are still leading, to reduction in the size and weight of batteries, and increase in data storage capacity, resulting in lighter tags that can record more data (Wilson, Shepard & Liebsch 2008).

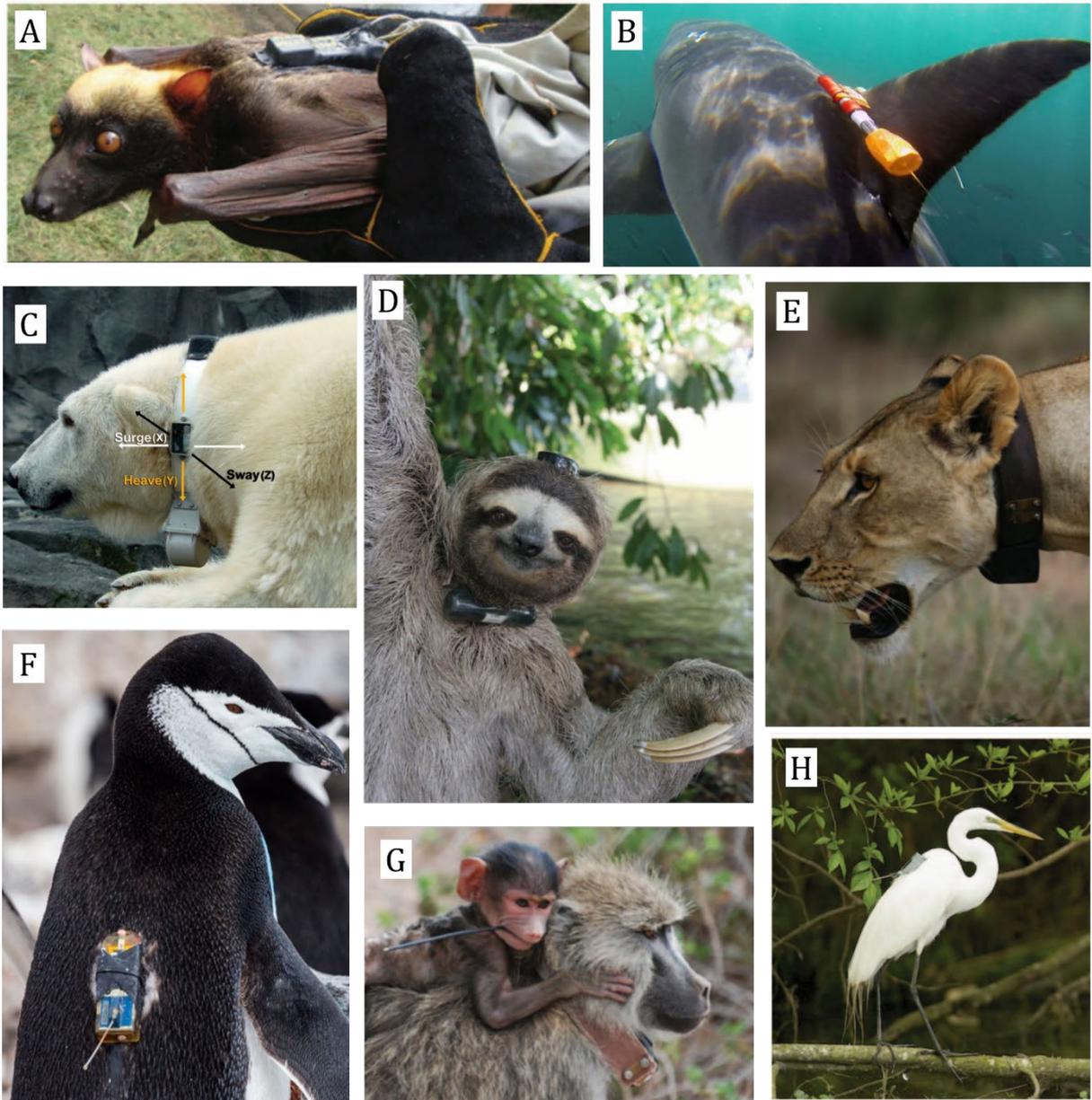


Figure 1.1. **The diversity of tagged animals.** (A) Lyle's flying fox (*Pteropus lylei*); (B) Great white shark (*Carcharodon carcharias*); (C) polar bear (*Ursus maritimus*); (D) three-toed sloth (*Bradypus variegatus*); (E) lion (*Panthera leo*); (F) chinstrap penguin (*Pygoscelis antarcticus*); (G) olive baboon (*Papio anubis*); and (H) great egret (*Ardea alba*). Images A, D, E, G, and H were taken from Kays et al. 2015. The copyright for image B belongs to Monterey Bay Aquarium. Image C was taken from Pagano et al. 2017. The copyright for image F belongs to Alex P. Taylor.

The two main applications of animal accelerometry have been to recognise common modes of behaviour (such as locomotion, resting, and feeding/foraging,

Brown et al. 2013), and estimate energy expenditure (Figure 1.2). The first application is valuable since it reveals ‘what’ animals are doing in addition to ‘where’ they are, the latter known through, say, GPS (Resheff et al. 2014). The high power consumption of GPS relative to that of accelerometers (about one order of magnitude larger, e.g. Table 11 in u-Blox Neo-M8 datasheet for GPS, and Table 5 in LSM303D datasheet for accelerometer), necessitating the use of large batteries, has meant that GPS-based behaviour recognition has been rare (Turner et al. 2000; Ungar et al. 2005; de Weerd et al. 2015). An advantage of accelerometers for behaviour recognition compared to other sensors that have been used to detect specific behaviours, such as stomach temperature sensors to detect feeding events in marine endotherms (Wilson et al. 1995), is the higher temporal resolution of behaviour. While stomach temperatures can take >1 hour to stabilise after a meal (Kuhn & Costa 2006), and feeding events <5 minutes apart cannot be distinguished in some cases (Ancel, Horning & Kooyman 1997), accelerometer-based detection of prey capture in a marine mammal can be achieved with a time resolution of <1.5 seconds (Volpov et al. 2015). The various methods employed to infer animal behaviour from accelerometer data are presented and discussed in detail in Chapter 2.

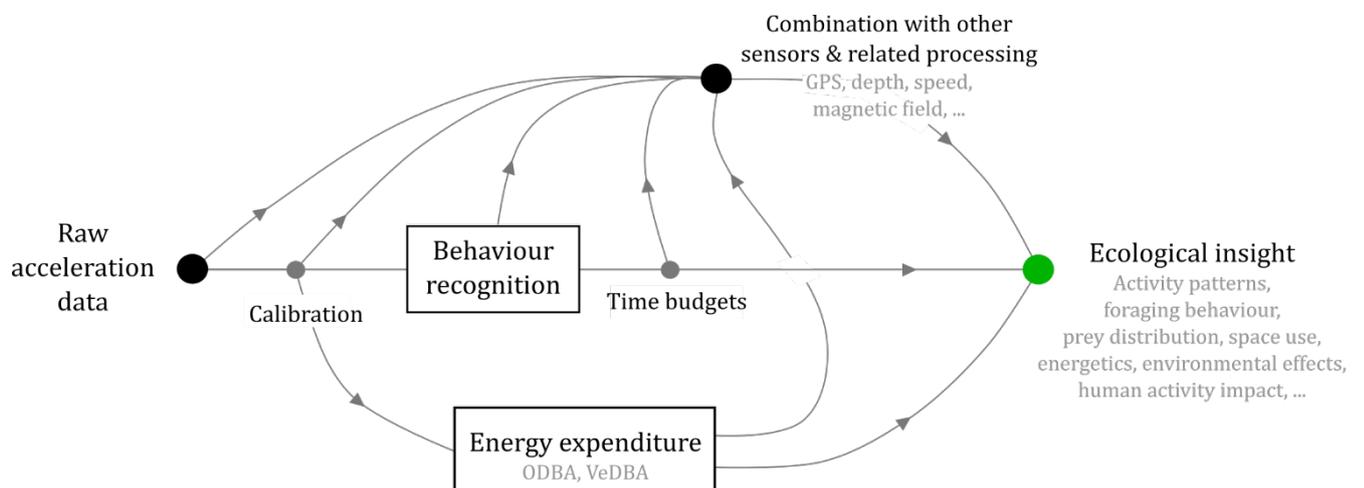


Figure 1.2. **Accelerometers: an important tool for understanding animal ecology.**

The two main applications of accelerometers have been to recognise behaviour and estimate energy expenditure. This is done either using the accelerometer alone or in combination with other complementary sensors such as GPS (for position).

The second application of the accelerometer, energy expenditure estimation, leverages the fact that the accelerometer is sensitive to both Earth's gravity ('static' acceleration) and nongravitational acceleration that is presumed to arise from acceleration of the animal's body (termed 'dynamic body acceleration', Wilson et al. 2006). Subtraction of the 'static' component from the 'dynamic' component using different methods (Qasem et al. 2012; Halsey et al. 2009) has given rise to commonly used metrics such as Overall Dynamic Body Acceleration (ODBA) (Wilson et al. 2006) and Vectorial Dynamic Body Acceleration (VeDBA) (Gleiss, Wilson & Shepard 2011) that correlate positively with the rate of oxygen consumption, an indirect measure of metabolic rate (Wilson et al. 2006; Halsey et al. 2008), across a wide range of species including mammals (e.g. Pagano & Williams 2019), birds (e.g. Elliott et al. 2013), fish (e.g. Wright et al. 2014), and reptiles (e.g. Halsey et al. 2011) (more references in Brown et al. 2013). Accelerometers, due to their non-invasiveness and ability to recognise specific behaviours, improved upon the limitations of the two existing techniques of doubly labelled water (DLW) and heart-rate loggers. DLW involves taking two blood samples from the animal (Schoeller & Van Santen 1982): DLW is injected into an animal, a blood sample is taken after DLW levels equilibrate throughout the body (30 mins to 6 hours depending on animal size), the animal is released to perform its natural behaviours, is recaptured after 24 h – 28 days (depending on animal size), another blood sample is taken, the levels of the final labelled isotope are measured, and the rate of oxygen consumption (indicative of energy expenditure) estimated. The heart rate method involves surgically implanting, and subsequently removing, a heart rate logger inside the body of the animal; energy expenditure is estimated through a calibrated linear relationship between heart rate and rate of oxygen consumption (Butler et al. 2004). DLW provides a 'lumped' value of energy expenditure for the duration of DLW elimination from the body, and thus has low temporal resolution and cannot provide behaviour-specific rates of energy expenditure, but can be used on animals weighing <1 kg. Heart-rate loggers can record at high temporal resolution for long periods and, combined with behavioural data, can provide behaviour-specific rates of energy expenditure. However, it is an invasive technique and, due to the size of the loggers and recording equipment, can only be used on animals weighing >1 kg (Butler et al. 2004).

Together, behaviour recognition and energy expenditure estimation using accelerometers as standalone sensors or in combination with other sensors, have provided valuable ecological insights (Figure 1.2) that would have been otherwise unobservable. Five diverse examples of such applications are provided here. It was shown that a central-place foraging seabird, the imperial cormorant (*Phalacrocorax atriceps*), preferentially selected foraging areas based on minimising energy expenditure rather than distance from the colony (Wilson, Quintana & Hobson 2012). In another central-place forager, the Cape gannet (*Morus capensis*), accelerometers revealed that these birds divided their foraging trips into three parts: the first part of the trip to feed themselves, then a period of resting on the water surface to digest food, and then the last part of the trip to forage once more to provide food for their offspring (Ropert-Coudert et al. 2004). Accelerometers deployed on Adélie penguins (*Pygoscelis adeliae*) coupled with a model to detect prey capture showed that two of the penguins' preferred prey items – krill and fish – had very different distributions with respect to depth below water surface, thus shedding light on the ecology of not only the tagged animal, but also its prey (Watanabe & Takahashi 2013). Accelerometers recording activity patterns in African wild dogs (*Lycaon pictus*), cheetahs (*Acinonyx jubatus*), spotted hyenas (*Crocuta crocuta*), and lions (*Panthera leo*) showed that, contrary to the idea that the subdominant wild dogs and cheetahs avoid encounters with the stronger, nocturnal spotted hyenas and lions by being active during the day and dusk, all four species could be active together to take advantage of hunting opportunities on luminous moonlit nights (Cozzi et al. 2012). Accelerometer data coupled with GPS data showed that human-induced behavioural changes in pumas (*Puma concolor*) led to higher energetic costs and resource requirements, thereby exacerbating the likelihood of human-wildlife conflict (Wang, Smith & Wilmers 2017).

Magnetometers

The first study, to my knowledge, to employ portable (fluxgate) magnetometers to study animal paths and orientation capabilities did so on green sea turtles (*Chelonia mydas*) and polar bears (*Ursus maritimus*) (Baldwin 1972). Thereafter, they have been used to obtain heading and, in combination with information on speed, perform dead-reckoning (Wilson et al. 1988; Wilson et al.

1991; Mitani et al. 2003; Bidder et al. 2015), which is the estimation of animal path and travel distance. In combination with accelerometers, magnetometers have been used to study the kinematics of swimming in marine animals (Martín López et al. 2015; Martín López et al. 2016). Magnetometer data have been combined with accelerometer and GPS data to develop movement and behaviour models for cattle (Guo et al. 2009). Studies with domestic dogs have shown that using magnetometer heading with accelerometer-based speed can reduce the required rate of GPS fix acquisition for dead-reckoning (Dewhirst et al. 2016). Magnetometers have been recently shown to detect behaviours that would be difficult to characterise using accelerometers (Williams et al. 2015; Williams et al. 2017), and that they are capable of measuring a component of angular velocity, akin to gyroscopes (Martín López et al. 2015).

1.2.3 | The effects of tags on animal behaviour and welfare¹

One underlying assumption in tagging studies is that insights gained from data collected on tagged individuals can be generalised to untagged animals of the same species, which is based on the assumption that tags do not significantly alter behaviour (Murray & Fuller 2000). Many studies do not document the effects of tags on their bearers (Vandenabeele, Wilson & Grogan 2011); however, studies generally tend to choose tagging methods they believe will minimise detrimental effects on the animal, although these decisions are often not data-driven (Murray & Fuller 2000). Some studies have, however, explicitly investigate the effects of tags on various aspects of the animal's life, e.g. on physiology and performance in fishes (Smircich & Kelly 2014), on diving behaviour of King Penguins (*Aptenodytes patagonicus*) (Ropert-Coudert et al. 2000), and on survival and foraging behaviour in meerkats (*Suricata suricatta*) (Golabek, Jordan & Clutton-Brock 2008). Some studies have also investigated the effects of tags using computer simulations (e.g. Vandenabeele et al. 2012; Pavlov, Wilson & Lucke 2007). Other than that, guidelines are available on the treatment of animals in behavioural research (Buchanan et al. 2012). Most studies

¹ Each time I've discussed or presented my project to outsiders (i.e. anyone who's not a field biologist, which happens to be quite a few people, surprisingly), I have invariably been asked, "*Et est-ce que ça les derange ?*" (and does tagging animals bother them?). Because of this, I decided to include this section giving some references that deal exclusively with this important issue.

today follow thumbs of rule for maximum tag weight relative to body mass: 2% for marine species (Winter 1996) and 5% for terrestrial species (Kenward 2000). This is, however, still an ongoing area of research where precise, data-driven guidelines have not yet been define, but it has been recommended for marine mammals, for instance, that tagging methodologies be standardised across studies to minimise the impact of tagging (Walker et al. 2012).

1.3 | Meerkats and Study Area

Meerkats (*Suricata suricatta*) are cooperatively breeding (Clutton-Brock & Manser 2016) generalist carnivorous mongooses inhabiting the semi-arid regions of southern Africa (Doolan & MacDonald 1996), with adults weighing between 0.5 – 1 kg. Meerkats live in groups of 2 – 50 individuals (Clutton-Brock & Manser 2016) (Figure 1.3A), where a single dominant female monopolises reproduction (Clutton-Brock et al. 1999), and all adults help to guard and feed the young (Clutton-Brock et al. 1998). Adults follow a coordinated system of vigilance, where they take turns to go on guard (Clutton-Brock et al. 1998) (Figure 1.3B), and communicate through calls to signal the presence or absence of predatory threat (Townsend, Zöttl & Manser 2011). Their main predators are martial eagles (*Polemaetus belliocous*), pale chanting goshawks (*Melierax canorus*), jackals (*Canis mesomelas*), and snakes (Manser 2001). Meerkats mainly feed on insects, with occasional contribution from reptiles and scorpions (Figure 1.3C) (Doolan & MacDonald 1996) (more on prey in Chapter 6). Much of their prey lies below the surface of the sand, which they dig out through complex, intensive digging manoeuvres involving alternating or synchronised forelimb swipes (more on digging style in Chapter 5). Meerkats face extreme seasonal fluctuations that not only affect their life-history responses (Paniw et al. 2019), but also their day-to-day behaviour: in the summer, meerkats avoid the hot midday sun by retreating into burrows for a ‘siesta’, whereas in winter they typically forage throughout the day (Doolan & MacDonald 1996).



Figure 1.3. **Meerkats (*Suricata suricatta*)**. (A) Group-living animals; (B) meerkats often display vigilance to keep an eye out for natural predators in the sky, such as the martial eagle (*Polemaetus bellicosus*) (not in picture); and (C) meerkats mainly prey on insects, with valuable supplements from larger prey such as geckos and, in the image shown, scorpions. Image credits for A and C belong to Arpat Ozgul. Image credit for B belongs to Kamiar Aminian.

The study area was the Kalahari Meerkat Project (KMP) situated on the Kuruman River Reserve (26° 59' S, 21° 50' E) in the South African Kalahari (Maag 2019). Meerkats with territories (of size ~2 km², Jordan, Cherry & Manser 2007) on the reserve have been habituated to close human observation (<1 m), and can be weighed up to thrice a day (morning, noon, evening) by making them climb onto electronic balances with small rewards of water or hard-boiled egg (Clutton-Brock & Manser 2016). Radio collars are fitted to the dominant individuals in each group to easily locate groups (Scantlebury et al. 2002). As noted earlier, radio collars have been found not to affect survival or foraging behaviour in meerkats (Golabek, Jordan & Clutton-Brock 2008). All individuals can be identified with the help of unique

dye-mark combinations on the body, and their ages are known precisely (± 5 days) (Bousquet 2011). The daily energy expenditure of females and pups has previously been measured using DLW (Scantlebury et al. 2002).

Typically, *ad libitum* focal observations (Altmann 1974) of ~15-minute duration of individuals are conducted as part of a long-term study running since 1993 at the KMP. However, behaviours typically noted down do not include the different foraging modes, such as searching, one-armed digging and two-armed digging (more on digging modes in Chapter 5), and doing so visually, accurately, and in real time would in any case be difficult given that meerkats transition between behaviours rapidly. This makes fine-scale observation of meerkat foraging behaviour for lengthy periods of time (say from dawn to dusk) very difficult to conduct manually, and thus detailed information on budgeting of time and energy between the different foraging modes in meerkats is neither available nor feasible to obtain using traditional methods of direct observation. Because of this, even though it is known that foraging competition interactions are affected by social hierarchy in meerkat groups (Madden et al. 2009; Flower 2011), little is known about the effects on foraging strategies.

1.4 | Thesis Goals

Despite the enormous progress in the use of accelerometers and magnetometers for the elucidation of various aspects of animal behaviour, there are nevertheless some gaps which, if filled, could enrich current knowledge and methodology.

For accelerometers, I identified four such gaps:

- (i) It remains unclear how precisely total acceleration can be split into ‘static’ (gravitational) and ‘dynamic’ (nongravitational) acceleration – an assumption upon which virtually all accelerometer-based behaviour recognition and energy expenditure estimation techniques are based. This is developed in Chapter 2.
- (ii) Some commonalities in signals recorded for various behaviours (such as the distinction of static versus dynamic behaviours, the concept of

vigorous activity producing high-amplitude signals, and periodicity of locomotion) and methodology for behaviour recognition (application of state-of-the-art, robust machine learning algorithms) have been hinted at in several studies but have not yet been brought together in a single, simple, standardised framework aiming to recognise behaviour across species. Chapter 3 targets a generic approach to address this issue.

- (iii) Fine-scale behaviours have not yet been detected with high accuracy, and there is a lack of methodology dedicated to automatically recognise the complex signals recorded during such activities. The new approach presented in Chapter 5 aims to fill this gap.
- (iv) The potential of other techniques to estimate energy expenditure from accelerometer data that have been used in human studies, such as those based on using ‘metabolic equivalent of task’ (MET) (Ainsworth et al. 2000; Bonomi et al. 2009), has not been explored for animal studies yet. In Chapter 6, a MET-based technique for energy expenditure estimation is proposed for animal studies, and compared with current approaches.

For magnetometers I identified two such gaps:

- (i) Though the magnetometer has been shown to be capable of measuring a dynamic component of movement – angular velocity (Martín López et al. 2015) – which, along with its ability to measure static tilt, makes it a good candidate for recognising behaviour, its limitations to measure angular velocity have not been made explicit yet. This is developed further in Chapter 2.
- (ii) That the magnetometer displays some qualitative similarities to the accelerometer (Williams et al. 2017) motivates the investigation of its potential to recognise common animal behaviours such as locomotion, resting, and feeding/foraging as the accelerometer has commonly been shown to do (Brown et al. 2013), but this hasn’t been explored yet. Doing so will especially be significant for studies looking to leverage the complementary strengths of multiple sensors present within the

same tagging device (Williams et al. 2019). Chapter 4 aims to address this issue.

This thesis has two goals. The first, an ‘engineering’ goal, is to address these gaps in a way that yields new methodology that can be applied accurately and robustly across species. The second, an ‘ecological’ goal, is to apply this methodology to investigate how meerkats allocate time and energy to different fine-scale foraging modes, and whether different foraging strategies arise in individuals with different traits, e.g. body mass and age.

1.5 | Thesis Outline

Chapter 2 presents an overview of currently used methods to infer animal behaviour using accelerometers, and provides a theoretical analysis – based on the physics of the sensing mechanism of the accelerometer – of the common assumption that a window-wise mean of the acceleration signal isolates the gravitational component of acceleration, and that subtracting it from total acceleration yields the acceleration of the animal’s body. Further, this chapter makes an explicit connection between magnetometer data and a dynamic aspect of movement: angular velocity. This, along with the magnetometer’s known capability of measuring static tilt, prepares the ground for usage of magnetometers in recognising common animal behaviours in much the same way as accelerometers.

Chapter 3 presents a new biomechanical approach to identify common ‘coarse-scale’ animal behaviours – such as resting, vigilance, foraging, and running – from accelerometer data. In this method, published in *Methods in Ecology and Evolution* in 2019, behaviour is characterised using only three features, one each to describe posture, movement intensity, and movement periodicity. Robust machine learning algorithms are used to find linear decision boundaries between classes in a hierarchical approach that first separates ‘static’ (resting, vigilance) and ‘dynamic’ (foraging, running) behaviours, and then the two sets of behaviours. The method is showcased on data collected on 10 meerkats.

Chapter 4 demonstrates that magnetometers can be used to identify the very same coarse-scale behaviours in Chapter 3, using the same number of features and biomechanical principles of posture, movement intensity, and movement periodicity and hierarchical separation of classes, with similar accuracy but greater robustness to inter-individual robustness compared to the accelerometer. This was published as a methodological paper in *Movement Ecology* in 2019.

Chapter 5 tackles the problem of identifying fine-scale behaviours from accelerometer data. Fine-scale behaviours often generate nonstationary signals that are difficult to characterise using the common technique based on computing statistical and frequency-domain features in a moving window of pre-set size. In the proposed method, complex behaviour is conceptualised as being composed of ‘microevents’: brief shock signals in accelerometer data. The ‘seek and learn’ approach first locates microevents in the signal using the technique of enveloping, and then, using features computed on a window centred at each microevent’s location and robust machine learning algorithms, learns differences between and separates different types of microevents in a hierarchical classifier. The method is then applied to separate fine-scale foraging modes in data collected on 10 meerkats – searching, one-armed digging, two-armed digging, and prey chewing – with high accuracy and robustness to inter-individual variation and changes in model parameters. This study is under review at *Methods in Ecology and Evolution*.

Chapter 6 has two components presented in a single study. In the first component, a new method to estimate energy expenditure, inspired from human studies, is presented. In this method, energy expenditure is equal to the sum of the products of metabolic equivalents of task (MET) values for each behaviour with duration of that behaviour. MET-based energy expenditure is compared with the common technique in animal studies of vectorial dynamic body acceleration (VeDBA). In the second component of this study, the full coarse- and fine-scale behaviour recognition (CoFiBRec) model along with MET-based energy expenditure estimation is first used to characterise energy expenditure from 1-6 days-long accelerometer data collected from 26 meerkats; it is found that this method provides values in kilo-Joules that are not statistically different from actual values measured in a previous study using doubly-labelled water. Further, we apply CoFiBRec and

MET-based energy expenditure estimation to investigate body mass-based differences in foraging strategy and efficiency in meerkats.

Chapter 7 provides a general discussion and critique of the overall results, and perspectives for the future.

References

- Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, A. M., Strath, S. J., ... & Jacobs, D. R. (2000). Compendium of physical activities: an update of activity codes and MET intensities. *Medicine & science in sports & exercise*, 32(9), S498-S516.
- Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour*, 49(3-4), 227-266.
- Ancel, A., Horning, M., & Kooyman, G. L. (1997). Prey ingestion revealed by oesophagus and stomach temperature recordings in cormorants. *Journal of Experimental Biology*, 200(1), 149-154.
- Arnould, J. P., Boyd, I. L., & Speakman, J. R. (1996). The relationship between foraging behaviour and energy expenditure in Antarctic fur seals. *Journal of Zoology*, 239(4), 769-782.
- Ayram, C. A. C., Mendoza, M. E., Etter, A., & Salicrup, D. R. P. (2017). Anthropogenic impact on habitat connectivity: A multidimensional human footprint index evaluated in a highly biodiverse landscape of Mexico. *Ecological Indicators*, 72, 895-909.
- Baldwin, H. A. (1972). Long range radio tracking of sea turtles and polar bear: Instrumentation and preliminary results.
- Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., ... & Viale, M. (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.
- Béliveau, A., Spencer, G. T., Thomas, K. A., & Roberson, S. L. (1999). Evaluation of MEMS capacitive accelerometers. *IEEE Design & Test of Computers*, 16(4), 48-56.
- Bennett, A. F. (1986). Measuring behavioral energetics. *Predator-prey relationships*, 69-81.

- Bidder, O. R., Walker, J. S., Jones, M. W., Holton, M. D., Urge, P., Scantlebury, D. M., ... & Wilson, R. P. (2015). Step by step: reconstruction of terrestrial animal movement paths by dead-reckoning. *Movement ecology*, 3(1), 23.
- Blanckenhorn, W. U., Preziosi, R. F., & Fairbairn, D. J. (1995). Time and energy constraints and the evolution of sexual size dimorphism—to eat or to mate?. *Evolutionary Ecology*, 9(4), 369-381.
- Block, B. A. (2005). Physiological ecology in the 21st century: advancements in biologging science. *Integrative and Comparative Biology*, 45(2), 305-320.
- Bodkin, J. L., Monson, D. H., & Esslinger, G. G. (2007). Activity budgets derived from time-depth recorders in a diving mammal. *The Journal of Wildlife Management*, 71(6), 2034-2044.
- Bonomi, A. G., Plasqui, G., Goris, A. H., & Westerterp, K. R. (2009). Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. *Journal of applied physiology*, 107(3), 655-661.
- Bost, C. A., Georges, J. Y., Guinet, C., Cherel, Y., Pütz, K., Charrassin, J. B., ... & Le Maho, Y. (1997). Foraging habitat and food intake of satellite-tracked king penguins during the austral summer at Crozet Archipelago. *Marine Ecology Progress Series*, 150, 21-33.
- Bousquet, CA (2011). *Group decision-making in meerkats (Suricata suricatta)* (Doctoral dissertation, University of Zurich).
- Bouten, C. V., Koekkoek, K. T., Verduin, M., Kodde, R., & Janssen, J. D. (1997). A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE transactions on biomedical engineering*, 44(3), 136-147.
- Brown, D. D., Kays, R., Wikelski, M., Wilson, R., & Klimley, A. P. (2013). Observing the unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry*, 1(1), 20.
- Buchanan, K., Burt de Perera, T., Carere, C., Carter, T., Hailey, A., Hubrecht, R., ... & Sneddon, L. (2012). Guidelines for the treatment of animals in behavioural research and teaching. *Animal Behaviour*, 83(1), 301-309.

- Burgess, W. C., Tyack, P. L., Le Boeuf, B. J., & Costa, D. P. (1998). A programmable acoustic recording tag and first results from free-ranging northern elephant seals. *Deep Sea Research Part II: Topical Studies in Oceanography*, 45(7), 1327-1351.
- Burton, T., Killen, S. S., Armstrong, J. D., & Metcalfe, N. B. (2011). What causes intraspecific variation in resting metabolic rate and what are its ecological consequences?. *Proceedings of the Royal Society B: Biological Sciences*, 278(1724), 3465-3473.
- Butler, P. J. (1989). Telemetric recording of physiological data from free-living animals. *Towards a More Exact Ecology (Grubb, PJ and Whittaker, JB, eds)*, 63-84.
- Butler, P. J., Green, J. A., Boyd, I. L., & Speakman, J. R. (2004). Measuring metabolic rate in the field: the pros and cons of the doubly labelled water and heart rate methods. *Functional ecology*, 18(2), 168-183.
- Cagnacci, F., Boitani, L., Powell, R. A., & Boyce, M. S. (2010). Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges.
- Calambokidis, J., Schorr, G. S., Steiger, G. H., Francis, J., Bakhtiari, M., Marshall, G., ... & Robertson, K. (2007). Insights into the underwater diving, feeding, and calling behavior of blue whales from a suction-cup-attached video-imaging tag (CRITTERCAM). *Marine Technology Society Journal*, 41(4), 19-29.
- Caro, T. (Ed.). (1998). *Behavioral ecology and conservation biology*. Oxford University Press on Demand.
- Caro, T. (1999). The behaviour–conservation interface. *Trends in ecology & evolution*, 14(9), 366-369.
- Carpenter, C. R. (1934). A field study of the behavior and social relations of howling monkeys. *Comparative psychology monographs*.
- Cavagna, G. A., Heglund, N. C., & Taylor, C. R. (1977). Mechanical work in terrestrial locomotion: two basic mechanisms for minimizing energy

- expenditure. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 233(5), R243-R261.
- Clark, P. E., Johnson, D. E., Kniep, M. A., Jermann, P., Huttash, B., Wood, A., ... & Titus, K. (2006). An advanced, low-cost, GPS-based animal tracking system. *Rangeland Ecology & Management*, 59(3), 334-340.
- Clemmons, J. R., & Buchholz, R. (Eds.). (1997). *Behavioral approaches to conservation in the wild*. Cambridge University Press.
- Clutton-Brock, T. H., Gaynor, D., Kansky, R., MacColl, A. D. C., McIlrath, G., Chadwick, P., ... & Skinner, J. D. (1998). Costs of cooperative behaviour in suricates (*Suricata suricatta*). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1392), 185-190.
- Clutton-Brock, T. H., Maccoll, A., Chadwick, P., Gaynor, D., Kansky, R., & Skinner, J. D. (1999). Reproduction and survival of suricates (*Suricata suricatta*) in the southern Kalahari. *African Journal of Ecology*, 37(1), 69-80.
- Clutton-Brock, T. H., & Manser, M. (2016). Meerkats: cooperative breeding in the Kalahari. *Cooperative breeding in vertebrates*, 294, 317.
- 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 & evolution*, 19(6), 334-343.
- Cooke, S. J., Thorstad, E. B., & Hinch, S. G. (2004). Activity and energetics of free-swimming fish: insights from electromyogram telemetry. *Fish and Fisheries*, 5(1), 21-52.
- Cozzi, G., Broekhuis, F., McNutt, J. W., Turnbull, L. A., Macdonald, D. W., & Schmid, B. (2012). Fear of the dark or dinner by moonlight? Reduced temporal partitioning among Africa's large carnivores. *Ecology*, 93(12), 2590-2599.
- Curio, E. (1996). Conservation needs ethology. *Trends in ecology & evolution*, 11(6), 260-263.
- Dadda, M., Koolhaas, W. H., & Domenici, P. (2010). Behavioural asymmetry affects escape performance in a teleost fish. *Biology letters*, 6(3), 414-417.

- Davies, N. B., Krebs, J. R., & West, S. A. (2012). *An introduction to behavioural ecology*. John Wiley & Sons.
- DeCesare, N. J., Squires, J. R., & Kolbe, J. A. (2005). Effect of forest canopy on GPS-based movement data. *Wildlife Society Bulletin*, 33(3), 935-941.
- Dell, A. I., Bender, J. A., Branson, K., Couzin, I. D., de Polavieja, G. G., Noldus, L. P., ... & Brose, U. (2014). Automated image-based tracking and its application in ecology. *Trends in ecology & evolution*, 29(7), 417-428.
- D'Eon, R. G., & Serrouya, R. (2005). Mule deer seasonal movements and multiscale resource selection using global positioning system radiotelemetry. *Journal of Mammalogy*, 86(4), 736-744.
- de Weerd, N., van Langevelde, F., van Oeveren, H., Nolet, B. A., Kölzsch, A., Prins, H. H., & de Boer, W. F. (2015). Deriving animal behaviour from high-frequency GPS: tracking cows in open and forested habitat. *Plos one*, 10(6).
- Dewhirst, O. P., Evans, H. K., Roskilly, K., Harvey, R. J., Hubel, T. Y., & Wilson, A. M. (2016). Improving the accuracy of estimates of animal path and travel distance using GPS drift-corrected dead reckoning. *Ecology and evolution*, 6(17), 6210-6222.
- Dill, L. M. (1987). Animal decision making and its ecological consequences: the future of aquatic ecology and behaviour. *Canadian Journal of Zoology*, 65(4), 803-811.
- Doolan, S. P., & MacDonald, D. W. (1996b). Diet and foraging behaviour of group-living meerkats, *Suricata suricatta*, in the southern Kalahari. *Journal of Zoology*, 239(4), 697-716.
- Dwyer, C. M., & Lawrence, A. B. (2005). A review of the behavioural and physiological adaptations of hill and lowland breeds of sheep that favour lamb survival. *Applied animal behaviour science*, 92(3), 235-260.
- Elliott, K. H., Le Vaillant, M., Kato, A., Speakman, J. R., & Ropert-Coudert, Y. (2013). Accelerometry predicts daily energy expenditure in a bird with high activity levels. *Biology letters*, 9(1), 20120919.

- Fieberg, J., Matthiopoulos, J., Hebblewhite, M., Boyce, M. S., & Frair, J. L. (2010). Correlation and studies of habitat selection: problem, red herring or opportunity?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550), 2233-2244.
- Flower, T. P. (2011). *Competition for food in meerkats (Suricata suricatta)* (Doctoral dissertation, University of Pretoria).
- Gleiss, A. C., Wilson, R. P., & Shepard, E. L. (2011). Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure. *Methods in Ecology and Evolution*, 2(1), 23-33.
- Golabek, K. A., Jordan, N. R., & Clutton-Brock, T. H. (2008). Radiocollars do not affect the survival or foraging behaviour of wild meerkats. *Journal of Zoology*, 274(3), 248-253.
- Gold, A. (1973). Energy expenditure in animal locomotion. *Science*, 181(4096), 275-276.
- Grémillet, D. J., & Plös, A. L. (1994). The use of stomach temperature records for the calculation of daily food intake in cormorants. *Journal of Experimental Biology*, 189(1), 105-115.
- Guo, Y., Poulton, G., Corke, P., Bishop-Hurley, G. J., Wark, T., & Swain, D. L. (2009). Using accelerometer, high sample rate GPS and magnetometer data to develop a cattle movement and behaviour model. *Ecological Modelling*, 220(17), 2068-2075.
- Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., ... & Cook, W. M. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science advances*, 1(2), e1500052.
- Halsey, L. G., Shepard, E. L., Hulston, C. J., Venables, M. C., White, C. R., Jeukendrup, A. E., & Wilson, R. P. (2008). Acceleration versus heart rate for estimating energy expenditure and speed during locomotion in animals: tests with an easy model species, *Homo sapiens*. *Zoology*, 111(3), 231-241.

- Halsey, L. G., Green, J. A., Wilson, R. P., & Frappell, P. B. (2009). Accelerometry to estimate energy expenditure during activity: best practice with data loggers. *Physiological and Biochemical Zoology*, *82*(4), 396-404.
- Halsey, L. G., Jones, T. T., Jones, D. R., Liebsch, N., & Booth, D. T. (2011). Measuring energy expenditure in sub-adult and hatchling sea turtles via accelerometry. *PLoS One*, *6*(8). Hebblewhite, M., & Merrill, E. H. (2007). Multiscale wolf predation risk for elk: does migration reduce risk?. *Oecologia*, *152*(2), 377-387.
- Hussey, N. E., Kessel, S. T., Aarestrup, K., Cooke, S. J., Cowley, P. D., Fisk, A. T., ... & Flemming, J. E. M. (2015). Aquatic animal telemetry: a panoramic window into the underwater world. *Science*, *348*(6240), 1255642.
- Jepsen, N., Koed, A., Thorstad, E. B., & Baras, E. (2002). Surgical implantation of telemetry transmitters in fish: how much have we learned?. In *Aquatic Telemetry* (pp. 239-248). Springer, Dordrecht.
- Johnson, C. J., & Gillingham, M. P. (2008). Sensitivity of species-distribution models to error, bias, and model design: an application to resource selection functions for woodland caribou. *Ecological Modelling*, *213*(2), 143-155.
- Jordan, N. R., Cherry, M. I., & Manser, M. B. (2007). Latrine distribution and patterns of use by wild meerkats: implications for territory and mate defence. *Animal Behaviour*, *73*(4), 613-622.
- Juell, J. E., & Westerberg, H. (1993). An ultrasonic telemetric system for automatic positioning of individual fish used to track Atlantic salmon (*Salmo salar* L.) in a sea cage. *Aquacultural engineering*, *12*(1), 1-18.
- Kato, A., Naito, Y., Watanuki, Y., & Shaughnessy, P. D. (1996). Diving pattern and stomach temperatures of foraging king cormorants at subantarctic Macquarie Island. *The Condor*, *98*(4), 844-848.
- Kawabe, R., Nashimoto, K., Hiraishi, T., Naito, Y., & Sato, K. (2003). A new device for monitoring the activity of freely swimming flatfish, Japanese flounder *Paralichthys olivaceus*. *Fisheries Science*, *69*(1), 3-10.

- Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, 348(6240), aaa2478.
- Kenward, R. E. (2000). *A manual for wildlife radio tagging*. Academic press.
- Komers, P. E. (1997). Behavioural plasticity in variable environments. *Canadian Journal of Zoology*, 75(2), 161-169.
- Kooyman, G. L. (1965). Techniques used in measuring diving capacities of Weddell seals. *Polar Record*, 12(79), 391-394.
- Kooyman, G. L., Ponganis, P. J., Castellini, M. A., Ponganis, E. P., Ponganis, K. V., Thorson, P. H., ... & LeMaho, Y. (1992). Heart rates and swim speeds of emperor penguins diving under sea ice. *Journal of Experimental Biology*, 165(1), 161-180.
- Kooyman, G. L. (2004). Genesis and evolution of bio-logging devices: 1963-2002.
- Korstjens, A. H., Lehmann, J., & Dunbar, R. I. M. (2010). Resting time as an ecological constraint on primate biogeography. *Animal Behaviour*, 79(2), 361-374.
- Krofel, M., Skrbinšek, T., & Kos, I. (2013). Use of GPS location clusters analysis to study predation, feeding, and maternal behavior of the Eurasian lynx. *Ecological research*, 28(1), 103-116.
- Le Boeuf, B. J., Naito, Y., Asaga, T., Crocker, D., & Costa, D. P. (1992). Swim speed in a female northern elephant seal: metabolic and foraging implications. *Canadian Journal of Zoology*, 70(4), 786-795.
- Lesmerises, R., Rebouillat, L., Dussault, C., & St-Laurent, M. H. (2015). Linking GPS telemetry surveys and scat analyses helps explain variability in black bear foraging strategies. *PloS one*, 10(7).
- Levin, S. A. (1992). The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology*, 73(6), 1943-1967.
- Lucas, M. C., Johnstone, A. D., & Priede, I. G. (1993). Use of physiological telemetry as a method of estimating metabolism of fish in the natural environment. *Transactions of the American Fisheries Society*, 122(5), 822-833.

- Maag, N. (2019). *Mechanisms and demographic consequences of dispersal in the cooperatively breeding meerkat* (Doctoral dissertation, University of Zurich).
- Madden, J. R., Drewe, J. A., Pearce, G. P., & Clutton-Brock, T. H. (2009). The social network structure of a wild meerkat population: 2. Intragroup interactions. *Behavioral Ecology and Sociobiology*, *64*(1), 81.
- Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1483), 2315-2324.
- Marshall, G., Bakhtiari, M., Shepard, M., Tweedy, I. I. I., Rasch, D., Abernathy, K., ... & Heithaus, M. R. (2007). An advanced solid-state animal-borne video and environmental data-logging device ("Critttercam") for marine research. *Marine Technology Society Journal*, *41*(2), 31-38.
- Martín López, L. M., Miller, P. J., de Soto, N. A., & Johnson, M. (2015). Gait switches in deep-diving beaked whales: biomechanical strategies for long-duration dives. *Journal of Experimental Biology*, *218*(9), 1325-1338.
- Martín López, L. M., de Soto, N. A., Miller, P., & Johnson, M. (2016). Tracking the kinematics of caudal-oscillatory swimming: a comparison of two on-animal sensing methods. *Journal of Experimental Biology*, *219*(14), 2103-2109.
- McFarland, D. J. (1977). Decision making in animals. *Nature*, *269*(5623), 15-21.
- McNamara, J. M., & Houston, A. I. (1986). The common currency for behavioral decisions. *The American Naturalist*, *127*(3), 358-378.
- Merrill, E., Sand, H., Zimmermann, B., McPhee, H., Webb, N., Hebblewhite, M., ... & Frair, J. L. (2010). Building a mechanistic understanding of predation with GPS-based movement data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1550), 2279-2288.
- Mitani, Y., Sato, K., Ito, S., Cameron, M. F., Siniff, D. B., & Naito, Y. (2003). A method for reconstructing three-dimensional dive profiles of marine mammals using geomagnetic intensity data: results from two lactating Weddell seals. *Polar Biology*, *26*(5), 311-317.

- Montoye, H. J., Washburn, R., Servais, S., Ertl, A., Webster, J. G., & Nagle, F. J. (1983). Estimation of energy expenditure by a portable accelerometer. *Medicine and Science in Sports and Exercise*, *15*(5), 403-407.
- Morris, J. R. W. (1973). Accelerometry—A technique for the measurement of human body movements. *Journal of biomechanics*, *6*(6), 729-736.
- Murray, D. L., & Fuller, M. R. (2000). A critical review of the effects of marking on the biology of vertebrates. *Research techniques in animal ecology: controversies and consequences*, 15-64.
- Nams, V. O. (2005). Using animal movement paths to measure response to spatial scale. *Oecologia*, *143*(2), 179-188.
- Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., & Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, *105*(49), 19052-19059.
- O'Dor, R. (2002). Telemetered cephalopod energetics: swimming, soaring, and blimping. *Integrative and Comparative Biology*, *42*(5), 1065-1070.
- Pagano, A. M., Rode, K. D., Cutting, A., Owen, M. A., Jensen, S., Ware, J. V., ... & Middel, K. R. (2017). Using tri-axial accelerometers to identify wild polar bear behaviors. *Endangered Species Research*, *32*, 19-33.
- Pagano, A. M., & Williams, T. M. (2019). Estimating the energy expenditure of free-ranging polar bears using tri-axial accelerometers: A validation with doubly labeled water. *Ecology and evolution*, *9*(7), 4210-4219.
- Paniw, M., Maag, N., Cozzi, G., Clutton-Brock, T., & Ozgul, A. (2019). Life history responses of meerkats to seasonal changes in extreme environments. *Science*, *363*(6427), 631-635.
- Partridge, B. L. (1982). The structure and function of fish schools. *Scientific american*, *246*(6), 114-123.
- Pavlov, V. V., Wilson, R. P., & Lucke, K. (2007). A new approach to tag design in dolphin telemetry: Computer simulations to minimise deleterious effects. *Deep Sea Research Part II: Topical Studies in Oceanography*, *54*(3-4), 404-414.

- Pinshow, B., Fedak, M. A., Battles, D. R., & Schmidt-Nielsen, K. (1976). Energy expenditure for thermoregulation and locomotion in emperor penguins. *American Journal of Physiology-Legacy Content*, 231(3), 903-912.
- Qasem, L., Cardew, A., Wilson, A., Griffiths, I., Halsey, L. G., Shepard, E. L., ... & Wilson, R. (2012). Tri-axial dynamic acceleration as a proxy for animal energy expenditure; should we be summing values or calculating the vector?. *PloS one*, 7(2).
- Reddy, N. P., Canilang, E. P., Casterline, J., Rane, M. B., Joshi, A. M., Thomas, R., & Candadai, R. (1991). Noninvasive acceleration measurements to characterize the pharyngeal phase of swallowing. *Journal of biomedical engineering*, 13(5), 379-383.
- Redmond, D. P., & Hegge, F. W. (1985). Observations on the design and specification of a wrist-worn human activity monitoring system. *Behavior Research Methods, Instruments, & Computers*, 17(6), 659-669.
- Reed, J. M. (1999). The role of behavior in recent avian extinctions and endangerments. *Conservation Biology*, 13(2), 232-241.
- Rempel, R. S., Rodgers, A. R., & Abraham, K. F. (1995). Performance of a GPS animal location system under boreal forest canopy. *The Journal of wildlife management*, 543-551.
- Renecker, L. A., & Hudson, R. J. (1986). Seasonal energy expenditures and thermoregulatory responses of moose. *Canadian Journal of Zoology*, 64(2), 322-327.
- Resheff, Y. S., Rotics, S., Harel, R., Spiegel, O., & Nathan, R. (2014). AcceleRater: a web application for supervised learning of behavioral modes from acceleration measurements. *Movement ecology*, 2(1), 27.
- Ropert-Coudert, Y., Bost, C. A., Handrich, Y., Bevan, R. M., Butler, P. J., Woakes, A. J., & Le Maho, Y. (2000). Impact of externally attached loggers on the diving behaviour of the king penguin. *Physiological and Biochemical Zoology*, 73(4), 438-444.

- Ropert-Coudert, Y., Grémillet, D., Kato, A., Ryan, P. G., Naito, Y., & Le Maho, Y. (2004). A fine-scale time budget of Cape gannets provides insights into the foraging strategies of coastal seabirds. *Animal Behaviour*, *67*(5), 985-992.
- Ropert-Coudert, Y., Beaulieu, M., Hanuise, N., & Kato, A. (2009). Diving into the world of biologging. *Endangered Species Research*, *10*, 21-27.
- Russell, D. J., McClintock, B. T., Matthiopoulos, J., Thompson, P. M., Thompson, D., Hammond, P. S., ... & McConnell, B. J. (2015). Intrinsic and extrinsic drivers of activity budgets in sympatric grey and harbour seals. *Oikos*, *124*(11), 1462-1472.
- Ryan, P. G., Petersen, S. L., Peters, G., & Grémillet, D. (2004). GPS tracking a marine predator: the effects of precision, resolution and sampling rate on foraging tracks of African Penguins. *Marine biology*, *145*(2), 215-223.
- Saunders, J. D. C., Inman, V. T., & Eberhardt, H. D. (1953). The major determinants in normal and pathological gait. *Journal of Bone and Joint Surgery*, 35-543.
- Scantlebury, M., Russell, A. F., McIlrath, G. M., Speakman, J. R., & Clutton-Brock, T. H. (2002). The energetics of lactation in cooperatively breeding meerkats *Suricata suricatta*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*(1505), 2147-2153.
- Schneirla, T. C. (1950). The relationship between observation and experimentation in the field study of behavior. *Annals of the New York Academy of Sciences*, *51*(6), 1022-1044.
- Schoeller, D. A., & Van Santen, E. (1982). Measurement of energy expenditure in humans by doubly labeled water method. *Journal of applied physiology*, *53*(4), 955-959.
- Scholander, P. F. (1940). *Experimental investigations on the respiratory function in diving mammals and birds* (No. 22). I kommisjon hos Jacob Dybwad.
- Sih, A., Ferrari, M. C., & Harris, D. J. (2011). Evolution and behavioural responses to human-induced rapid environmental change. *Evolutionary applications*, *4*(2), 367-387.

- Simpson, S. J., Raubenheimer, D., Charleston, M. A., & Clissold, F. J. (2010). Modelling nutritional interactions: from individuals to communities. *Trends in Ecology & Evolution*, 25(1), 53-60.
- Sims, D. W., Witt, M. J., Richardson, A. J., Southall, E. J., & Metcalfe, J. D. (2006). Encounter success of free-ranging marine predator movements across a dynamic prey landscape. *Proceedings of the Royal Society B: Biological Sciences*, 273(1591), 1195-1201.
- Sims, D. W., Southall, E. J., Humphries, N. E., Hays, G. C., Bradshaw, C. J., Pitchford, J. W., ... & Morritt, D. (2008). Scaling laws of marine predator search behaviour. *Nature*, 451(7182), 1098-1102.
- Smircich, M. G., & Kelly, J. T. (2014). Extending the 2% rule: the effects of heavy internal tags on stress physiology, swimming performance, and growth in brook trout. *Animal Biotelemetry*, 2(1), 16.
- Stearns, S. C. (1992). *The evolution of life histories* (No. 575 S81).
- Strandburg-Peshkin, A., Farine, D. R., Couzin, I. D., & Crofoot, M. C. (2015). Shared decision-making drives collective movement in wild baboons. *Science*, 348(6241), 1358-1361.
- Sutherland, W. J. (1998). The importance of behavioural studies in conservation biology. *Animal behaviour*, 56(4), 801-809.
- Thomson, J. A., & Heithaus, M. R. (2014). Animal-borne video reveals seasonal activity patterns of green sea turtles and the importance of accounting for capture stress in short-term biologging. *Journal of experimental marine biology and ecology*, 450, 15-20.
- Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für tierpsychologie*, 20(4), 410-433.
- Townsend, C. R., & Calow, P. (1981). *Physiological ecology: an evolutionary approach to resource use* (No. Sirsi) i9780878938278).
- Townsend, S. W., Zöttl, M., & Manser, M. B. (2011). All clear? Meerkats attend to contextual information in close calls to coordinate vigilance. *Behavioral Ecology and Sociobiology*, 65(10), 1927-1934.

- Tuomainen, U., & Candolin, U. (2011). Behavioural responses to human-induced environmental change. *Biological Reviews*, *86*(3), 640-657.
- Turner, L. W., Udal, M. C., Larson, B. T., & Shearer, S. A. (2000). Monitoring cattle behavior and pasture use with GPS and GIS. *Canadian Journal of Animal Science*, *80*(3), 405-413.
- Ungar, E. D., Henkin, Z., Gutman, M., Dolev, A., Genizi, A., & Ganskopp, D. (2005). Inference of animal activity from GPS collar data on free-ranging cattle. *Rangeland Ecology & Management*, *58*(3), 256-266.
- Urbano, F., Cagnacci, F., Calenge, C., Dettki, H., Cameron, A., & Neteler, M. (2010). Wildlife tracking data management: a new vision. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1550), 2177-2185.
- Vandenabeele, S. P., Wilson, R. P., & Grogan, A. (2011). Tags on seabirds: how seriously are instrument-induced behaviours considered?. *Animal Welfare-The UFAW Journal*, *20*(4), 559.
- Vandenabeele, S. P., Shepard, E. L., Grogan, A., & Wilson, R. P. (2012). When three per cent may not be three per cent; device-equipped seabirds experience variable flight constraints. *Marine Biology*, *159*(1), 1-14.
- Volpov, B. L., Hoskins, A. J., Battaile, B. C., Viviant, M., Wheatley, K. E., Marshall, G., ... & Arnould, J. P. (2015). Identification of prey captures in Australian fur seals (*Arctocephalus pusillus doriferus*) using head-mounted accelerometers: field validation with animal-borne video cameras. *PLoS One*, *10*(6).
- Walker, K. A., Trites, A. W., Haulena, M., & Weary, D. M. (2012). A review of the effects of different marking and tagging techniques on marine mammals. *Wildlife Research*, *39*(1), 15-30.
- Walker, J. S., Jones, M. W., Laramée, R. S., Holton, M. D., Shepard, E. L., Williams, H. J., ... & Bidder, O. R. (2015). Prying into the intimate secrets of animal lives; software beyond hardware for comprehensive annotation in 'Daily Diary' tags. *Movement ecology*, *3*(1), 29.

- Wang, Y., Smith, J. A., & Wilmers, C. C. (2017). Residential development alters behavior, movement, and energetics in an apex predator, the puma. *PloS one*, *12*(10).
- Watanabe, Y. Y., & Takahashi, A. (2013). Linking animal-borne video to accelerometers reveals prey capture variability. *Proceedings of the National Academy of Sciences*, *110*(6), 2199-2204.
- Weathers, W. W., Buttemer, W. A., Hayworth, A. M., & Nagy, K. A. (1984). An evaluation of time-budget estimates of daily energy expenditure in birds. *The Auk*, *101*(3), 459-472.
- Webber, D. M., McKinnon, G. P., & Claireaux, G. (2001). Evaluating differential pressure in the european sea bass (*Dicentrarchus labrax*) as a telemetered index of swimming speed. In *Electronic tagging and tracking in marine fisheries* (pp. 297-313). Springer, Dordrecht.
- Weimerskirch, H., Salamolard, M., & Jouventin, P. (1992). Satellite telemetry of foraging movements in the wandering albatross. *Wildlife Telemetry. Remote Monitoring and Tracking of Animals*, 185-198.
- Weimerskirch, H., Salamolard, M., Sarrazin, F., & Jouventin, P. (1993). Foraging strategy of wandering albatrosses through the breeding season: a study using satellite telemetry. *The Auk*, *110*(2), 325-342.
- Weimerskirch, H., Wilson, R. P., & Lys, P. (1997). Activity pattern of foraging in the wandering albatross: a marine predator with two modes of prey searching. *Marine Ecology Progress Series*, *151*, 245-254.
- Weimerskirch, H., Pinaud, D., Pawlowski, F., & Bost, C. A. (2007). Does prey capture induce area-restricted search? A fine-scale study using GPS in a marine predator, the wandering albatross. *The American Naturalist*, *170*(5), 734-743.
- Williams, H. J., Shepard, E. L. C., Duriez, O., & Lambertucci, S. A. (2015). Can accelerometry be used to distinguish between flight types in soaring birds?. *Animal Biotelemetry*, *3*(1), 45.

- Williams, H. J., Holton, M. D., Shepard, E. L., Largey, N., Norman, B., Ryan, P. G., ... & Marks, N. J. (2017). Identification of animal movement patterns using triaxial magnetometry. *Movement ecology*, 5(1), 6.
- Williams, H. J., Taylor, L. A., Benhamou, S., Bijleveld, A. I., Clay, T. A., de Grissac, S., ... & Griffiths, R. C. (2019). Optimizing the use of biologgers for movement ecology research. *Journal of Animal Ecology*.
- Wilson, R. P., & Bain, C. A. (1984). An inexpensive speed meter for penguins at sea. *The Journal of wildlife management*, 1360-1364.
- Wilson, R. P., & Wilson, M. P. (1988). Dead reckoning: a new technique for determining penguin movements at sea. *Meeresforschung (Hamburg)*, 32(2), 155-158.
- Wilson, R. P., Wilson, M. P. T., Link, R., Mempel, H., & Adams, N. J. (1991). Determination of movements of African penguins *Spheniscus demersus* using a compass system: dead reckoning may be an alternative to telemetry. *Journal of Experimental Biology*, 157(1), 557-564.
- Wilson, R. P., & Culik, B. M. (1993). Activity-specific metabolic rates from double labeled water studies: are activity costs underestimated?. *Ecology*, 74(4), 1285-1287.
- Wilson, R. P., Pütz, K., Grémillet, D., Culik, B. M., Kierspel, M., Regel, J., ... & Cooper, J. (1995). Reliability of stomach temperature changes in determining feeding characteristics of seabirds. *Journal of Experimental Biology*, 198(5), 1115-1135.
- Wilson, R. P., Grémillet, D., Syder, J., Kierspel, M. A., Garthe, S., Weimerskirch, H., ... & Nel, D. (2002). Remote-sensing systems and seabirds: their use, abuse and potential for measuring marine environmental variables. *Marine Ecology Progress Series*, 228, 241-261.
- Wilson, R. P., White, C. R., Quintana, F., Halsey, L. G., Liebsch, N., Martin, G. R., & Butler, P. J. (2006). Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *Journal of Animal Ecology*, 75(5), 1081-1090.

- Wilson, R. P., Shepard, E. L. C., & Liebsch, N. (2008). Prying into the intimate details of animal lives: use of a daily diary on animals. *Endangered species research*, 4(1-2), 123-137.
- Wilson, R. P., Quintana, F., & Hobson, V. J. (2012). Construction of energy landscapes can clarify the movement and distribution of foraging animals. *Proceedings of the Royal Society B: Biological Sciences*, 279(1730), 975-980.
- Winter, J. D. (1996). Advances in underwater biotelemetry. *Fisheries techniques*, 555-590.
- Wong, T. C., Webster, J. G., Montoye, H. J., & Washburn, R. (1981). Portable accelerometer device for measuring human energy expenditure. *IEEE Transactions on Biomedical Engineering*, (6), 467-471.
- Wright, S., Metcalfe, J. D., Hetherington, S., & Wilson, R. (2014). Estimating activity-specific energy expenditure in a teleost fish, using accelerometer loggers. *Marine Ecology Progress Series*, 496, 19-32.
- Yates, C. A., Erban, R., Escudero, C., Couzin, I. D., Buhl, J., Kevrekidis, I. G., ... & Sumpter, D. J. (2009). Inherent noise can facilitate coherence in collective swarm motion. *Proceedings of the National Academy of Sciences*, 106(14), 5464-5469.
- Ydenberg, R. C., Welham, C. V. J., Schmid-Hempel, R., Schmid-Hempel, P., & Beauchamp, G. (1994). Time and energy constraints and the relationships between currencies in foraging theory. *Behavioral Ecology*, 5(1), 28-34.
- Yoda, K., Sato, K., Niizuma, Y., Kurita, M., Bost, C., Le Maho, Y., & Naito, Y. (1999). Precise monitoring of porpoising behaviour of Adélie penguins determined using acceleration data loggers. *Journal of Experimental Biology*, 202(22), 3121-3126

CHAPTER **2**

**State of the Art in
Animal Behaviour
Recognition, and
Theoretical Analysis of
Sensor Data
Interpretation**

Summary

Inertial sensors such as accelerometers and magnetometers are used to quantify animal behaviour due to their sensitivity to body movement. When the animal is stationary, the signal doesn't change with time, and when the animal moves, the signal starts varying with time. The static signal, despite remaining flat, contains information on sensor tilt with respect to a global vector field: Earth's gravity in the case of accelerometers, and Earth's magnetic field in the case of magnetometers. These simple principles, in tandem with a staggering suite of powerful machine learning techniques, have been leveraged to recognise common animal behaviours such as locomotion, resting, and feeding/foraging. And yet there is room for new applications, such as recognition of finer behaviours, for which it may be necessary to re-evaluate how inertial sensor data are interpreted. This chapter shall first review existing literature on methods employed to infer behaviour from inertial sensor data, then provide a theoretical analysis of the limits of some of the key assumptions involved in data interpretation, and then pave the way for subsequent chapters.

2.1 | State of the art²

2.1.1 | The use of accelerometers to infer animal behaviour

Accelerometers are not sentient (as far as one can tell), and must “re”cognise based on a human observer’s cognition. Remote recognition of behaviour using accelerometers, therefore, requires one or the other form of “groundtruth” data, or reference knowledge, which is based on human observation. Practically speaking, the process of groundtruthing typically involves time-synchronising the acceleration signals with simultaneously recorded video, and then having an expert (who is familiar with the behaviour of the animal being studied) annotate the video. This links the recorded acceleration signal to the stream of observed animal behaviours that produced it.

The task of inferring behaviour from recorded acceleration is facilitated by calibrating the accelerometer such that data are in units of m/s^2 or g (acceleration due to Earth’s gravity, approximately equal to 9.81 m/s^2) (and not, say, in units of milli-Volts). There are a number of methods used for calibrating accelerometers, and virtually all of them use g as a reference. Consequently, the general principle behind such methods is similar: that each axis of the accelerometer must measure a magnitude of acceleration equal to g when aligned in the vertical direction (Parvis, Ferraris & Grimaldi 1995), or that static acceleration values in different orientations of the sensor must lie on a sphere centred at the origin with radius g (Fong et al. 2008), or that gently tumbling the sensor through different orientations will generate points on a sphere centred at the origin with radius g (Wang et al. 2015 adapted a method developed for magnetometers – Gebre-Egziabher et al. 2006 – for application on accelerometers).

All behaviour is performed for a finite duration, following which the animal transitions to a different behaviour. The animal may be static for a while (e.g. resting), then begin foraging (involving displacement, and hence movement), and then, perhaps perceiving predatory presence, run (involving vigorous, periodic

² The intended meaning of this term is closer to the French *‘état de l’art’* and the way this term was reported to be used for the first time in English in 1910 (Wikipedia article on ‘State of the art’): ‘the state of current knowledge in the domain’. This is different from its meaning in contemporary English: ‘latest/most recent’.

motion) toward safety. Accelerometers, being sensitive to movement and the lack of it, can, in theory, detect when these different behaviours are performed. Thus, sections of the accelerometer signal must somehow be able to quantify ‘static’-ness, or ‘medium-intensity dynamicity’, or ‘vigorous periodicity’. This is commonly done by chopping up acceleration signals into finite sections, called windows, and then engineering quantities called “features” that summarise the characteristics of the acceleration signal in the window. A wide range of features have been derived from running windows of acceleration data of pre-set size, including: (a) statistical summaries such as mean, standard deviation, skewness, kurtosis (e.g. Nathan et al. 2012; Resheff et al. 2014), (b) frequency-domain features based on the Fourier transform such as dominant frequency, peak amplitude, energy, entropy (e.g. Laich et al. 2008; Marais et al. 2014; Fehlmann et al. 2017), (c) wavelet-domain features derived after continuous wavelet transformation of the acceleration signal (Sakamoto et al. 2009), (d) biomechanically significant features such as static acceleration (which is equivalent to the mean of the signal) (e.g. Fehlmann et al. 2017), ‘dynamic body acceleration’ (obtained by subtracting static acceleration from total acceleration according to Wilson et al. 2006), body segment inclination obtained from sensor tilt (e.g. Laich et al. 2008; Alvarenga et al. 2016), and (e) miscellaneous features such as pairwise correlation between axes (e.g. Marais et al. 2014; Nathan et al. 2012), signal magnitude area (e.g. Gao et al. 2013; Alvarenga et al. 2016). This list is representative, not exhaustive, and the diversity of features continues to grow. Feature engineering is sometimes followed by a “feature selection” step to reduce the number of candidate features to a few that discriminate best between behavioural classes. Examples of feature selection/reduction techniques include: principal component analysis (e.g. Soltis et al. 2012), embedded methods (e.g. Kamminga et al. 2018), feature relative importance using Random Forest (e.g. Alvarenga et al. 2016), and custom greedy selection (Marais et al. 2014).

There is considerable variation in the techniques employed to infer behaviour from acceleration data, and methods adopted by some studies do not require or perform all of the steps described so far. Till date, there have been five main methods to infer behaviour from acceleration data (Figure 2.1). Early studies recognised behaviours of interest from raw acceleration data visually: this was either based on expert knowledge of the animal species (e.g. Yoda et al. 1999; Yoda et al. 2001) or

on visual comparison of signals with those recorded during an initial groundtruthing phase (Ropert-Coudert et al. 2004). Common trends in acceleration patterns observed and analysed visually across species were identified in Shepard et al. 2008a. The first step towards automated identification of behaviour, to my knowledge, was taken through the use of empirically determined thresholds to separate high and low activity counts in dairy cows (Müller & Schrader 2003); the use of empirical thresholds was soon extended to recognise multiple behaviours in other studies, e.g. in terrestrial tortoises (*Testudo graeca*) (Lagarde et al. 2008) and imperial cormorants (*Phalacrocorax atriceps*) (Laich et al. 2008). The logistical difficulty of the process of groundtruthing, where recorded acceleration signals must be associated to time-synchronised videos annotated by an expert, was recognised early on, and there was consequently an interest in unsupervised learning approaches such as k-means clustering that does not require groundtruthed data (Sakamoto et al. 2009), although ‘k’ and the physical interpretation of the clusters (i.e. which behaviour each cluster corresponds to) found by this method must be defined and interpreted, respectively, by the user. This approach, based on computation of wavelet-based metrics, has since been used in several studies (e.g. Sato et al. 2009; Brownscombe et al. 2014; Broell et al. 2016). It has been reported, however, that since all the features used by this method are derived from dynamic acceleration, certain basic behaviours differing mainly in posture are missed (Sakamoto et al. 2009; Nathan et al. 2012). The interest in such approaches continues (e.g. k-means clustering in Shahriar et al. 2016; hidden Markov models in Leos-Barajas et al. 2017), and an accurate and robust unsupervised approach would be of great value to the field.

A semi-automated technique that has sought to reduce the difficulty of collecting groundtruthed data is user-vetted template matching, where a matching algorithm takes as input one example pattern of a given behaviour, then identifies candidate matches throughout a given acceleration signal, with the user finally accepting or rejecting the match (Walker et al. 2015). Another semi-automated technique has sought to identify behavioural elements in composite animal behaviours: here, manually chosen thresholds based on parameters describing the shape of the acceleration signal are used to separate behaviours using a Boolean decision tree (Wilson et al. 2018).

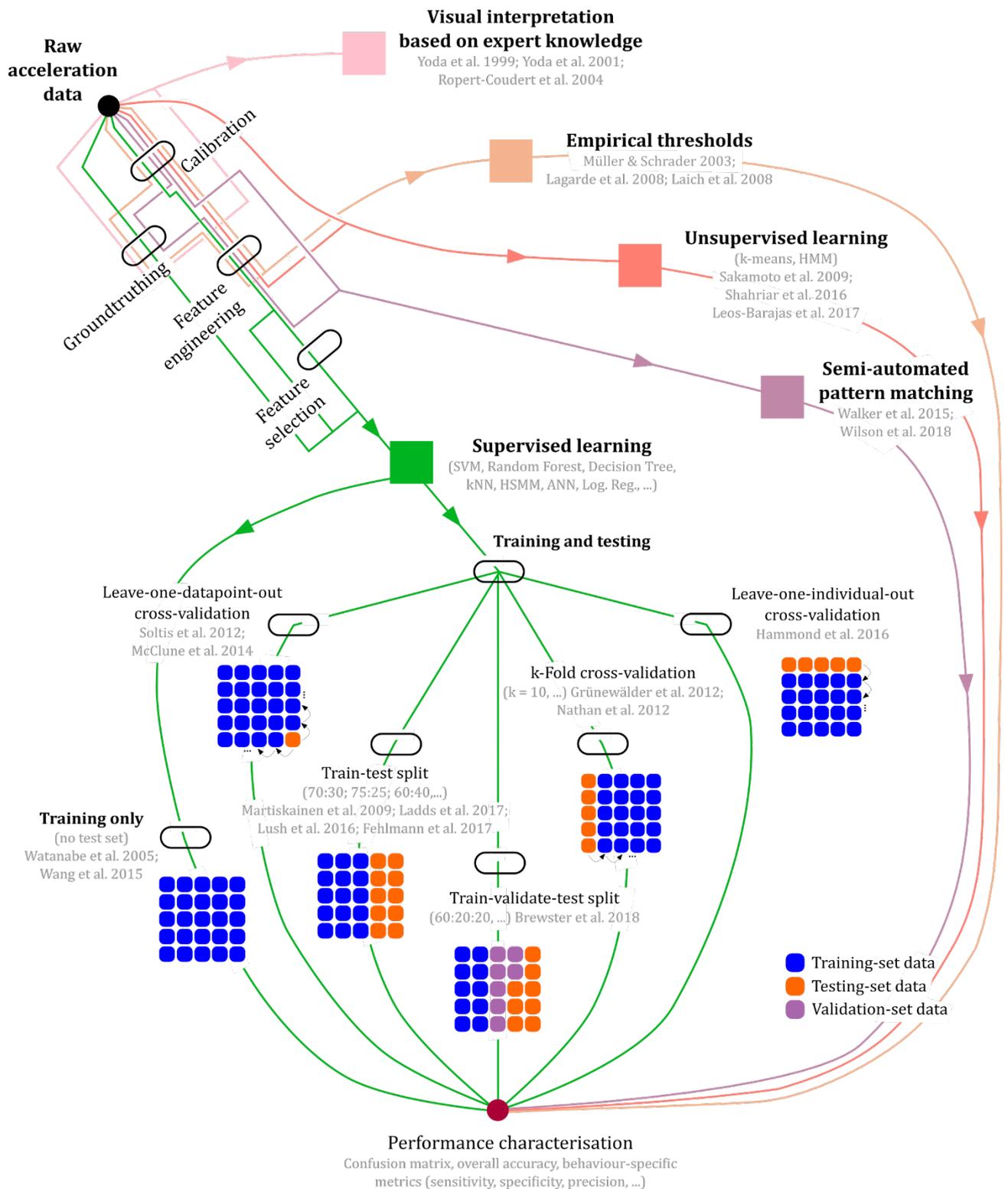


Figure 2.1. **Metro map of animal behaviour recognition techniques.** Paths emanate from “raw acceleration data”. Of the five techniques that have been used, supervised learning is the most common and diverse. References shown here are not exhaustive. SVM: Support Vector Machine; kNN: k-Nearest Neighbours; ANN: Artificial Neural Network; H(S)MM: Hidden (Semi) Markov Model; Log. Reg.: Logistic Regression.

The need for validated behavioural information on the one hand, and that for complete automation of behaviour recognition in light of the high amounts of data collected today on the other, have pushed the majority of recent studies to adopt supervised machine learning approaches. This has been facilitated by the widespread availability of user-friendly implementations of robust machine learning algorithms in open-source (Pedregosa et al. 2011; Bischl et al. 2016) and commercial (e.g. MATLAB) software. This approach involves two additional steps post feature selection (Figure):

(i) *Choosing one or several machine learning models to test*

This step involves the application of one or several available supervised machine learning algorithms such as Support Vector Machine (SVM) (e.g. Martiskainen et al. 2009; Carroll et al. 2014), Random Forest (e.g. Nathan et al. 2012; Pagano et al. 2017), Decision Tree (e.g. McClune et al. 2014; Nishizawa et al. 2013), Logistic Regression (e.g. Robert et al. 2009; Ladds et al. 2016), and Artificial Neural Networks (Brewster et al. 2018; Nathan et al. 2012). Three notable variants are: (a) hidden semi-Markov models (HSMM) (e.g. Hammond et al. 2016), where information on transitions between and durations of individual behavioural states is also incorporated, and (b) k-nearest neighbours (KNN), where acceleration data are directly given as input to the KNN algorithm without computing features (e.g. Bidder et al. 2014; Sur et al. 2017), although KNN can also be performed post feature computation. A description of common machine learning algorithms employed for animal behaviour recognition has been given in Nathan et al. 2012.

(ii) *Cross-validation*

A wide variety of validation approaches have been used to quantify the accuracy of behaviour recognition. A few studies trained their algorithm and tested its performance on the same data (Watanabe et al. 2005; Wang et al. 2015); this method may be used to assess the separability of behavioural classes, but may not be best-suited for prediction on unseen data since it yields overoptimistic results (Larson

1931; Arlot & Celisse 2010).

Typically, however, the dataset is split into two parts: a training set, which is used to train the model, and a testing set, which is withheld from the training process and is used to test model performance on unseen data. When this split is carried out only once, the process is referred to as ‘hold-out’ or ‘validation’, and when results are averaged over several splits, the process is called ‘cross-validation’ (CV) (Arlot & Celisse 2010). Several studies have used validation splits such as 70:30 (70% data used for training, 30% for testing), 75:25, and 60:40 (e.g. Martiskainen et al. 2009; Ladds et al. 2017; Lush et al. 2016; Fehlmann et al. 2017). Three kinds of CV techniques have been used (Figure): leave-one-datapoint-out (e.g. McClune et al. 2014; Soltis et al. 2012; Grünewälder et al. 2012), k-fold (with $k = 10$, typically) (e.g. Nathan et al. 2012; Kamminga et al. 2018), and leave-one-individual-out (Hammond et al. 2016).

In leave-one-datapoint-out CV, each datapoint is successively left out from training data and is used for testing (Arlot & Celisse 2010). In k-fold CV, the dataset is partitioned into k subsamples, and the validation process is carried out until each subsample has been successively left out and used as a testing set (Geisser 1974; Arlot & Celisse 2010). In leave-one-individual-out (LOIO) CV, training is performed using data from all individuals but one, and the left out individual’s data is used as the testing set; the process is carried out until each individual’s data has been the testing set exactly once.

The dataset can also be divided into three parts: one for training, one for validation, and one for testing. Here, after fitting the model using training data, two types of errors are estimated: prediction error using the validation set to select among a given set of candidate models, and generalisation error using the testing set to assess the performance of the final chosen model (Chapter 7.2 in Hastie, Tibshirani & Friedman 2009). To my knowledge, this has been done only in one recent study on lemon sharks (*Negaprion brevirostris*) (Brewster et al. 2018).

In all cross-validation techniques except LOIO, both training and testing sets contain datapoints extracted from a single continuous

recording on the same individual. This violates these methods' assumption (Arlot & Celisse 2010) that these datapoints are independent and identically distributed since they are extracted from the same time series. The effects of non-independence of data have been shown to be mitigated through LOIO CV in human neuroimaging studies (Esterman et al. 2010). LOIO CV is likely the most stringent of these cross-validation techniques, since the trained model must deal with an entirely independent time series with possibly different tag orientation and the movement-related idiosyncrasies of a hitherto unseen animal, and has, to my knowledge been previously performed in only one study (Hammond et al. 2016).

2.1.2 | The use of magnetometers to infer animal behaviour

Magnetometers are built based on one of several sensing mechanisms, e.g. the most commonly used Hall effect (Popovic 2014), fluxgate (Primdahl 1979), and magnetic tunnelling junctions (Parkin et al. 2004); see Lenz 1990 for several other examples of sensing mechanisms. The role of magnetometers in animal studies has largely been restricted to obtaining heading (e.g. Wilson & Wilson 1988; Wilson et al. 1991; Mitani et al. 2003), and no study has reported automatic inference of common animal behaviours from magnetometer data. Like accelerometers, magnetometers also need to be calibrated; however, since the magnetic field of the Earth changes significantly with location (much more than the change in gravity with location), and magnetometers are sensitive to hard and soft iron errors (Caruso 2000), calibration methods seek to normalise the output of a magnetometer such that all datapoints lie on a sphere of radius 1 (e.g. Gebre-Egziabher et al. 2001; Bonnet et al. 2009; Williams et al. 2017). In addition, however, magnetometers have been reported to be sensitive to dynamic movement as well: cheetah (*Acinonyx jubatus*) running behaviour has been reported to produce large oscillations in magnetometer readings (Wilson, Shepard & Liebsch 2008). Recently, it was shown that even relatively slow dynamic movements, such as soaring in birds, can be distinguished using magnetometers (Williams et al. 2015; Williams et al. 2017). Further, human studies have shown that it might be possible to use metrics based on magnetometer-

derived angular velocity for wearable applications (Kunze et al. 2010) such as the measurement of walking and running cadence (Abdulla et al. 2013).

2.1.3 | Conclusion

The use of accelerometers has by far outstripped that of magnetometers for animal behaviour recognition, and a large suite of methodology is now available for inferring animal behaviours from accelerometer data. However, the facts that accelerometers and magnetometers both measure static and dynamic components of movement, and that accelerometers have met with enormous success in recognising animal behaviour, suggests that magnetometers, too, could be capable of recognising the common behaviours that accelerometers can: resting, locomotion, and feeding/foraging (Brown et al. 2013). Further, since magnetometers are based on a different sensing mechanism compared to accelerometers, some complementary benefits for recognition of even common behaviours (apart from those already observed for specific behaviours reported in Williams et al. 2017) might be revealed in a one-to-one comparison of behaviour recognition using accelerometers, and that using magnetometers.

Both behaviour recognition and energy expenditure estimation (Chapter 1) from acceleration data rely on running mean-based gravity compensation to separate static and dynamic components of movement. However, whether running means actually isolate gravity or not, and whether gravity-compensated dynamic acceleration actually reflects animal body acceleration given that sensors will inevitably wobble about their location of attachment, has not yet been critically analysed in animal studies.

Secondly, the ability and limits of the magnetometer's potential to characterise dynamic movement have not been explicitly formulated yet. If the local magnetic field intensity vector does not vary with time and remains uniform in the space where measurements are performed, the magnetometer will be insensitive to translation, and the only variation in time will be due to sensor rotation. Thus, variation of magnetometer signals with time is due directly to angular velocity, which motivated previous attempts to derive angular velocity from magnetometer data (Kunze et al. 2010). However, when the axis of rotation becomes parallel to the

local magnetic field vector, the magnetometer will not be able to record angular velocity, but there is no relationship, to my knowledge, that quantifies this dependency on the angle between the rotation axis and magnetic field vector.

The next section provides theoretical results that address these two issues.

2.2 | Theoretical analysis

2.2.1 | On the separation of static and dynamic acceleration

Accelerometers are sensitive to both gravitational and nongravitational acceleration (IEEE Std 528–2019), commonly assumed to be equal to static and ‘dynamic body acceleration’ in animal accelerometry literature, respectively. The two measurements cannot be recorded separately, and both are present in the same recording signal. This is not problematic when the behaviours of interest are simply active or inactive (e.g. Van Oort et al. 2004; Whitney et al. 2007; Gervasi, Brunberg & Swenson 2006), where only nongravitational acceleration can separate the two behavioural classes. However, separating the two acceleration components becomes important when two or more dynamic behaviours need to be discriminated, such as prey capture from swimming in marine predators (e.g. Suzuki et al. 2009; Volpov et al. 2015), and for estimating energy expenditure (Wilson et al. 2006). The mean of the acceleration signal in a running window (e.g. of size 1, 2, 3 seconds, etc.) is assumed to quantify static acceleration resulting from body angle with respect to gravity (e.g. Yoda et al. 2001; Wilson et al. 2006). The robustness of static acceleration computation to the choice of window size has been experimentally investigated for steady, constant-speed locomotion as a model behaviour (Shepard et al. 2008b). The subtraction of the mean from total acceleration, used to derive Overall Dynamic Body Acceleration (ODBA) which has been found to correlate with energy expenditure in a number of species (Wilson et al. 2006), has been said to result from movement of the animal’s body parts (Shepard et al. 2008b). The precise definition has been given to be: ‘*A thorough description of ODBA states that it is a measure of dynamic acceleration induced about the centre of an animal’s mass as a result of the movement of body parts, being derived from recordings of acceleration*

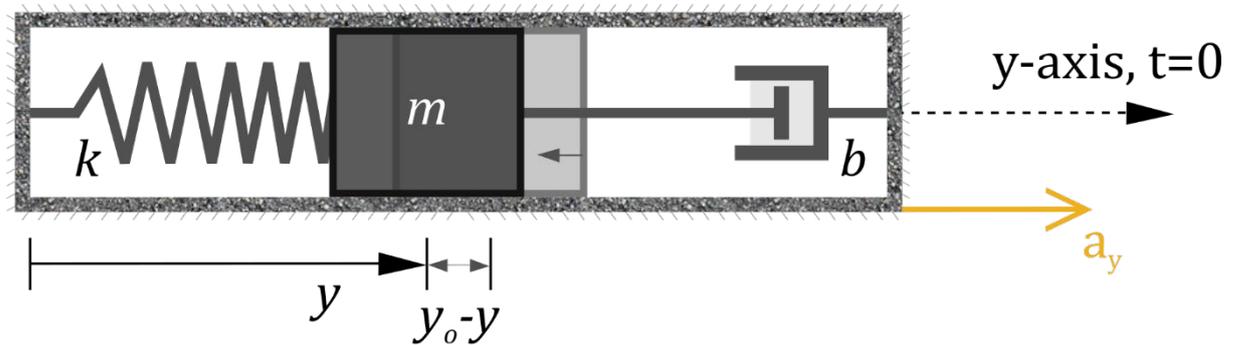
in the three spatial dimensions by a data logger placed on a fixed point of an animal.' (Halsey, Shepard & Wilson 2011).

This definition, however, does not take into account movement of loosely attached sensors with respect to the animal or of the animal's soft tissues with respect to joint motions at the location of tag attachment. The former problem was noticed in an early study on primates, where a loose collar bouncing against the animal's body was thought to introduce high-frequency artefacts in the acceleration signal (Sellers & Crompton 2004). The latter problem has received much attention in human studies involving accelerometers (e.g. Forner-Cordero et al. 2008). Experimental measurement of human soft tissue artefacts is known to require clusters of several markers (Barré et al. 2015), which would be impossible to deploy on wild animals. In this section, a theoretical model of the accelerometer is used to derive analytical expressions for static acceleration under conditions of possibly accelerating locomotion, and the impact of sensor wobbling on measured acceleration.

2.2.1.1 | Computing external acceleration from spring-mass-damper system dynamics

In this section, using the spring-mass-damper system (Figure 2.2) as a model for the commonly used capacitive accelerometer (Béliveau et al. 1999; Suzuki & Tai 2006; Button 2015), an equation relating external acceleration to the resulting displacement of the accelerometer's proof mass is derived similar to Aminian 2019. Consider the uniaxial accelerometer shown in Figure 2.2A.

(A)



(B)



Figure 2.2. **The spring-mass-damper model for a uniaxial capacitive accelerometer.** (A) A horizontal acceleration a_y acts on the accelerometer casing at time $t = 0$, causing a leftwards displacement of the proof mass from its position at y_0 (where the spring with stiffness k is undeformed) to a new position y . (B) Free-body diagram of the proof mass: forces due to the compression of the spring and resistance from the damper (with proportionality constant b) oppose the pseudo-force generated by a_y .

The proof mass is constrained to move along the accelerometer's sensitive axis (here, the y -axis), and it can do so without friction. For time $t < 0$, the proof mass is at rest in the horizontal plane with no forces acting on it in the y -direction. The spring has a stiffness constant of k and natural length y_0 , and the damper's proportionality constant is b . At $t = 0$, an external acceleration a_y starts acting on the accelerometer casing towards the right, displacing the proof mass to the left with velocity $\dot{y} = \frac{dy}{dt}$. With $\ddot{y} = \frac{d^2y}{dt^2}$, one finds the following equation from the free-body diagram in Figure 2.2B:

$$m\ddot{y} = -ma_y - b\dot{y} - k(y - y_0)$$

from where we obtain the desired relationship:

$$a_y = -\ddot{y} - \frac{b}{m}\dot{y} - \frac{k}{m}(y - y_o) \quad (2.1)$$

One can use equation (2.1) to find an expression for a_y in more complex situations involving more forces, where y can be found using Newton's laws but not a_y .

2.2.1.2 | Accelerometer response under the influence of time-varying horizontal acceleration and sensor wobbling

Consider a triaxial accelerometer composed of three uniaxial accelerometers (one along each of the three mutually perpendicular axes), attached to an animal accelerating during a locomotion bout. The ‘true’ dynamic body acceleration vector with respect to the orientation given by the animal’s anatomical frame of reference is denoted by: $\vec{A}_{AF} = A_{x,AF}\hat{l}_{AF} + A_{y,AF}\hat{j}_{AF} + A_{z,AF}\hat{k}_{AF}$, where \hat{l}_{AF} , \hat{j}_{AF} , \hat{k}_{AF} refer to the x (heave axis: ventral-to-dorsal), y (surge axis: tail-to-head) and z (sway axis: medial-to-lateral) axes of the anatomical frame, respectively. The subscript AF refers to “anatomical frame”. We assume that the orientation of AF does not change with respect to the global frame during the locomotion bout under consideration. The components of \vec{A}_{AF} are given by:

$$\begin{aligned} A_{x,AF} &= \sum_{n=1}^N (A_{1,n} \cos \omega nt + B_{1,n} \sin \omega nt) \\ A_{y,AF} &= A_0 + \sum_{n=1}^N (A_{2,n} \cos \omega nt + B_{2,n} \sin \omega nt) \\ A_{z,AF} &= \sum_{n=1}^N (A_{1,n} \cos \omega nt + B_{1,n} \sin \omega nt) \end{aligned} \quad (2.2)$$

where the acceleration signal has a period of $2\pi/\omega$ seconds, with ω denoting the angular frequency, and is described with sufficient accuracy using N harmonics. Note also that the constant A_0 implies a net acceleration, and thus displacement (i.e. after integrating $A_{y,AF}$ twice over a full cycle of length $2\pi/\omega$), in the positive surge direction; the other axes do not have a non-zero constant term to indicate that there is

no net displacement in the heave or sway axes. The gravity vector \vec{g} , with magnitude $g = 9.81 \text{ m/s}^2$, makes angles of $\alpha_1, \alpha_2, \alpha_3$ with $\hat{i}_{AF}, \hat{j}_{AF}, \hat{k}_{AF}$, respectively such that, for instance, $g \cos \alpha_2$ is the component of \vec{g} aligned along the surge axis. We assume that at time $t = 0$, the accelerometer's axes $\hat{i}_{SF}, \hat{j}_{SF}, \hat{k}_{SF}$ (x, y and z, respectively; the subscript *SF* denotes 'sensor frame') are perfectly aligned with $\hat{i}_{AF}, \hat{j}_{AF}, \hat{k}_{AF}$, respectively. The sensor is loosely attached, and I model its wobbling as planar rotation about the sway axis (\hat{k}_{AF}) such that the angle ϕ between \hat{j}_{SF} and \hat{j}_{AF} varies with time as: $\phi = \phi_o \sin \omega t$, where ϕ_o is a positive constant in radians corresponding to the maximum value of ϕ (Figure 2.3A). Since the animal's acceleration produces this wobbling, it is reasonable to assume that sensor rotation will occur with the same frequency as that of locomotion, ω .

Given this set-up, we derive an equation for the acceleration measured by the uniaxial accelerometer aligned along \hat{j}_{SF} , i.e. the sensor's y-axis, using the free-body diagram in Figure 2.3B.

$$\begin{aligned}
m\ddot{y} = & -k(y - y_o) - b\dot{y} + m\phi_o^2\omega^2 \cos^2 \omega t \cdot y \\
& - mA_{y,AF} \cos(\phi_o \sin \omega t) \\
& - mA_{x,AF} \sin(\phi_o \sin \omega t) \\
& - mg \cos \alpha_2 \cos(\phi_o \sin \omega t) \\
& - mg \cos \alpha_1 \sin(\phi_o \sin \omega t)
\end{aligned} \tag{2.3}$$

where $(m\phi_o^2\omega^2 \cos^2 \omega t \cdot y)$ is the centrifugal force experienced by the proof mass, the terms involving g are due to gravitational force, and terms involving A are due to the animal's body acceleration. Equation (2.3) does not have an analytical solution with a finite number of terms when ϕ_o is large (i.e. comparable to 1). However, for small ϕ_o , where $\phi_o \ll 1$ radian such that $\sin(\phi_o \sin \omega t) \approx \phi_o \sin \omega t$, and $\cos(\phi_o \sin \omega t) \approx 1$ (first-order Taylor series approximation), the above equation can be simplified to:

$$\begin{aligned}
m\ddot{y} = & -k(y - y_o) - b\dot{y} \\
& - mA_{y,AF} - mA_{x,AF}\phi_o \sin \omega t - mg \cos \alpha_2 \\
& - mg \cos \alpha_1 \phi_o \sin \omega t
\end{aligned}$$

from which, upon substitution of equations (2.1) and (2.2), we obtain an expression for the measured acceleration a_y as:

$$\begin{aligned}
a_y &= g \cos \alpha_2 + g \cos \alpha_1 \phi_o \sin \omega t \\
&+ \left(A_0 + \sum_{n=1}^N (A_{2,n} \cos \omega n t + B_{2,n} \sin \omega n t) \right) \\
&+ \phi_o \sin \omega t \left(\sum_{n=1}^N (A_{1,n} \cos \omega n t + B_{1,n} \sin \omega n t) \right)
\end{aligned}$$

which can be resolved into:

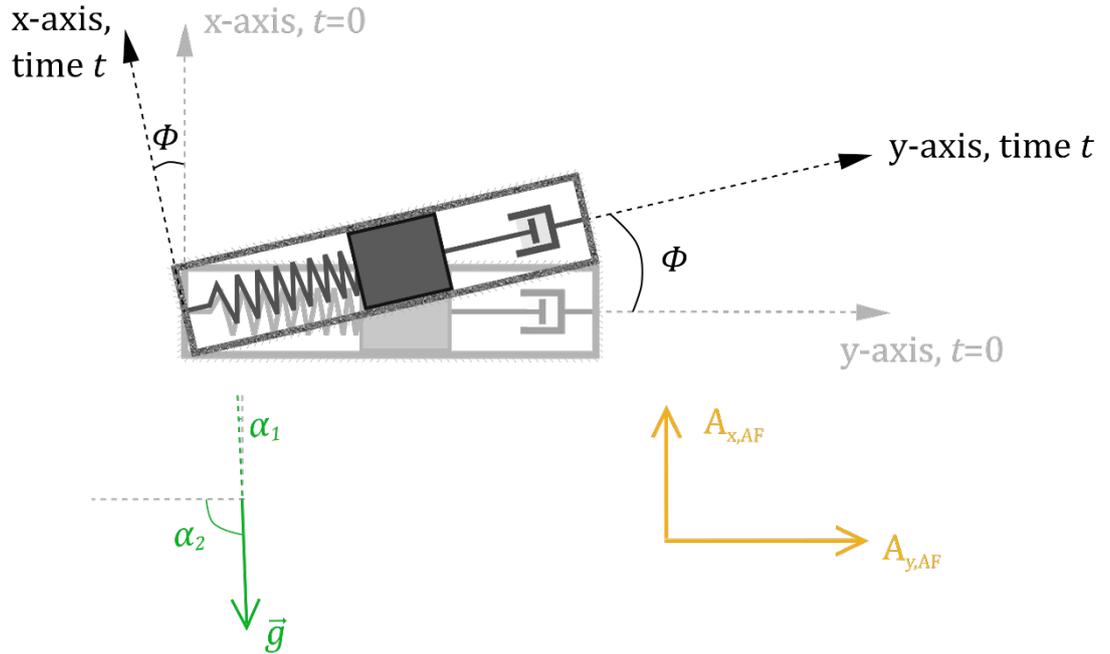
$$\begin{aligned}
a_y &= g \cos \alpha_2 + A_0 + \frac{\phi_o B_{1,1}}{2} + g \cos \alpha_1 \phi_o \sin \omega t + A_{2,1} \cos \omega t + B_{2,1} \sin \omega t \\
&+ \frac{\phi_o A_{1,1}}{2} \sin 2\omega t - \frac{\phi_o B_{1,1}}{2} \cos 2\omega t \\
&+ \left(\sum_{n=2}^N \left(\frac{\phi_o B_{1,n}}{2} \cos \omega(n-1)t + A_{2,n} \cos \omega n t - \frac{\phi_o B_{1,n}}{2} \cos \omega(n+1)t \right) \right. \\
&\left. - \frac{\phi_o A_{1,n}}{2} \sin \omega(n-1)t + B_{2,n} \sin \omega n t + \frac{\phi_o B_{1,n}}{2} \sin \omega(n+1)t \right)
\end{aligned} \tag{2.4}$$

With equation (2.4) at hand, we can now test: (i) whether gravitational acceleration is obtained upon averaging a_y , (ii) whether dynamic body acceleration is obtained upon subtracting the mean of a_y from total acceleration, and (iii) what the effect of sensor rotation is. Stated differently, and applied to the case under consideration here, the first two questions require that: (i) mean acceleration along \hat{j}_{SF} equal $g \cos \alpha_2$, and (ii) no gravity-related terms appear in the expression for dynamic body acceleration, and that the latter's value along \hat{j}_{SF} equal $A_{y,AF}$.

Let τ be the size of the window over which mean acceleration \bar{a}_y is computed as follows:

$$\bar{a}_y = \frac{1}{\tau} \int_0^\tau a_y dt \tag{2.5}$$

(A)



(B)

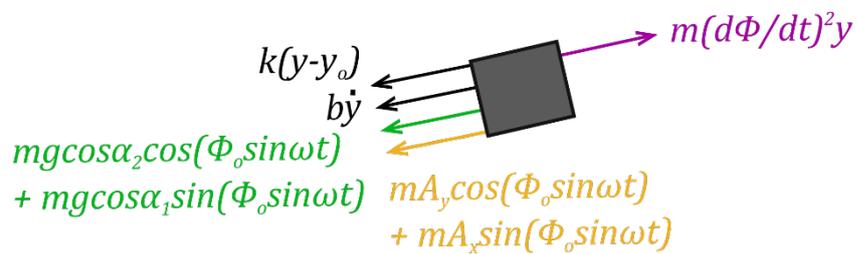


Figure 2.3. **Translating and rotating accelerometer in the presence of gravity.** The triaxial accelerometer (only one axis shown here for clarity) is presumed to be attached to an animal accelerating during a locomotion bout. The directions of the animal's acceleration vector's components $A_{y,AF}$ and $A_{x,AF}$ are specified with respect to the animal's anatomical frame (subscript AF) which is assumed to not change orientation with respect to a global Earth frame. The gravity vector \vec{g} makes angles of $\cos\alpha_2$ and $\cos\alpha_1$ with respect to AF . (A) The accelerometer rotates by an angle $\phi = \phi_0 \sin\omega t$. (B) Free-body diagram of the proof mass at time t .

We observe that, in general, all terms in equation (2.4) will appear in the expression for a_y , and thus \bar{a}_y will not equal static acceleration. Even if $\tau = 2\pi k/\omega$, where an integer number of complete cycles of the base frequency ω fit in the

window such that the integral of all cosine and sine terms is zero, we obtain the following value of \bar{a}_y and dynamic body acceleration along \hat{j}_{SF} ($a_y - \bar{a}_y$):

$$\begin{aligned} \bar{a}_y &= g \cos \alpha_2 + A_0 + \frac{\phi_o B_{1,1}}{2} \\ &\text{and} \\ a_y - \bar{a}_y &= g \cos \alpha_1 \phi_o \sin \omega t + A_{2,1} \cos \omega t \\ &+ B_{2,1} \sin \omega t \frac{\phi_o A_{1,1}}{2} \sin 2\omega t - \frac{\phi_o B_{1,1}}{2} \cos 2\omega t \\ &+ \left(\sum_{n=2}^N \left(\frac{\phi_o B_{1,n}}{2} \cos \omega(n-1)t + A_{2,n} \cos \omega n t - \frac{\phi_o B_{1,n}}{2} \cos \omega(n+1)t \right) \right. \\ &\left. - \left(\frac{\phi_o A_{1,n}}{2} \sin \omega(n-1)t + B_{2,n} \sin \omega n t + \frac{\phi_o B_{1,n}}{2} \sin \omega(n+1)t \right) \right) \end{aligned} \quad (2.6)$$

The terms in \bar{a}_y will not disappear or diminish with increasing values of τ either, although the animal is not likely to continue accelerating for long periods of time. Even when there is no sensor rotation ($\phi_o = 0$), $\bar{a}_{y,\phi_o=0} = g \cos \alpha_2 + A_0$, which is not equal to the assumed value of $g \cos \alpha_2$. When $\phi_o = 0$, however, $a_y - \bar{a}_y$ becomes equal to dynamic body acceleration, $A_{y,SF}$. Finally, equation (2.4) shows that sensor rotation affects the frequency content of the signal by: (i) changing the coefficients of sines and cosines of harmonics that were already present in the animal's acceleration, and (ii) by adding additional harmonics $\omega(n+1)$ that were not present in the animal's acceleration.

2.2.1.3 | Discussion

The strengths and limitations of using the prevalent technique of extracting static acceleration by averaging the measured acceleration signal over a window of given length, and extracting dynamic body acceleration by subtracting static acceleration from total acceleration, were analysed theoretically. Using the spring-mass-damper model for a capacitive accelerometer, I showed that even when the sensor is not subject to rotation, if translational acceleration involves constant acceleration terms during the time period τ of signal averaging, the mean of the signal contains dynamic acceleration terms along with the desired tilt with respect to gravity. Consequently, these constant acceleration terms, erroneously attributed to static acceleration, do not appear in the expression for dynamic body acceleration

since they get subtracted out. If, however, total acceleration does not contain constant acceleration terms during τ , the results correspond to expected values: static acceleration only measures tilt with respect to gravity, and dynamic body acceleration is fully obtained by subtracting static acceleration from total acceleration.

The theoretical formulation of the impact of sensor rotation during time-varying translational acceleration is, to my knowledge, new, and shows that because of sensor rotation: (i) static acceleration can contain components of translational acceleration from multiple body axes (surge and heave in the case considered), and (ii) dynamic body acceleration can include components of gravity. Both these effects are undesirable and deviations from expectation. Further, sensor rotation modifies the frequency content of the underlying body acceleration signal, adding an extra harmonic in the case considered. This is consistent with a previous study where the justification for using a low-pass filter was to eliminate the high-frequency artefacts introduced by a bouncing collar attached to a primate (Sellers et al. 2004). Note that the analysis in this section can be further developed to include the case where, in addition to sensor wobbling, the anatomical frame can also change orientation with respect to the global frame.

In real situations, where animals tune their locomotion frequency according to the requirements of the moment (Ahlborn, Blake & Megill 2006), the net mean of time-varying acceleration cycles will not be zero and will erroneously appear as part of static acceleration. Equations (2.5) and (2.6) (for a_y) nevertheless show that the effect of sinusoidal variation on static acceleration (and hence estimation of dynamic body acceleration) will reduce with increase in the window size over which acceleration measurements are averaged, in line with previous experimental evidence (Shepard et al. 2008b). The appropriateness of prevalent methods to estimate static acceleration and dynamic body acceleration will depend on the application: the impact of errors will be less severe for a behaviour recognition application where differences between behaviours, rather than absolute parameter values, are more important, and more severe for, say, a dead-reckoning application, where lack of precise removal of gravity will lead to errors in position estimation that scale with the square of the time period of integration.

2.2.2 | Magnetometer: can a rotating magnetometer measure angular velocity?

One feature of the magnetometer is that it only measures static tilt. The dynamic component of movement must therefore be obtained indirectly, for instance, through differentiation of the magnetometer signal with respect to time (Kunze et al. 2010). The interest here is to make explicit the extent to which time-differentiation of magnetometer signals can estimate angular velocity. This would form the theoretical foundation needed to develop meaningful features for standalone behaviour recognition using magnetometers in Chapter 4.

2.2.2.1 | Computing angular velocity using a magnetometer

I extend a previous formulation of time-variation in magnetometer signals (Section 6.2 in Afzal, Renaudin & Lachapelle 2011; their formulation is used in equations 2.7 and 2.8 here). Let \vec{b}_G be the Earth's magnetic field intensity vector in an Earth-fixed global frame of reference (hereafter referred to as the global frame). Let the field intensity vector measured by the magnetometer in its frame of reference (hereafter referred to as the sensor frame) be \vec{b}_S . Let the calibration matrix to compute \vec{b}_G from \vec{b}_S be given by the matrix C_S^G . Thus, we obtain:

$$\vec{b}_G = C_S^G \vec{b}_S$$

Conversely, the inverse of this matrix, $(C_S^G)^{-1} = C_G^S$, will permit computation of \vec{b}_G from \vec{b}_S . Thus, we have:

$$C_G^S \vec{b}_G = \vec{b}_S \quad (2.7)$$

Let the sensor frame rotate with an angular velocity of $\vec{\omega}_S$ with respect to the global frame of reference. Upon differentiating equation (2.7) with respect to time, we obtain:

$$C_G^S \dot{\vec{b}}_G = \dot{\vec{b}}_S + (\vec{\omega}_S \times \vec{b}_S)$$

where $\dot{\vec{b}}_G$ is the time-derivative of \vec{b}_G , and $\dot{\vec{b}}_S$ is the time-derivative of \vec{b}_S . According to our assumptions from before, where local magnetic field strength does

not vary with time and is uniform in space, the left-hand side of the above equation is zero. Applying the assumption and rearranging, we obtain:

$$\dot{\vec{b}}_S = -(\vec{\omega}_S \times \vec{b}_S) \quad (2.8)$$

Taking the magnitudes of the quantities on both sides of the above equation, we obtain:

$$|\dot{\vec{b}}_S| = |\vec{\omega}_S| \cdot |\vec{b}_S| \cdot \sin\theta \quad (2.9)$$

where θ is the angle between \vec{b}_S and $\vec{\omega}_S$. Thus, the time-derivative of the measured magnetic field is proportional to the angular speed. However, the proportionality constant depends on the sine of the angle between the field intensity and angular velocity vectors. Thus, when the two vectors are parallel, the magnetometer is incapable of measuring a dynamic component of movement (i.e. $|\dot{\vec{b}}_S| = 0$). In fact, a corollary of equation (2.8) is that the magnetometer cannot be used to measure the component of angular velocity parallel to the magnetic field vector \vec{b}_S , and is only sensitive to the component $\vec{\omega}_{S,\perp}$ which is perpendicular to \vec{b}_S :

$$\vec{\omega}_{S,\perp} = -\frac{\vec{b}_S \times \dot{\vec{b}}_S}{\|\vec{b}_S\|^2} \quad (2.10)$$

2.2.2.2 | Discussion

A theoretical basis for describing the capability and limitation of magnetometers to capture the dynamic component of movement was provided. It was found that the magnitude of the time-derivative of the magnetic field intensity vector is directly proportional to the magnitude of angular velocity, but that this proportionality constant depends on the sine of the angle between the angular velocity and magnetic field vectors. In particular, when these vectors are parallel, the magnetometer can only measure static tilt. This is a disadvantage of magnetometers compared to gyroscopes. However, not only do animal movement paths typically involve frequent direction changes (Benhamou 2004; Sims et al. 2008) but sensor tilt can also change significantly during dynamic movement (e.g. Wilson, Shepard &

Liebsch 2008). Thus, parallel alignment of the two vectors would be expected to occur only rarely, and the time-derivative of magnetometer data could therefore be used to quantify dynamicity of movement, at least for behaviour recognition applications (as opposed to biomechanical applications, where the unknown angle between the axis of rotation and local magnetic field intensity vector could be problematic).

To conclude, this analysis provides a theoretical basis for the magnetometer's capacity to measure a dynamic component of movement: angular velocity. This is significant for behaviour recognition applications for two reasons: (i) meaningful features can be engineered from time-differentiated magnetometer signals, which could help build transparent classification models (Chapter 4), and (ii) angular velocity data acquired using gyroscopes have previously been used in tandem with acceleration data for accurate activity recognition in human studies (Najafi et al. 2003; Paraschiv-Ionescu et al. 2004; Coley et al. 2005). Magnetometers have lower power consumption compared to gyroscopes (Katevas, Haddadi & Tokarchuk 2016) and, if capable of measuring angular velocity, would represent a viable alternative to gyroscopes in future studies.

2.4 | Conclusion

The most widely used inertial sensor, by far, in inferring animal behaviour is the accelerometer. Accelerometers have been used to successfully infer coarse-scale animal behaviours – resting, locomotion, feeding/foraging – across a wide variety of animal species. However, the common principles of static and dynamic acceleration, and periodicity of locomotion, have still not been brought together under a single, general framework that could be employed across species. Further, most cross-validation approaches adopted to quantify model performance have not addressed the main goal of such studies, which is to perform well on data collected on new, unseen individuals. These two issues are addressed in Chapter 3. Magnetometers, despite their potential to measure static (through sensor tilt) and dynamic (through angular velocity) components of movement, and potential for offering complementary

benefits compared to the accelerometer, have not yet been applied to recognise common behaviours. This gap is addressed in Chapter 4. Theoretical analysis of an accelerometer attached to an animal accelerating during locomotion showed that acceleration measurements do not equal dynamic body acceleration when the sensor can wobble about the location of attachment. However, the fact that the sensor can move independently of body acceleration may present the opportunity to recognise fine-scale behaviours; this is exploited in Chapter 5. Theoretical limitations to the extraction of dynamic body acceleration, upon which common techniques of energy expenditure estimation are based, motivate the development of other techniques that do not make assumptions about the nature of the recorded acceleration signal; this is explored in Chapter 6.

References

- Abdulla, U. A., Taylor, K., Barlow, M., & Naqshbandi, K. Z. (2013, July). Measuring walking and running cadence using magnetometers. In *2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications* (pp. 1458-1462). IEEE.
- Afzal, M. H., Renaudin, V., & Lachapelle, G. (2011). Use of earth's magnetic field for mitigating gyroscope errors regardless of magnetic perturbation. *Sensors, 11*(12), 11390-11414.
- Ahlborn, B. K., Blake, R. W., & Megill, W. M. (2006). Frequency tuning in animal locomotion. *Zoology, 109*(1), 43-53.
- Alvarenga, F. A. P., Borges, I., Palkovič, L., Rodina, J., Oddy, V. H., & Dobos, R. C. (2016). Using a three-axis accelerometer to identify and classify sheep behaviour at pasture. *Applied Animal Behaviour Science, 181*, 91-99.
- Aminian, K. (2019). *Chapter 6: Capacitive sensors* (course notes from *Sensors in Medical Instrumentation*, Ecole Polytechnique Fédérale de Lausanne).
- Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys, 4*, 40-79.
- Barré, A., Jolles, B. M., Theumann, N., & Aminian, K. (2015). Soft tissue artifact distribution on lower limbs during treadmill gait: influence of skin markers' location on cluster design. *Journal of biomechanics, 48*(10), 1965-1971.
- Béliveau, A., Spencer, G. T., Thomas, K. A., & Roberson, S. L. (1999). Evaluation of MEMS capacitive accelerometers. *IEEE Design & Test of Computers, 16*(4), 48-56.
- Benhamou, S. (2004). How to reliably estimate the tortuosity of an animal's path:: straightness, sinuosity, or fractal dimension?. *Journal of theoretical biology, 229*(2), 209-220.
- Bidder, O. R., Campbell, H. A., Gómez-Laich, A., Urgé, P., Walker, J., Cai, Y., ... & Wilson, R. P. (2014). Love thy neighbour: automatic animal behavioural

- classification of acceleration data using the k-nearest neighbour algorithm. *PloS one*, 9(2).
- Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., ... & Jones, Z. M. (2016). mlr: Machine Learning in R. *The Journal of Machine Learning Research*, 17(1), 5938-5942.
- Bonnet, S., Bassompierre, C., Godin, C., Leseq, S., & Barraud, A. (2009). Calibration methods for inertial and magnetic sensors. *Sensors and Actuators A: Physical*, 156(2), 302-311.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.
- Brewster, L. R., Dale, J. J., Guttridge, T. L., Gruber, S. H., Hansell, A. C., Elliott, M., ... & Gleiss, A. C. (2018). Development and application of a machine learning algorithm for classification of elasmobranch behaviour from accelerometry data. *Marine biology*, 165(4), 62.
- Broell, F., Taylor, A. D., Litvak, M. K., Bezanson, A., & Taggart, C. T. (2016). Post-tagging behaviour and habitat use in shortnose sturgeon measured with high-frequency accelerometer and PSATs. *Animal Biotelemetry*, 4(1), 11.
- Brownscombe, J. W., Gutowsky, L. F., Danylchuk, A. J., & Cooke, S. J. (2014). Foraging behaviour and activity of a marine benthivorous fish estimated using tri-axial accelerometer biologgers. *Marine Ecology Progress Series*, 505, 241-251.
- Button, V. (2015). *Principles of Measurement and transduction of biomedical variables*. Academic Press.
- Carroll, G., Slip, D., Jonsen, I., & Harcourt, R. (2014). Supervised accelerometry analysis can identify prey capture by penguins at sea. *Journal of Experimental Biology*, 217(24), 4295-4302.
- Caruso, M. J. (2000, March). Applications of magnetic sensors for low cost compass systems. In *IEEE 2000. Position Location and Navigation Symposium (Cat. No. 00CH37062)* (pp. 177-184). IEEE.

- Cheng, H., & Gupta, K. C. (1989). An historical note on finite rotations. *Journal of Applied Mechanics*, 139-145.
- Chimienti, M., Cornulier, T., Owen, E., Bolton, M., Davies, I. M., Travis, J. M., & Scott, B. E. (2016). The use of an unsupervised learning approach for characterizing latent behaviors in accelerometer data. *Ecology and evolution*, 6(3), 727-741.
- Coley, B., Najafi, B., Paraschiv-Ionescu, A., & Aminian, K. (2005). Stair climbing detection during daily physical activity using a miniature gyroscope. *Gait & posture*, 22(4), 287-294.
- Esterman, M., Tamber-Rosenau, B. J., Chiu, Y. C., & Yantis, S. (2010). Avoiding non-independence in fMRI data analysis: leave one subject out. *Neuroimage*, 50(2), 572-576.
- Fehlmann, G., O'Riain, M. J., Hopkins, P. W., O'Sullivan, J., Holton, M. D., Shepard, E. L., & King, A. J. (2017). Identification of behaviours from accelerometer data in a wild social primate. *Animal Biotelemetry*, 5(1), 6.
- Fong, W. T., Ong, S. K., & Nee, A. Y. C. (2008). Methods for in-field user calibration of an inertial measurement unit without external equipment. *Measurement Science and technology*, 19(8), 085202.
- Forner-Cordero, A., Mateu-Arce, M., Forner-Cordero, I., Alcántara, E., Moreno, J. C., & Pons, J. L. (2008). Study of the motion artefacts of skin-mounted inertial sensors under different attachment conditions. *Physiological measurement*, 29(4), N21.
- Gao, L., Campbell, H. A., Bidder, O. R., & Hunter, J. (2013). A Web-based semantic tagging and activity recognition system for species' accelerometry data. *Ecological Informatics*, 13, 47-56.
- Gebre-Egziabher, D., Elkaim, G. H., Powell, J. D., & Parkinson, B. W. (2001, May). A non-linear, two-step estimation algorithm for calibrating solid-state strapdown magnetometers. In *8th International St. Petersburg Conference on Navigation Systems (IEEE/AIAA)*.

- Gebre-Egziabher, D., Elkaim, G. H., David Powell, J., & Parkinson, B. W. (2006). Calibration of strapdown magnetometers in magnetic field domain. *Journal of Aerospace Engineering*, 19(2), 87-102.
- Geisser, S. (1974). A predictive approach to the random effect model. *Biometrika*, 61(1), 101-107.
- Gervasi, V., Brunberg, S., & Swenson, J. E. (2006). An individual-based method to measure animal activity levels: a test on brown bears. *Wildlife Society Bulletin*, 34(5), 1314-1319.
- Grünewälder, S., Broekhuis, F., Macdonald, D. W., Wilson, A. M., McNutt, J. W., Shawe-Taylor, J., & Hailes, S. (2012). Movement activity based classification of animal behaviour with an application to data from cheetah (*Acinonyx jubatus*). *PloS one*, 7(11).
- Guilford, T., Meade, J., Willis, J., Phillips, R. A., Boyle, D., Roberts, S., ... & Perrins, C. M. (2009). Migration and stopover in a small pelagic seabird, the Manx shearwater *Puffinus puffinus*: insights from machine learning. *Proceedings of the Royal Society B: Biological Sciences*, 276(1660), 1215-1223.
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar), 1157-1182.
- Halsey, L. G., Shepard, E. L., & Wilson, R. P. (2011). Assessing the development and application of the accelerometry technique for estimating energy expenditure. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 158(3), 305-314.
- Hammond, T. T., Springthorpe, D., Walsh, R. E., & Berg-Kirkpatrick, T. (2016). Using accelerometers to remotely and automatically characterize behavior in small animals. *Journal of Experimental Biology*, 219(11), 1618-1624.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.
- Hochscheid, S., & Wilson, R. P. (1999). A new method for the determination of at-sea activity in sea turtles. *Marine ecology progress series*, 185, 293-296.

- IEEE Std 528–2019, IEEE Standard for Inertial Sensor Terminology.
- Jeanniard-du-Dot, T., Guinet, C., Arnould, J. P., Speakman, J. R., & Trites, A. W. (2017). Accelerometers can measure total and activity-specific energy expenditures in free-ranging marine mammals only if linked to time-activity budgets. *Functional Ecology*, *31*(2), 377-386.
- Kamminga, J. W., Le, D. V., Meijers, J. P., Bisby, H., Meratnia, N., & Havinga, P. J. (2018). Robust sensor-orientation-independent feature selection for animal activity recognition on collar tags. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(1), 1-27.
- Katevas, K., Haddadi, H., & Tokarchuk, L. (2016, September). Sensingkit: Evaluating the sensor power consumption in ios devices. In *2016 12th International Conference on Intelligent Environments (IE)* (pp. 222-225). IEEE.
- Kooyman, G. L. (2004). Genesis and evolution of bio-logging devices: 1963-2002.
- Kunze, K., Bahle, G., Lukowicz, P., & Partridge, K. (2010, October). Can magnetic field sensors replace gyroscopes in wearable sensing applications?. In *International Symposium on Wearable Computers (ISWC) 2010* (pp. 1-4). IEEE.
- Ladds, M. A., Thompson, A. P., Slip, D. J., Hocking, D. P., & Harcourt, R. G. (2016). Seeing it all: evaluating supervised machine learning methods for the classification of diverse otariid behaviours. *PLoS One*, *11*(12).
- Ladds, M. A., Thompson, A. P., Kadar, J. P., Slip, D. J., Hocking, D. P., & Harcourt, R. G. (2017). Super machine learning: improving accuracy and reducing variance of behaviour classification from accelerometry. *Animal Biotelemetry*, *5*(1), 8.
- Lagarde, F., Guillon, N., Dubroca, L., Bonnet, X., Kaddour, K. B., Slimani, T., & El Mouden, E. H. (2008). Slowness and acceleration: a new method to quantify the activity budget of chelonians.

- Laich, A. G., Wilson, R. P., Quintana, F., & Shepard, E. L. (2008). Identification of imperial cormorant *Phalacrocorax atriceps* behaviour using accelerometers. *Endangered species research*, 10, 29-37.
- Larson, S. C. (1931). The shrinkage of the coefficient of multiple correlation. *Journal of Educational Psychology*, 22(1), 45.
- Lenz, J. E. (1990). A review of magnetic sensors. *Proceedings of the IEEE*, 78(6), 973-989.
- Lush, L., Ellwood, S., Markham, A., Ward, A. I., & Wheeler, P. (2016). Use of tri-axial accelerometers to assess terrestrial mammal behaviour in the wild. *Journal of Zoology*, 298(4), 257-265.
- Marais, J., Le Roux, S. P., Wolhuter, R., & Niesler, T. (2014, November). Automatic classification of sheep behaviour using 3-axis accelerometer data. In *Proceedings of the twenty-fifth annual symposium of the Pattern Recognition Association of South Africa (PRASA)* (pp. 97-102).
- Martiskainen, P., Järvinen, M., Skön, J. P., Tiirikainen, J., Kolehmainen, M., & Mononen, J. (2009). Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Applied animal behaviour science*, 119(1-2), 32-38.
- McClune, D. W., Marks, N. J., Wilson, R. P., Houghton, J. D., Montgomery, I. W., McGowan, N. E., ... & Scantlebury, M. (2014). Tri-axial accelerometers quantify behaviour in the Eurasian badger (*Meles meles*): towards an automated interpretation of field data. *Animal Biotelemetry*, 2(1), 5.
- Mitani, Y., Sato, K., Ito, S., Cameron, M. F., Siniff, D. B., & Naito, Y. (2003). A method for reconstructing three-dimensional dive profiles of marine mammals using geomagnetic intensity data: results from two lactating Weddell seals. *Polar Biology*, 26(5), 311-317.
- Moreau, M., Siebert, S., Buerkert, A., & Schlecht, E. (2009). Use of a tri-axial accelerometer for automated recording and classification of goats' grazing behaviour. *Applied animal behaviour science*, 119(3-4), 158-170.

- Müller, R., & Schrader, L. (2003). A new method to measure behavioural activity levels in dairy cows. *Applied Animal Behaviour Science*, 83(4), 247-258.
- Najafi, B., Aminian, K., Paraschiv-Ionescu, A., Loew, F., Bula, C. J., & Robert, P. (2003). Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. *IEEE Transactions on biomedical Engineering*, 50(6), 711-723.
- 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.
- Nishizawa, H., Noda, T., Yasuda, T., Okuyama, J., Arai, N., & Kobayashi, M. (2013). Decision tree classification of behaviors in the nesting process of green turtles (*Chelonia mydas*) from tri-axial acceleration data. *Journal of ethology*, 31(3), 315-322.
- Pagano, A. M., Rode, K. D., Cutting, A., Owen, M. A., Jensen, S., Ware, J. V., ... & Middel, K. R. (2017). Using tri-axial accelerometers to identify wild polar bear behaviors. *Endangered Species Research*, 32, 19-33.
- Paraschiv-Ionescu, A., Buchser, E. E., Rutschmann, B., Najafi, B., & Aminian, K. (2004). Ambulatory system for the quantitative and qualitative analysis of gait and posture in chronic pain patients treated with spinal cord stimulation. *Gait & posture*, 20(2), 113-125.
- Parkin, S. S., Kaiser, C., Panchula, A., Rice, P. M., Hughes, B., Samant, M., & Yang, S. H. (2004). Giant tunnelling magnetoresistance at room temperature with MgO (100) tunnel barriers. *Nature materials*, 3(12), 862-867.
- Parvis, M., Ferraris, F., & Grimaldi, U. (1995). Procedure for effortless in-field calibration of three-axis rate gyros and accelerometers. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.

- Popovic, R. S. (2014, May). High resolution Hall magnetic sensors. In *2014 29th International Conference on Microelectronics Proceedings-MIEL 2014* (pp. 69-74). IEEE.
- Primdahl, F. (1979). The fluxgate magnetometer. *Journal of Physics E: Scientific Instruments*, *12*(4), 241.
- Resheff, Y. S., Rotics, S., Harel, R., Spiegel, O., & Nathan, R. (2014). AcceleRater: a web application for supervised learning of behavioral modes from acceleration measurements. *Movement ecology*, *2*(1), 27.
- Robert, B., White, B. J., Renter, D. G., & Larson, R. L. (2009). Evaluation of three-dimensional accelerometers to monitor and classify behavior patterns in cattle. *Computers and Electronics in Agriculture*, *67*(1-2), 80-84.
- Ropert-Coudert, Y., Grémillet, D., Kato, A., Ryan, P. G., Naito, Y., & Le Maho, Y. (2004). A fine-scale time budget of Cape gannets provides insights into the foraging strategies of coastal seabirds. *Animal Behaviour*, *67*(5), 985-992.
- Sato, K., Sakamoto, K. Q., Watanuki, Y., Takahashi, A., Katsumata, N., Bost, C. A., & Weimerskirch, H. (2009). Scaling of soaring seabirds and implications for flight abilities of giant pterosaurs. *PloS one*, *4*(4).
- Sellers, W. I., & Crompton, R. H. (2004). Automatic monitoring of primate locomotor behaviour using accelerometers. *Folia primatologica*, *75*(4), 279-293.
- Shahriar, M. S., Smith, D., Rahman, A., Freeman, M., Hills, J., Rawnsley, R., ... & Bishop-Hurley, G. (2016). Detecting heat events in dairy cows using accelerometers and unsupervised learning. *Computers and electronics in agriculture*, *128*, 20-26.
- Sims, D. W., Southall, E. J., Humphries, N. E., Hays, G. C., Bradshaw, C. J., Pitchford, J. W., ... & Morritt, D. (2008). Scaling laws of marine predator search behaviour. *Nature*, *451*(7182), 1098-1102.
- Shepard, E. L., Wilson, R. P., Quintana, F., Laich, A. G., Liebsch, N., Albareda, D. A., ... & Newman, C. (2008a). Identification of animal movement patterns using tri-axial accelerometry. *Endangered species research*, *10*, 47-60.

- Shepard, E. L., Wilson, R. P., Halsey, L. G., Quintana, F., Laich, A. G., Gleiss, A. C., ... & Norman, B. (2008b). Derivation of body motion via appropriate smoothing of acceleration data. *Aquatic Biology*, 4(3), 235-241.
- Soltis, J., Wilson, R. P., Douglas-Hamilton, I., Vollrath, F., King, L. E., & Savage, A. (2012). Accelerometers in collars identify behavioral states in captive African elephants *Loxodonta africana*. *Endangered Species Research*, 18(3), 255-263.
- Sur, M., Suffredini, T., Wessells, S. M., Bloom, P. H., Lanzone, M., Blackshire, S., ... & Katzner, T. (2017). Improved supervised classification of accelerometry data to distinguish behaviors of soaring birds. *PloS one*, 12(4).
- Suzuki, Y. S., & Tai, Y. C. (2006). Micromachined high-aspect-ratio parylene spring and its application to low-frequency accelerometers. *Journal of Microelectromechanical Systems*, 15(5), 1364-1370.
- Suzuki, I., Naito, Y., Folkow, L. P., Miyazaki, N., & Blix, A. S. (2009). Validation of a device for accurate timing of feeding events in marine animals. *Polar Biology*, 32(4), 667-671.
- Van Oort, B. E., Tyler, N. J., Storeheier, P. V., & Stokkan, K. A. (2004). The performance and validation of a data logger for long-term determination of activity in free-ranging reindeer, *Rangifer tarandus* L. *Applied Animal Behaviour Science*, 89(3-4), 299-308.
- Volpov, B. L., Hoskins, A. J., Battaile, B. C., Viviant, M., Wheatley, K. E., Marshall, G., ... & Arnould, J. P. (2015). Identification of prey captures in Australian fur seals (*Arctocephalus pusillus doriferus*) using head-mounted accelerometers: field validation with animal-borne video cameras. *PLoS One*, 10(6).
- Walker, J. S., Jones, M. W., Laramée, R. S., Holton, M. D., Shepard, E. L., Williams, H. J., ... & Bidder, O. R. (2015). Prying into the intimate secrets of animal lives; software beyond hardware for comprehensive annotation in 'Daily Diary' tags. *Movement ecology*, 3(1), 29.
- 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), 2.

- Watanabe, S., Izawa, M., Kato, A., Ropert-Coudert, Y., & Naito, Y. (2005). A new technique for monitoring the detailed behaviour of terrestrial animals: a case study with the domestic cat. *Applied Animal Behaviour Science*, 94(1-2), 117-131.
- Whitney, N. M., Papastamatiou, Y. P., Holland, K. N., & Lowe, C. G. (2007). Use of an acceleration data logger to measure diel activity patterns in captive whitetip reef sharks, *Triaenodon obesus*. *Aquatic Living Resources*, 20(4), 299-305.
- Williams, H. J., Shepard, E. L. C., Duriez, O., & Lambertucci, S. A. (2015). Can accelerometry be used to distinguish between flight types in soaring birds?. *Animal Biotelemetry*, 3(1), 45.
- Williams, H. J., Holton, M. D., Shepard, E. L., Largey, N., Norman, B., Ryan, P. G., ... & Marks, N. J. (2017). Identification of animal movement patterns using triaxial magnetometry. *Movement ecology*, 5(1), 6.
- Wilson, R. P., & Wilson, M. P. (1988). Dead reckoning: a new technique for determining penguin movements at sea. *Meeresforschung (Hamburg)*, 32(2), 155-158.
- Wilson, R. P., Wilson, M. P. T., Link, R., Mempel, H., & Adams, N. J. (1991). Determination of movements of African penguins *Spheniscus demersus* using a compass system: dead reckoning may be an alternative to telemetry. *Journal of Experimental Biology*, 157(1), 557-564.
- Wilson, R. P., White, C. R., Quintana, F., Halsey, L. G., Liebsch, N., Martin, G. R., & Butler, P. J. (2006). Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *Journal of Animal Ecology*, 75(5), 1081-1090.
- Wilson, R. P., Shepard, E. L. C., & Liebsch, N. (2008). Prying into the intimate details of animal lives: use of a daily diary on animals. *Endangered species research*, 4(1-2), 123-137.

- Wilson, R. P., Holton, M. D., di Virgilio, A., Williams, H., Shepard, E. L., Lambertucci, S., ... & Srivastava, M. (2018). Give the machine a hand: A Boolean time-based decision-tree template for rapidly finding animal behaviours in multisensor data.
- Yoda, K., Sato, K., Niizuma, Y., Kurita, M., Bost, C., Le Maho, Y., & Naito, Y. (1999). Precise monitoring of porpoising behaviour of Adélie penguins determined using acceleration data loggers. *Journal of Experimental Biology*, 202(22), 3121-3126.
- Yoda, K., Naito, Y., Sato, K., Takahashi, A., Nishikawa, J., Ropert-Coudert, Y., ... & Le Maho, Y. (2001). A new technique for monitoring the behaviour of free-ranging Adélie penguins. *Journal of Experimental Biology*, 204(4), 685-690.

CHAPTER 3

Coarse-Scale Behaviour Recognition Using Accelerometers³

³ This chapter was published as, “Chakravarty, P., Cozzi, G., Ozgul, A., & Aminian, K. (2019). A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods in Ecology and Evolution*, 10(6), 802-814.”

Abstract

1. Data from animal-borne inertial sensors are widely used to investigate several aspects of an animal's life, such as energy expenditure, daily activity patterns, and behaviour. Accelerometer data used in conjunction with machine learning algorithms has been the tool of choice for characterising animal behaviour. Although machine learning models perform reasonably well, they may not rely on meaningful features, nor lend themselves to physical interpretation of the classification rules. This lack of interpretability and control over classification outcomes is of particular concern where different behaviours have different frequency of occurrence and duration, as in most natural systems, and calls for the development of alternative methods. Biomechanical approaches to human activity classification could overcome these shortcomings, yet their full potential remains untapped for animal studies.
2. We propose a general framework for behaviour recognition using accelerometers, and develop a hybrid model where 1) biomechanical features characterise movement dynamics, and 2) a node-based hierarchical classification scheme employs simple machine learning algorithms at each node to find feature-value thresholds separating different behaviours. Using tri-axial accelerometer data collected on ten wild Kalahari meerkats, and annotated video recordings of each individual as groundtruth, this hybrid model was validated in three scenarios: (1) when each behaviour was equally represented (EQDIST), (2) when naturally imbalanced datasets were considered (STRAT), and (3) when data from new individuals were considered (LOIO).
3. A linear-kernel Support Vector Machine at each node of our classification scheme yielded an overall accuracy of >95% for each scenario. Our hybrid approach had a 2.7% better average overall accuracy than top-performing classical machine learning approaches. Further, we showed that not all models with high overall accuracy returned accurate behaviour-specific performance, and good performance during EQDIST did not always generalise to STRAT and LOIO.
4. Our hybrid model took advantage of robust machine learning algorithms for automatically estimating decision boundaries between behavioural classes.

This not only achieved high classification performance but also permitted biomechanical interpretation of classification outcomes. The framework presented here provides the flexibility to adapt models to required levels of behavioural resolution, and has the potential to facilitate meaningful model sharing between studies.

3.1 | Introduction

An in-depth understanding of wild animal behaviour and movement has assumed prime importance in recent years in light of an urgent need to augment our forecasting, conservation and management capacities in the face of rapid environmental change. Since gathering data on wild animals in their natural habitats is often precluded by logistical difficulty, animal-borne sensor systems that offer the possibility of continuously and remotely recording data as the animal goes about its daily life have been developed (Ropert-Coudert and Wilson 2005). Over the two decades since their first application to remote monitoring of animal behaviour (e.g., Sellers et al. 1998), animal-borne accelerometers have been employed to gain insights into the life histories of species as far apart in size, type of habitat, speed and mode of locomotion as chipmunks (*Tamias alpinus*) (Hammond et al. 2016), seals (*Leptonychotes weddellii*) (Naito et al. 2010), African wild dogs (*Lycaon pictus*) and cheetahs (*Acinonyx jubatus*) (Cozzi et al. 2012), vultures (*Gyps fulvus*) (Nathan et al. 2012), and blue whales (*Balaenoptera musculus*) (Goldbogen et al. 2011). To keep pace with the increasing ubiquity of accelerometry, developing methods applicable across species has become essential. For instance, general methods have been developed to infer animal energy expenditure from recorded acceleration (Wilson et al. 2006). However, despite progress in recognising animal behaviour from recorded acceleration, there is as yet no single technique that combines under one framework the virtues of easy-to-implement machine learning on the one hand, and the interpretability and robustness of biomechanically defined classification rules on the other.

Machine learning, often rather unenviably referred to as a black box (e.g. McClune et al. 2014), has been widely employed to infer animal behaviour from raw acceleration data (Nathan et al. 2012; Grünewälder et al. 2012; Gao et al. 2013; Bidder et al. 2014; Resheff et al. 2014; Wang et al. 2015). Despite the power and ease of use of machine learning approaches, the need to develop behaviour recognition tools generalizable across species has brought to light the importance of being able to physically interpret classification rules, even at the cost of small gains in classification accuracy (Nathan et al. 2012). Model interpretability becomes particularly significant when accurately recognising relatively rarer behaviours is important. When machine learning approaches are applied to datasets that are heavily

skewed in the frequency and duration of different behaviours (Watanabe et al. 2005; Grünewälder et al. 2012; Resheff et al. 2014; Wang et al. 2015), large overall accuracies may be obtained even when the recognition accuracy of under-represented behaviours is poor (He and Garcia 2009); the issue of imbalanced classes has even been called a ‘curse’ in machine learning literature (Lemaitre, Nogueira, and Aridas 2017). This occurs because machine learning algorithms typically seek to minimise the overall misclassification rate, and thus tend to optimise for the most frequent behaviours since they contribute most to overall accuracy – the classification performance of rare behaviours may nevertheless remain poor. Examination of the classification rules would help understand how to improve the recognition of rarer behaviours. However, the use of many features, which is typical with machine learning approaches (>15: Nathan et al. 2012; Resheff et al. 2014; Wang et al. 2015), implies that the resulting classification rules are a function of that many variables, and deciphering these high-dimensional complex rules and pinpointing the reasons behind misclassification of under-represented (rarer) but important behaviours quickly becomes intractable.

Confining the role of machine learning algorithms to threshold-finding within a classification scheme pre-defined on the basis of biomechanical knowledge of animal movement dynamics, and using biomechanically significant movement descriptors (features) within the scheme, may help solve these difficulties. Such an approach is made possible by the fact that there are certain natural commonalities in the movement biomechanics of any animal: different behaviours may be carried out more or less intensely, in characteristic postures, erratically or in a periodic manner. Even though the actual names for different animal behaviours will depend on the environment and animal’s context, broad behavioural categories will still lend themselves to description through the biomechanical principles of intensity, posture, and periodicity.

Several studies have employed biomechanical considerations to characterise behaviours central to an animal’s life, such as foraging, locomotion, and resting (e.g. Shepard et al. 2008); in these studies, a majority of the acceleration-derived metrics used for behaviour recognition were aimed at quantifying the three biomechanical principles of intensity, posture, and periodicity. For instance, one study that categorised acceleration data in Adélie penguins (*Pygoscelis adeliae*) (Yoda et al.

2001) considered the division of behavioural classes into static and dynamic categories, discriminated between static behaviours based on posture, and identified locomotion as being periodic. The use of periodicity metrics may help tease apart even highly context-, environment-, and mode-of-locomotion-dependent behaviours such as locomotion and foraging. For instance, periodicity metrics have since been used to characterise locomotor movement across avian (Laich et al. 2008), marine (Kawabe et al. 2003), and terrestrial (Soltis et al. 2012) species; the aperiodicity of foraging has been alluded to for Little Penguins (*Eudyptula minor*) diving for unpredictably located and mobile prey (Ropert-Coudert et al. 2006; Watanabe and Takahashi 2013). Despite extensive use of such biomechanically significant parameters in their behaviour recognition schemes, previous methods leave room for improvement – parameter thresholds for building classification rules were manually chosen (Laich et al. 2008), and training data used for creating the model came only from a few captive animals (McClune et al. 2014; Soltis et al. 2012).

In this study, we present a general behaviour recognition framework in the form of a hybrid model that combines general biomechanical principles on the one hand, and machine learning tools on the other. The proposed hybrid model consists of a tree-like classification framework pre-defined on the basis of biomechanical considerations, where specific combinations of acceleration-derived biomechanical descriptors capturing movement biomechanics of behavioural categories studied across a range of species are used in conjunction with robust machine learning algorithms at each node of the tree. The use of biomechanics to instruct the classification makes the model interpretable, and the use of machine learning at each node of the tree completely automates the search for optimal metric thresholds separating different behaviours. We showcase the application and benefits of this hybrid behaviour recognition framework on data collected on wild meerkats (*Suricata suricatta*), where the classification of their main activities such as vigilance, foraging, resting, and running is essential for characterising their social and individual behaviour. We validated this hybrid model against synchronised, annotated video-camera footage considered as the groundtruth, and compared its performance with benchmark measures obtained with alternative classification methods based entirely on classical machine learning.

3.2 | Material and Methods

3.2.1 | Biomechanically Driven Behaviour Recognition and Validation

3.2.1.1 | Quantifying movement biomechanics through acceleration-derived features

Raw tri-axial acceleration data may be summarised in the form of quantifiable biomechanical descriptors of movement, which can then be used as features capable of discriminating between different behavioural categories. The features to be developed will be strongly dependent on the desired ethogram; here we shall consider an archetypal, general ethogram consisting of three common behaviours – locomotion, resting, and foraging. We focused on three biomechanical principles to characterise these behaviours: posture, movement intensity and periodicity. These principles have been previously used to discriminate between behaviours for a range of species; we aim to synthesise and combine these existing but scattered principles within one biomechanically defined classification framework. We quantify these three descriptors for each fixed-duration sliding window w containing N tri-axial acceleration values recorded along the surge (a_{surge}), sway (a_{sway}) and heave (a_{heave}) axes as follows.

Posture

Common measures of posture involve estimating how the sensor is oriented with respect to Earth’s gravity. Computation is done by averaging acceleration recorded along each axis of the accelerometer, assuming that acceleration due to dynamic bodily movement shows up as oscillations around a constant, static value corresponding to Earth’s gravity (e.g. Yoda et al. 2001). For instance, the mean of surge acceleration in window w , $a_{surge,w}$, may be computed as:

$$meanSurge_w = \frac{\sum_N a_{surge,w}}{N} \quad (1)$$

Intensity

Contrary to posture estimation, movement intensity is commonly characterised by quantifying dynamic acceleration, which is obtained by removing static acceleration corresponding to Earth’s gravity from total recorded acceleration.

For this, we considered the use of $stdNorm_w$, the standard deviation of the Euclidean norm of the tri-axial acceleration vector ($|\vec{a}_w|$), which is equivalent to computing Vectorial Dynamic Body Acceleration (VeDBA, McGregor et al. 2009):

$$stdNorm_w = std(|\vec{a}_w|) \quad (2)$$

Periodicity

To quantify movement periodicity, we analysed the frequency content of the signal through the use of the Fourier transform (FT). The FT of aperiodic signals such as acceleration recorded during foraging will be relatively flat, whereas that computed for periodic signals such as those recorded during locomotion will be marked by the presence of a clear, tall peak, usually at a characteristic frequency. This peak height was considered as a measure of signal periodicity, and may be computed as follows. For each window w , the acceleration signal from each of the three axes recorded at a sampling frequency of F_s may be: 1) normalised (zero mean and unit energy); 2) low-pass filtered; 3) zero-padded and windowed; and 4) FT-ed with a resolution of U Hz (corresponding to FT computation at $L = F_s/U$ equally spaced frequencies between 0 and F_s). From the FT of acceleration along each axis, the square of the magnitude of each Fourier coefficient ($c_{f_i,w,sway}^2, c_{f_i,w,surge}^2, c_{f_i,w,heave}^2, i \in 1 \dots L$), corresponding to the power of the signal at frequency f_i , may be computed and averaged over the three axes. Finally, from this axis-averaged FT, the maximum power obtained across all frequencies $f_i (i \in 1 \dots L)$ may be chosen as a measure of the periodicity of the signal. This feature will henceforth be referred to as *fftPeakPowerAvg*.

$$fftPeakPowerAvg_w = \max_L \frac{c_{f_i,w,sway}^2 + c_{f_i,w,surge}^2 + c_{f_i,w,heave}^2}{3} \quad (3)$$

Note that actual values of the signal processing parameters, such as type, order and cut-off frequency of the low-pass filter, length of zero-padding, type of window, and U , will depend on the characteristics of the recorded acceleration signal, such as signal bandwidth and sampling frequency.

3.2.1.2 | A biomechanically defined hierarchical classification scheme with automated feature-threshold computation

The proposed approach involves pre-defining a hierarchical tree-like scheme that classifies broader behavioural categories into increasingly specific ones up to the desired level of behavioural resolution. Each node of this tree uses one or more features tailored to the classification at that node. A pre-defined hierarchical scheme has two interesting characteristics that make it more advantageous over the common classical machine learning approach of directly classifying behaviours at their highest resolution. The first advantage concerns the ease of dealing with imbalanced classes. It is quite probable that when specific behaviours are grouped into compound categories at the higher-placed nodes of the tree, a relatively rarer behaviour gets clubbed into the same compound category with a more frequent behaviour. As a result, the problem of imbalanced classes is not encountered at the higher-placed node, and is deferred to a lower node where the rare behaviour can no longer be grouped into a compound category with the more frequent behaviour. Thus, if poor recognition accuracy of a rare behaviour is encountered, one can precisely identify the node at which the misclassification occurs. Moreover, since the input features are tailored to each node, one can understand which features to add or modify at the culprit node to improve classification. This process of optimising the model for each individual behaviour may be very difficult to do with machine learning approaches that classify directly up to the finest behaviour resolution level using classification rules which are a function of many (>15) features. Secondly, when higher behavioural resolution is desirable, specific behaviours themselves can be considered as compound categories and separated into finer behaviours by adding nodes below them. For instance, adding a node below 'locomotion' could distinguish between slow and fast locomotion. In classical machine learning, if behavioural classes are added retrospectively, the entire model would have to be built anew.

The precise form of the classification tree for a specific application will depend on the ethogram of interest, so will the features to be given as inputs to each node. Here we demonstrate the construction of a classification tree for the archetypal ethogram considered in Section 2.1.1 consisting of locomotion, resting, and foraging. For the classification of these behaviours, the first node would separate the static (resting) and dynamic (locomotion and foraging) behavioural categories. At the

second stage of this tree, one node below the dynamic category would separate the two dynamic behaviours: locomotion and foraging. In this scheme, each node divides a parent behavioural category into exactly two daughter behavioural types. To accomplish this binary classification at each node, appropriate user-chosen biomechanically significant features (cf. Section 2.1.1) may be given as inputs to commonly employed binary machine learning algorithms (such as Support Vector Machines (SVM)) to obtain optimal feature-value thresholds in a completely automated fashion. For instance, at the first node of the classification tree described here, *stdNorm*, a measure of movement intensity, may be used as a single-feature input to an SVM that will separate behaviours into the static and dynamic behavioural categories.

3.2.1.3 | Model validation

To test the utility of a behaviour recognition method, one needs to validate the predictions made by it against groundtruthed data – the latter typically coming from direct observation or video annotation of the behaviours of interest. Typically, a video camera synchronised with the animal-borne accelerometer is used to film the animal while the accelerometer records data; the groundtruth is then obtained by having an expert assign behaviour labels to sections of the video based on a suitably defined ethogram. This process is then repeated for multiple individuals to capture inter-individual variation in behaviours. Finally, windows of acceleration corresponding to observed behaviours of interest are extracted from the continuous acceleration data stream to obtain bouts of labelled behaviour of fixed duration. To evaluate the effect of imbalanced datasets and inter-individual variability on model classification, we discuss three different forms of cross-validation.

Equally distributed behaviour ten-fold cross-validation (EQDIST)

EQDIST evaluates model performance when the dataset has an equal number of bouts of each behaviour. It involves conducting standard ten-fold cross-validation on sub-sampled datasets where the sample size for each behaviour is made equal. This is done by first pooling data from all individuals together, and then randomly selecting as many bouts from each behaviour as the one with the least number of bouts.

Stratified ten-fold cross-validation (STRAT)

STRAT evaluates model performance when some behaviours may be under-represented or rarer in the dataset than others. It involves pooling data from all individuals together and then dividing the pooled dataset into ten equal parts in such a way that the proportion of bouts from each type of behaviour in each fold is equal to that in the entire, unmodified dataset.

Leave-one-individual-out cross-validation (LOIO)

LOIO evaluates model performance when inter-individual variation is taken into account. It involves training a model using data pooled over all individuals except one, and then testing this model on data from the individual left out. This process is repeated until each individual has been the ‘test’ individual once.

We used confusion matrix-based metrics to evaluate and compare model performance. These performance statistics included three behaviour-specific metrics (sensitivity, precision and specificity), and overall model accuracy (see Appendix S3.1 for definitions).

3.2.2 | Case Study: Kalahari Meerkats

3.2.2.1 | Data collection and groundtruthing

Fieldwork was conducted at the Kalahari Meerkat Project, a long-term study of wild meerkats in the South African Kalahari Desert (Clutton-Brock et al. 1999). Eleven data-recording sessions were done on ten adult meerkats (7 males, 3 females, body mass 667 ± 98 grams, age 24 ± 15 months); one of the individuals was recorded twice. For each individual, data was collected over three hours in the morning, which corresponds to the duration of normal morning meerkat activity (meerkats become inactive as temperatures soar around the midday hours). Individuals were captured using methodology described in Jordan et al. (2009) and fitted with a collar equipped with an inertial measurement unit (IMU) (adapted version of Physilog IV, GaitUp SA, Switzerland) of size 35 mm x 29 mm x 19 mm and measuring tri-axial acceleration at 100 Hz/axis with a range of $\pm 156.96 \text{ m/s}^2$ (corresponding to ± 16

times the acceleration due to Earth's gravity) with 16-bit resolution. The overall collar weight was < 25 grams; collars of this size and weight have been shown not to affect meerkat behaviour (Golabek, Jordan and Clutton-Brock, 2008). The collar was positioned so that the axes of the accelerometer were oriented as shown in Figure 3.1. The accelerometer was calibrated prior to each capture according to a standard procedure (Ferraris et al. 1995). After release, the collared animal was filmed using a handheld video camera recording at 25 frames/second and synchronised electronically with the IMU system. All videos were annotated using Solomon Coder (Version: beta 17.03.22). This video annotation served as the groundtruthing for our behaviour recognition scheme.

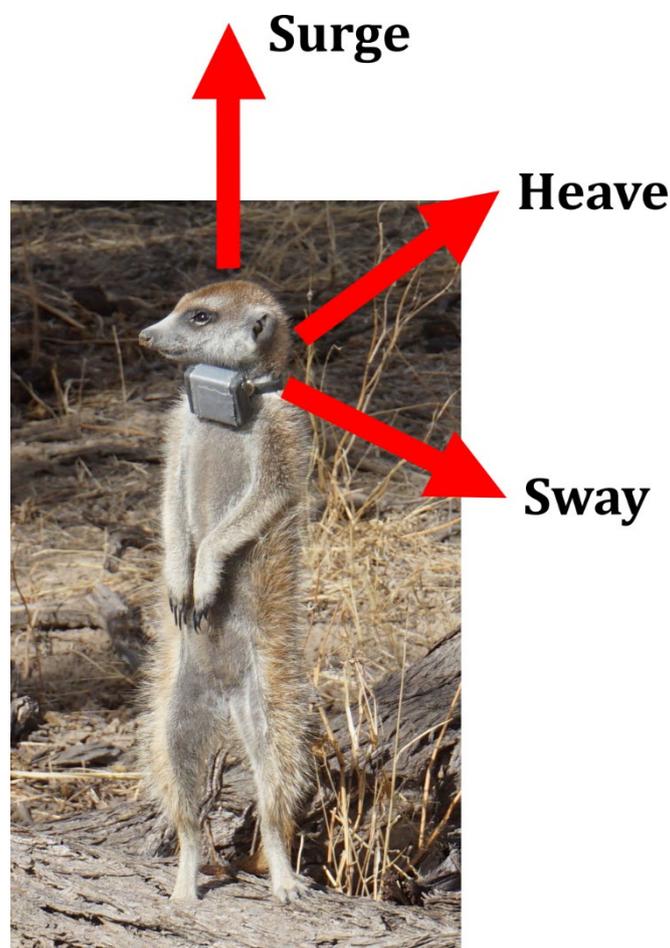


Figure 3.1. **Sensor axes orientation.** The image shows the animal displaying typical bipedal vigilance behaviour.

3.2.2.2 | Meerkat behaviours and hierarchical classification scheme

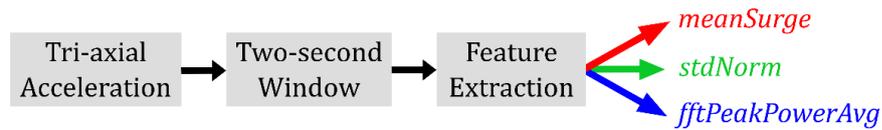
Four different behaviours (Figure 3.2) were considered for the ethogram based on their biological significance:

- *Vigilance*: representative of the animal's general stress or alertness level. The meerkat is stationary and lifts its head and torso up to survey its surroundings.
- *Foraging*: can help derive proxies for body condition, which could, in turn, play a role in determining the animals' survival and social status (Hodge et al. 2008). Most meerkat prey live underground; foraging involves scanning the ground, digging, and handling and ingesting prey.
- *Running*: high-speed locomotion from one point to another with presumably high energy expenditure. Prolonged running events are rare and typically mark important events such as aggressive interactions with rival meerkat groups.
- *Resting*: periods of inactivity, mainly due to fatigue or excessive heat. Typically, the meerkat lies down with its body flat on the ground.

Social context-dependent meerkat behaviours such as grooming and territory marking were excluded from the ethogram.

Based on the description of the behaviours of interest here and the framework presented in Section 2.1, a three-node hierarchical scheme was devised to classify meerkat behaviour, as shown in Figure 3.2. Movement intensity (*stdNorm*) and posture (*meanSurge*) were used to separate static (vigilance and resting) and dynamic (foraging and running) behaviours in the first node. At the second node, posture (*meanSurge*) was used to distinguish vigilance from resting – while the animal's torso is typically upright during vigilance, it is flat during resting. At the third node, movement intensity (*stdNorm*) and periodicity (*fftPeakPowerAvg*) were used to separate foraging from running – running was observed to involve faster and more periodic limb movements compared to foraging. At each node, a separate machine learning algorithm (M1, M2 and M3 in Figure 3.2) was chosen from four candidates: Naïve-Bayes (NB), Linear Discriminant Analysis (LDA), Logistic Regression (LR), and linear-kernel Support Vector Machine (SVM). We shall henceforth refer to each such M1-M2-M3 combination as one 'hybrid model'. Thus, since four candidates were possible at each of the three nodes, a total of $4 \times 4 \times 4 = 64$ hybrid models were tested to find the best one. Scikit-learn (Pedregosa et al. 2011, version 0.19.0) was used to implement all machine learning models (using their default configurations) in Python.

(A)



(B)

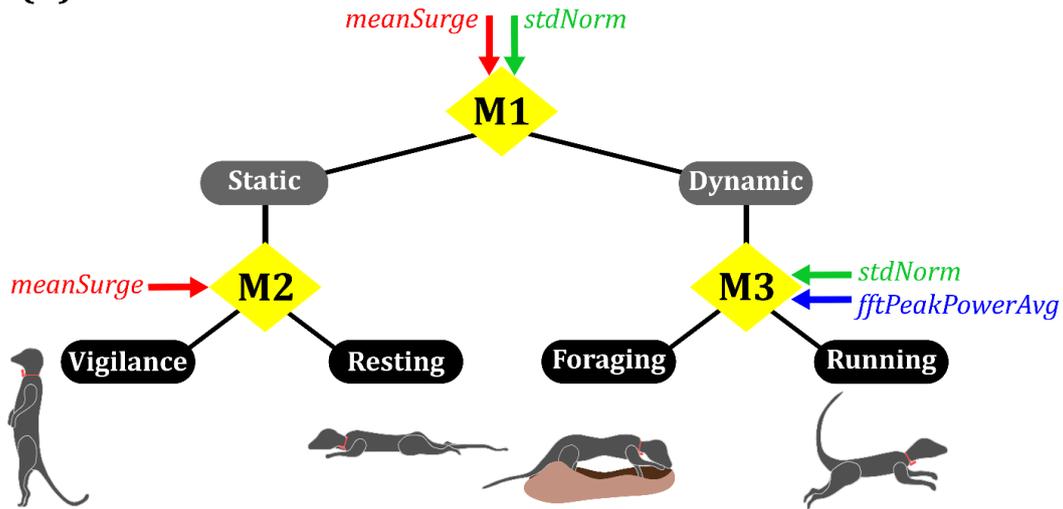


Figure 3.2. **Biomechanically informed behaviour recognition scheme.** (a) Workflow showing feature extraction from triaxial acceleration; (b) At each of the three nodes of the classification scheme, a separate machine learning algorithm (M1, M2, M3) is trained with specific biomechanical features encapsulating information on posture (*meanSurge*), movement intensity (*stdNorm*) and periodicity (*fftPeakPowerAvg*).

3.2.2.3 | Feature computation

A sliding window w of size two seconds was considered for feature computation; this provided sufficient temporal resolution of behaviour and was short enough to capture bouts of running, the behaviour with the shortest duration. Successive windows had an overlap of 50%. Windows containing transitions between different behaviours were excluded; each window thus contained acceleration data corresponding to exactly one video-labelled behaviour.

For each window w , acceleration along the surge axis ($a_{surge,w}$) only was averaged ($meanSurge_w$) and used to estimate neck inclination (eq. 1), as values along this axis are least susceptible to changes due to possible rotations of the collar around

the axis of the meerkat's cylindrical neck. For *fftPeakPowerAvg* computation (cf. Section 2.1.1 under 'Periodicity'), the parameters used were: $F_s = 100$ Hz, $U = 0.01$ Hz, $L = 10000$. The low-pass filter was a fourth-order Butterworth filter with a cut-off frequency of 10 Hz. The length of zero-padding was one second, and the Blackman-Harris windowing function was used. All feature computations were done using MATLAB R2016b. MATLAB code to compute *fftPeakPowerAvg* is supplied in Supplementary Information.

3.2.2.4 | Model validation

For EQDIST, ten synthetically equalised datasets were generated using the Imbalanced-learn (Lemaitre et al. 2017, version 0.3.0) package in Python. STRAT was implemented using Scikit-learn (Pedregosa et al. 2011). For both EQDIST and STRAT, confusion matrices obtained from each test fold were added together to produce an aggregated confusion matrix from which performance statistics were calculated. LOIO was done for individuals for which all four behaviours were recorded. Performance statistics for each individual were computed separately, and their mean and standard deviation across individuals reported.

3.2.2.5 | Alternative classification methods: benchmarking against classical machine learning approaches

To benchmark our hybrid model's results against those obtained with alternative classification methods based entirely on classical machine learning, four commonly employed algorithms were considered: Naïve-Bayes (NB), K-Nearest Neighbours (KNN, with $K = 5$), Random Forest (RF, with 10 trees), and Support Vector Machine (SVM) with a linear kernel. Features presented in a recent review of animal behaviour recognition (Nathan et al. 2012) were considered as input to these machine learning algorithms. They considered statistical features – mean, standard deviation, skewness, kurtosis, maximum and minimum – computed from data from each accelerometer axis as well as the acceleration norm. For consistency, we retained only those features that either used only the surge axis or all three axes together – 16 such features were identified. Further, to meaningfully compare results

with our three-feature hybrid model, we selected the top three features from this set of 16, thereby ensuring that both approaches would have the same complexity in terms of feature-space dimensionality. Further details on feature selection can be found in Appendix S3.2.

3.3 | Results

3.3.1 | Observed Behaviour

A total of 105,604 two-second bouts of video-labelled behaviour were collected. 12.3% of these bouts contained transitions from one observed behaviour to another, in 6.2% of them the animal was not caught on camera, and 3.3% contained social behaviour such as grooming (cf. Section 2.2.2) – these bouts were excluded from the dataset. The remaining 82550 bouts corresponded to the four behaviours of interest, as presented in Table 3.1. Foraging (56.2% of retained bouts) and vigilance (38.2%) were the most common behaviours, whereas running (1%) was the rarest. Typical signals recorded for each behaviour are shown in Figure 3.3.

3.3.2 | Performance Evaluation

Out of the 64 possible combinations for the M1-M2-M3 hybrid model, we found that SVM-SVM-SVM performed the best across all three cross-validation methods. The linear-kernel SVM not only fully automated the search for robust feature-value thresholds, but also yielded simple linear decision boundaries, thereby rendering classification rules transparent and intuitive, as shown in Figure 3.4. The next three sub-sections give details on the performance of the SVM-SVM-SVM hybrid model for each cross-validation method tested, and benchmark them against results obtained with classical machine learning using the same number of features. Results obtained with classical machine learning using all 16 features are provided in Appendix S3.3.

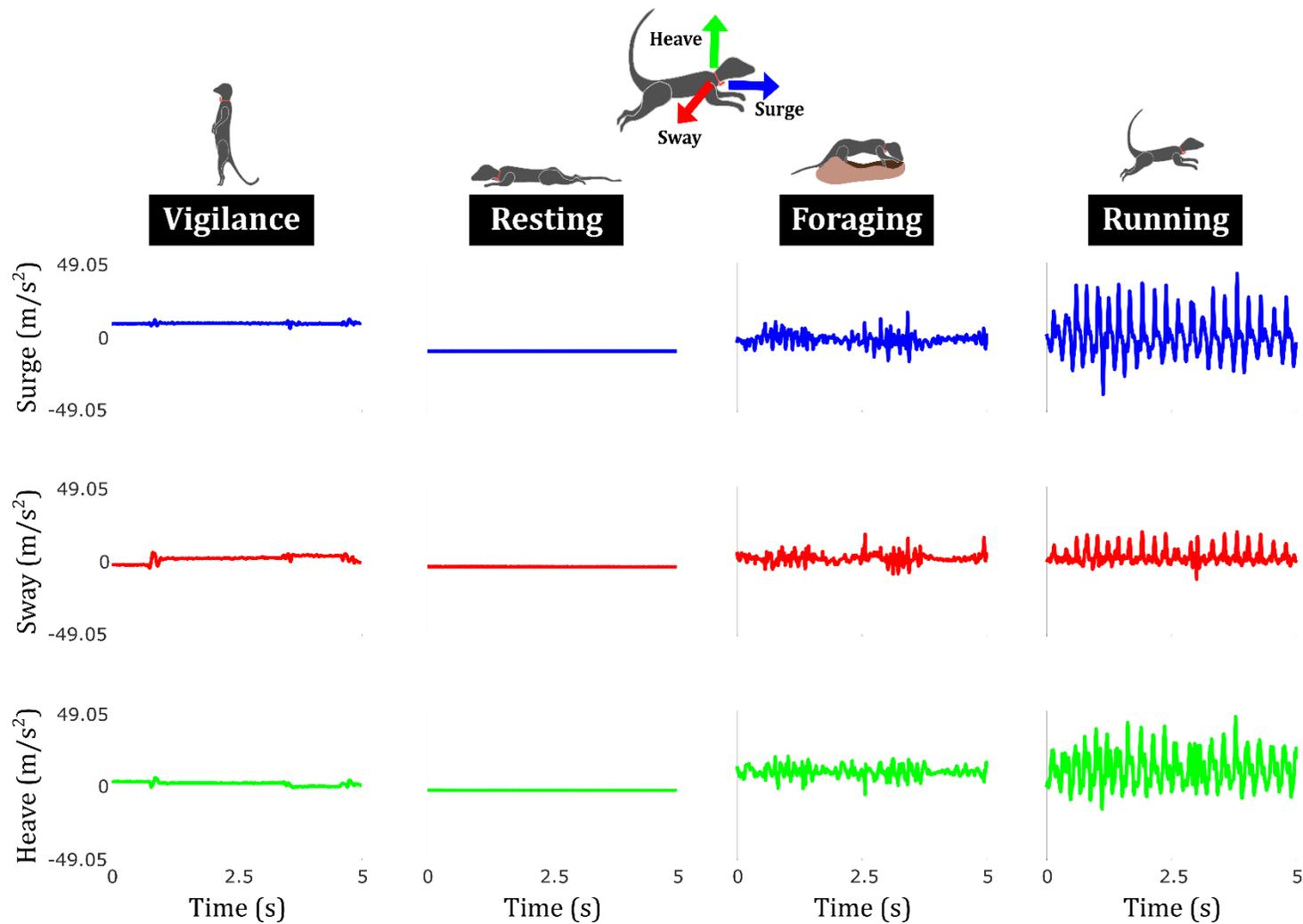


Figure 3.3. **Five-second snapshots of the four behaviours of interest for a typical individual (dataset #1).** During vigilance, the meerkat is still; short perturbations in the signal occur when the head turns to scan the surroundings. During resting, the meerkat remains still and the surge signal has a different intercept compared to that during vigilance. Foraging produces erratically varying signals due to site-dependent digging and manoeuvring. Running is a fast, rhythmic activity marked by a highly periodic signal.

Recording Session Number	Vigilance	Resting	Foraging	Running	Bouts per Recording Session
1	4594	2114	1562	69	8339
2	3896	120	5315	29	9360
3	1453	0	6278	38	7769
4	5221	0	2823	161	8205
5	1890	0	6134	169	8193
6	1639	744	4438	98	6919
7	4785	156	3498	40	8479
8	71	0	4841	20	4932
9	4283	0	1713	43	6039
10	1906	0	4407	84	6397
11	1782	661	5398	77	7918
Bouts per Activity	31520	3795	46407	828	82550 (total bouts)

Table 3.1. **Summary of data collected.** Data were collected on ten unique individuals; data from recording sessions #4 and #7 were collected on the same individual. A bout refers to a window w of two seconds containing one video-labelled behaviour.

3.3.2.1 | EQDIST

Results for EQDIST in (Table 3.2) show that for behaviour-specific metrics, the best hybrid model (SVM-SVM-SVM) performed better in terms of sensitivity (average 3.5% across the four behaviours), precision (average 3.3%) and specificity (average 1.1%) compared to the best machine learning model (K-Nearest Neighbours).

3.3.2.2 | STRAT

Results for STRAT in (Table 3.3) show that for behaviour-specific metrics, the best hybrid model (SVM-SVM-SVM) performed better in terms of sensitivity (average 5.1% across the four behaviours), precision (average 3.1%) and specificity (average 0.7%) compared to the best machine learning model (K-Nearest Neighbours).

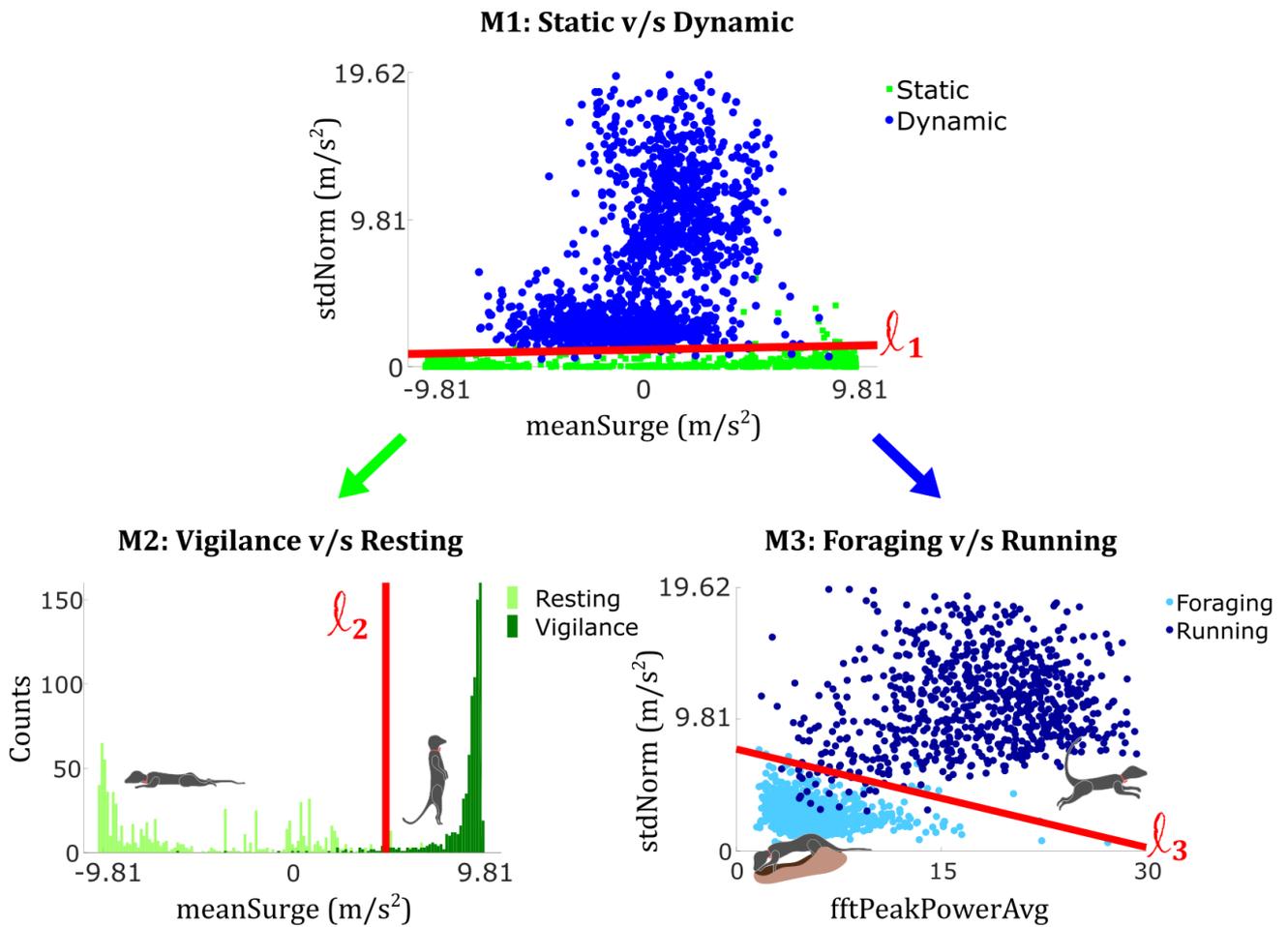


Figure 3.4. **Transparent classification.** Linear decision boundaries (l_1 , l_2 and l_3) found by the SVM-SVM-SVM hybrid model (for data used during EQDIST) render the classification scheme transparent and physically intuitive.

3.3.2.3 | LOIO

For LOIO, data from recording sessions numbered 3, 4, 5, 8, 9 and 10 (see Table 3.1) were discarded since they did not contain any resting behaviour. The mean and standard deviation of all performance metrics for LOIO with data from the five retained recording sessions (corresponding to five unique individuals) are shown in (Table 3.4). For behaviour-specific metrics, the best hybrid model (SVM-SVM-SVM) performed better in terms of mean sensitivity (average 1.1% across the four

behaviours), mean precision (average 4.3%) and mean specificity (average 1.3%) compared to the best machine learning model (Random Forest).

3.4 | Discussion

We presented a physically intuitive behaviour recognition framework based on a hybrid model that combines movement biomechanics and robust machine learning. We showed that with our biomechanically defined node-based hierarchical classification framework, acceleration-derived features summarising movement biomechanics could be used in conjunction with existing machine learning algorithms to recognise behaviour from tri-axial acceleration data. Using data collected on ten wild meerkats, we demonstrated the efficacy of our hybrid model in scenarios where one or more behaviours are rarer or under-represented in the dataset compared to others, and when there is inter-individual variability. We showed that our hybrid model's predictions were more accurate than those obtained with alternative classification methods based on classical machine learning using the same number of features.

Our results shed light on the complex interplay of several factors involved in behaviour recognition: feature selection, class imbalance, class separability, model selection, and persistent misclassification. While the feature selection methods used in classical machine learning found measures of movement intensity (*stdNorm*, *stdSurge*; Appendix S3.2), they failed to select a measure of posture. This proved to be especially problematic for distinguishing the two static behaviours, vigilance and resting, since they primarily differ in posture. Consistently poorer resting and vigilance detection performance resulted when a measure of posture was excluded (classical machine learning) compared to when it was included (*meanSurge* in the hybrid model) even in EQDIST, where matters are not yet complicated by class imbalance.

Approach	Model	Vigilance			Resting			Foraging			Running			Overall Accuracy (%)
		Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
Hybrid	SVM-SVM-SVM	95.2	98.2	94.7	93.7	98.6	95.8	97	98	94.3	96.7	99.3	97.9	95.7
Classical Machine Learning	NB	43.5	96.7	81.6	88.9	82.3	62.6	95.7	97.5	92.6	97.8	98.8	96.5	81.5
	KNN	91.3	94.5	84.6	82.7	97.9	93	96.9	97.9	93.9	97.9	99.3	97.9	92.2
	RF	90.5	94.4	84.3	83.3	97.5	91.8	96.7	98	94.2	97.5	99.4	98.2	92
	SVM	53.6	96.2	82.3	87.6	85.7	67.1	97.3	97.5	92.8	97.7	99.4	98.2	84

Table 3.2. **EQDIST cross-validation results.** The performance of the SVM-SVM-SVM hybrid model is benchmarked against that of classical machine learning. NB: Naïve-Bayes; KNN: K-Nearest Neighbours; RF: Random Forest; SVM: Support Vector Machine.

Approach	Model	Vigilance			Resting			Foraging			Running			Overall Accuracy (%)
		Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
Hybrid	SVM-SVM-SVM	97.1	98.8	98.1	85	99.4	87.1	99.3	97.8	98.3	85.9	99.9	92.1	97.7
Classical Machine Learning	NB	96.4	92	88.2	0	100	0	98	96.6	97.3	94.7	99.3	58.1	92.9
	KNN	95.8	96.6	94.5	64.1	99.3	81.4	98.8	97.4	98	88.4	99.9	89.4	96
	RF	95.7	96.5	94.4	64.7	99.2	80.2	98.8	97.5	98.1	86.7	99.9	90.7	95.9
	SVM	96.8	92.1	88.4	0	100	0	99.2	96.6	97.4	82.6	99.9	91.3	93.6

Table 3.3. **STRAT cross-validation results.** The performance of the SVM-SVM-SVM hybrid model is benchmarked against that of classical machine learning. NB: Naïve-Bayes; KNN: K-Nearest Neighbours; RF: Random Forest; SVM: Support Vector Machine.

Approach	Model	Vigilance			Resting			Foraging			Running			Overall Accuracy (%)
		Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
Hybrid	SVM-SVM-SVM	95.8 ± 2.8	98.4 ± 1.2	96.4 ± 4.5	71.4 ± 23.6	98.9 ± 1.2	81.1 ± 28.0	98.8 ± 1.0	97.4 ± 1.5	95.3 ± 7.0	86.3 ± 13.2	99.9 ± 0.1	89.1 ± 11.1	96.5 ± 1.8
Classical Machine Learning	NB	76.8 ± 22.4	82.6 ± 22.3	78.7 ± 14.4	31.7 ± 45.5	90.9 ± 12.7	2.1 ± 2.9	97.3 ± 1.0	96.1 ± 1.8	94.4 ± 6.6	98.6 ± 2.0	99.3 ± 0.8	61.9 ± 20.6	78.9 ± 8.2
	KNN	92.5 ± 2.6	93.1 ± 6.2	88.8 ± 8.6	62.6 ± 18.4	98.3 ± 0.8	63.4 ± 29.8	98.2 ± 0.9	97.4 ± 1.1	96.2 ± 4.5	90.2 ± 12.5	99.9 ± 0.2	90.7 ± 10.1	93.1 ± 2.8
	RF	92.7 ± 3.2	93.7 ± 5.5	89.7 ± 7.3	68.5 ± 17.8	98.4 ± 1.0	66.9 ± 28.4	98.0 ± 1.0	97.6 ± 1.1	96.6 ± 3.8	88.8 ± 12.0	99.9 ± 0.2	91.4 ± 11.7	93.6 ± 2.4
	SVM	95.5 ± 2.6	82.0 ± 21.8	79.3 ± 14.8	0	100	0	98.7 ± 1.0	96.1 ± 1.8	94.5 ± 6.5	84.8 ± 12.0	99.9 ± 0.1	90.8 ± 9.4	88.2 ± 10.5

Table 3.4. **LOIO cross-validation results.** The performance of the SVM-SVM-SVM hybrid model is benchmarked against classical machine learning. Performance metrics were calculated separately for each test individual and their mean and standard deviation across test individuals are shown here. NB: Naïve-Bayes; KNN: K-Nearest Neighbours; RF: Random Forest; SVM: Support Vector Machine.

The effect of not including a measure of posture gets amplified when naturally occurring class imbalance is introduced in STRAT: the classical machine learning models NB and SVM completely miss the much rarer resting behaviour (vigilance bouts outnumber resting bouts 8.3:1), whereas for KNN and RF, resting detection sensitivity plummets to below 65% and precision to below 82%. Model response to class imbalance may vary unpredictably: classical machine learning with KNN and RF yields fair performance across EQDIST, STRAT, and LOIO; SVM, on the other hand, first recognises resting during EQDIST (albeit poorly) despite the absence of a feature describing posture, then completely misses resting behaviour during STRAT and LOIO (even though overall model accuracy still remains high: 93.6% and 88.2% respectively), and then performs well when all 16 features are included (Appendix S3.3). This indicates the difficulty of generalising model behaviour across datasets when feature choice is left to an automatic algorithm. However, the hybrid model, where the SVM algorithm was only used to find feature-value thresholds within a biomechanically pre-defined tree-like classification structure with task-specific features (i.e. biomechanically appropriate features for each node in the classification tree), consistently performs well across EQDIST, STRAT, and LOIO. Class imbalance may be less problematic when class separability is high. Running, despite being the rarest class in our dataset (foraging instances outnumber running instances 56:1) is recognised fairly accurately across EQDIST, STRAT, and LOIO for the KNN, RF, and SVM classical machine learning models, and the hybrid model. This might be because measures of intensity (*stdNorm*, *stdSurge*) were particularly effective at separating foraging and running: this can be seen from Figure 3.4, where the magnitude of <1 of the slope of the decision boundary ℓ_3 implies that the feature on the vertical axis (*stdNorm*) is more discriminating than the one on the horizontal axis (*fftPeakPowerAvg*).

In addition to exponentially greater model complexity, perhaps the greatest disadvantage of increasing the number of features in classical machine learning is the loss of ability to understand why misclassifications persist. The performance of classical machine learning becomes comparable to that of the hybrid model when all 16 features are used (Appendix S3.3), yet even with a five-fold increase in the number of features (from 3 to 16), resting detection performance still remains poor. Examination of the classification rules to understand the source of misclassification

is rendered unmanageable by the fact that, for 16 features, 16-dimensional space will need to be analysed. With the hybrid model, however, due to the hierarchical nature of the classification scheme, it is easy to pinpoint where and why resting misclassification occurs. The rarer resting behaviour is clubbed with a frequent behaviour, vigilance, into the static behavioural category, and the rarer running behaviour is clubbed with a frequent behaviour, foraging, (foraging bouts outnumber running bouts 56:1) into the dynamic category. Thereafter, since the static and dynamic behavioural categories are separated accurately (Table S3.7), it is easy to see that the primary deterioration of resting recognition accuracy must occur at the M2 node. This node employs an SVM which uses only one feature as input, *meanSurge* (measure of posture) – this thus indicated that there may have been limitations to our hypothesis about posture during resting and/or vigilance. Indeed, re-consulting the groundtruthing video data, we discovered that the major source of erroneous resting detection was the disproportionately large contribution of an additional unexpected curled-up resting posture (different from the belly-flat position typically observed) of a single individual (see Appendix S3.4). Thus, new insight into the groundtruthing data itself was obtained because of the interpretability offered by our hybrid model; this is in contrast to machine learning approaches, which have to rely completely on groundtruthing data.

Our behaviour recognition framework offers two other key advantages over existing methods. Firstly, in our hybrid model, robust machine learning algorithms were tasked with searching for feature-value thresholds, thus making the mechanism of decision boundary selection automatic, general, clear, and easy to implement. In the tree-based classification model presented by McClune et al. 2014, analyses were based on data from a single, tame individual, and it was not clear how their iterative feature-threshold selection method could be extended to data from more individuals. Secondly, our classification scheme has the potential to enable meaningful model sharing across studies by offering the added advantage of being adaptable to desired levels of behavioural resolution. For studies where coarse behavioural resolution is sufficient, our hybrid model can be used to simply determine when the animal was static or dynamic. For studies requiring higher behavioural resolution, our scheme may be used as a template upon which new nodes, possibly using additional biomechanically derived features, may be added further down the tree to

accommodate new behaviours when needed. For the meerkat ethogram chosen for this study, three biomechanically interpretable features proved to be enough but, for instance, if one were interested in characterising meerkat foraging effort, one could add an additional node below ‘foraging’ and split it into ‘digging’ and ‘non-digging’ through a peak-detection-based metric indicating front-paw swipes made during digging. In a classical machine learning approach, the whole model would need to be rebuilt from scratch if new behaviours were to be added; in our approach, adding higher behavioural resolution to a given ‘coarser’ model will not impact the existing model parameters. This could enable separate studies on the same species to add upon a single model until the required behavioural resolution has been achieved. This high-resolution model may then be made available for future studies on that species.

Choosing appropriate sensor parameters may be a crucial component of achieving accurate behaviour classification. Even though Gao et al. 2013 followed a hierarchical classification scheme employing SVMs, the web-based system they developed limited the input sampling frequency to only 1 Hz which, they found for some species, was insufficient to extract meaningful information from the frequency-domain features they used. In our meerkat study, requisite signal processing techniques employed to compute movement periodicity through *fftPeakPowerAvg* might have played an important role in successfully distinguishing meerkat running from foraging despite high class imbalance. This was due, in part, to the choice of a sufficiently high sampling frequency of 100 Hz. We found that meerkat running had a characteristic frequency of around 4 Hz, and a good rule-of-thumb is to oversample by about 20 times when using a noisy sensor (Boyd et al. 2010), although the sampling frequency could theoretically be reduced to the Nyquist limit of twice the maximum frequency of interest. Finally, while *stdNorm* and *fftPeakPowerAvg*, used in foraging versus running classification, make no assumptions about sensor orientation with respect to the animal, *meanSurge*, used in vigilance versus resting classification, assumes knowledge of sensor placement around the meerkat’s neck. Higher running classification accuracy (compared to that for resting) might thus indicate that if the features used are independent of sensor orientation with respect to the animal, classification may be more robust for some species, since sensor fixation will inevitably be slightly different across individuals or species (e.g. collars in mammals and back-packs in birds). Studies on energy expenditure (e.g. Gleiss et al.

2011) also suggested that when accelerometers cannot be accurately placed on the animal, using information from all axes together (Vectorial Dynamic Body Acceleration) may perform better than treating each axis independently (Overall Dynamic Body Acceleration).

Through the use of general biomechanical principles characterising animal movement, our conceptually simple, robust classification method may be applicable across a range of species, with different behaviour labels depending on the species' specific context. The biomechanically defined node-based hierarchical format permitting model adaptation to coarser or finer behavioural resolution makes it apt for meaningful model sharing between studies on a given species. The advantage and novelty of our method is that it allows a high classification performance and, at the same time, a physical and biomechanical interpretation of the classification outcomes. The incorporation of common machine learning algorithms found in all popular, existing packages in Python, MatLab and R makes this method simple and accessible.

Appendices

Appendix S3.1: Metrics for performance evaluation of classification models

We used a confusion matrix $[N_{jk}]$ to summarise model performance, where $j, k \in \{1: \text{vigilance}, 2: \text{resting}, 3: \text{foraging}, 4: \text{running}\}$ and N_{jk} denotes the number of samples of behaviour ‘ j ’ that were predicted by the model to be behaviour ‘ k ’. From this confusion matrix, we computed three behaviour-specific metrics - sensitivity (often referred to as recall in the literature), precision, and specificity, as well as overall model accuracy. These measures are computed as shown in Table S3.1.

Performance statistic	Expression	Interpretation
Sensitivity (or Recall)	$S_k = \frac{N_{kk}}{\sum_{i=1}^4 N_{ki}}$	How often behaviour k is correctly identified by the model
Precision	$P_k = \frac{N_{kk}}{\sum_{i=1}^4 N_{ik}}$	How often the model is right when it predicts behaviour k
Specificity	$Sp_k = \frac{\sum_{i \neq k} \sum_{j \neq k} N_{ij}}{\sum_{i \neq k} \sum_{j=1}^4 N_{ij}}$	How often the absence of behaviour k is correctly identified by the model
Overall accuracy	$A = \frac{\sum_{i=1}^4 N_{ii}}{\sum_{i=1}^4 \sum_{j=1}^4 N_{ij}}$	How often the model is right, across all behaviours

Table S3.1. **Measures to quantify model performance.** Sensitivity (or recall), precision, and specificity are behaviour-specific metrics while overall accuracy summarises overall model performance across the four behaviours

Appendix S3.2: Top 3 feature selection from Nathan et al., 2012 feature set

From the set of all features presented by Nathan et al., 2012, we first left out features that assumed knowledge of the sway and heave axes (since these axes are highly susceptible to changes in collar orientation around the animal's neck), and calculated features only for the surge axis and acceleration norm. These features are listed in Table S3.2.

S.No.	Feature name	Feature description
1.	<i>meanSurge</i>	Mean of surge acceleration
2.	<i>meanNorm</i>	Mean of acceleration norm
3.	<i>stdSurge</i>	Standard deviation of surge acceleration
4.	<i>stdNorm</i>	Standard deviation of acceleration norm
5.	<i>skewnessSurge</i>	Skewness of surge acceleration
6.	<i>skewnessNorm</i>	Skewness of acceleration norm
7.	<i>kurtosisSurge</i>	Kurtosis of surge acceleration
8.	<i>kurtosisNorm</i>	Kurtosis of acceleration norm
9.	<i>maxSurge</i>	Maximum value of surge acceleration
10.	<i>maxNorm</i>	Maximum value of acceleration norm
11.	<i>minSurge</i>	Minimum value of surge acceleration
12.	<i>minNorm</i>	Minimum value of acceleration norm
13.	<i>autocorrSurge</i>	Autocorrelation (for a displacement of one measurement) of surge acceleration
14.	<i>autocorrNorm</i>	Autocorrelation (for a displacement of one measurement) of acceleration norm
15.	<i>trendSurge</i>	The coefficient for a linear regression through the data for surge acceleration
16.	<i>trendNorm</i>	The coefficient for a linear regression through the data for acceleration norm

Table S3.2. Features retained from feature set presented by Nathan et al., 2012.

Then, five different feature selection methods were run to select the top 3 features in this feature set. These methods were:

1. `f_classif` (filter method)
2. `mutual_info_classif` (filter method)
3. RFE (recursive feature elimination, wrapper method)
4. `Linear_model.LogisticRegression` (embedded method)
5. `ExtraTreesClassifier` (feature importance with forests of trees, embedded method)

Test Name	Filter Methods		Wrapper Method	Embedded Methods	
	F_classif (order matters)	Mutual_info_classif (order matters)	RFE (order doesn't count)	LogReg (order doesn't count)	ExtraTreesClassifier (order matters)
Top5 features	<i>meanSurge</i>	<i>stdNorm</i>	<i>meanNorm</i>	<i>meanNorm</i>	<i>minSurge</i>
	<i>minSurge</i>	<i>minSurge</i>	<i>stdSurge</i>	<i>stdSurge</i>	<i>stdSurge</i>
	<i>minNorm</i>	<i>stdSurge</i>	<i>stdNorm</i>	<i>stdNorm</i>	<i>meanSurge</i>
	<i>stdSurge</i>	<i>meanSurge</i>	<i>minNorm</i>	<i>minNorm</i>	<i>minNorm</i>
	<i>stdNorm</i>	<i>minNorm</i>	<i>autocorrSurge</i>	<i>autocorrSurge</i>	<i>stdNorm</i>

Table S3.3. Top five features found using five different feature selection techniques.

From Table S3.3, one sees that the following 3 features occur the most often, and were the ones selected for comparison with the hybrid model:

- *stdSurge* (all 5 methods)
- *stdNorm* (all 5 methods)
- *minNorm* (all 5 methods)

Appendix S3.3: Results with machine learning using all 16 features from Nathan et al., 2012

Here we provide results obtained when classical machine learning algorithms using all 16 features are tested for the three cross-validation methods: EQDIST (Table S3.4), STRAT (Table S3.5), and LOIO (Table S3.6).

Model	Vigilance			Resting			Foraging			Running			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
NB	95.2	99	96.9	95	99.1	97.4	98.5	97.3	92.5	96.9	99.7	99.1	96.4
KNN	95.4	96.6	90.3	91.7	98.6	95.6	94.6	97.8	93.4	95.6	99.5	98.6	94.3
RF	96.7	99	96.9	96.6	99.1	97.2	97.6	98.6	95.9	97.6	99.5	98.6	97.1
SVM	96.4	98.9	96.8	96.4	99	97	97.4	98.8	96.3	98.2	99.5	98.4	97.1

Table S3.4 **EQDIST cross-validation results.** Performance of classical machine learning algorithms using all 16 features in Table S3.2. NB: Naïve-Bayes; KNN: K-Nearest Neighbours; RF: Random Forest; SVM: Support Vector Machine.

Model	Vigilance			Resting			Foraging			Running			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
NB	95.5	99.5	99.2	94.1	99.1	83	99.1	97.5	98.1	93.7	99.7	74.3	97.4
KNN	97.5	98.3	97.3	83	99.6	90	99.1	98.1	98.5	84.3	99.9	91.8	97.6
RF	97.8	98.5	97.6	87.9	99.5	89.9	99.1	98.7	99	88.3	99.9	93.8	98
SVM	97.5	98.9	98.2	89.2	99.5	90.1	99.3	98.3	98.7	88.8	99.9	89.5	98

Table S3.5. **STRAT cross-validation results.** Performance of classical machine learning algorithms using all 16 features Table S3.2. NB: Naïve-Bayes; KNN: K-Nearest Neighbours; RF: Random Forest; SVM: Support Vector Machine.

Model	Vigilance			Resting			Foraging			Running			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
NB	93.6 ± 4.2	99.2 ± 0.8	98.0 ± 2.9	86.4 ± 18.8	98.6 ± 1.2	81.2 ± 1.7	98.6 ± 1.0	97.4 ± 1.5	95.5 ± 6.7	94.6 ± 6.0	99.5 ± 0.7	70.6 ± 19.3	96.3 ± 1.7
KNN	96.4 ± 2.1	97.0 ± 0.8	94.2 ± 4.6	66.8 ± 21.6	99.2 ± 0.6	80.9 ± 30.4	98.8 ± 0.5	97.8 ± 1.5	96.0 ± 6.3	83.7 ± 16.1	99.9 ± 0.1	96.8 ± 4.9	96.3 ± 1.6
RF	96.7 ± 3.7	97.7 ± 1.6	95.5 ± 5.5	71.6 ± 24.7	99.2 ± 0.9	83.4 ± 26.7	98.8 ± 1.0	98.8 ± 0.8	97.6 ± 4.0	91.0 ± 12.0	100 ± 0	98.5 ± 2.5	97.2 ± 1.2
SVM	96.1 ± 2.7	98.4 ± 0.8	96.6 ± 3.5	79.9 ± 17.9	99.1 ± 0.8	82.9 ± 25.5	99.0 ± 0.6	98.3 ± 1.2	96.6 ± 5.6	90.4 ± 13.7	99.9 ± 0.1	95.7 ± 4.2	97.2 ± 1.0

Table S3.6. **LOIO cross-validation results.** Performance of classical machine learning algorithms using all 16 features in Table S3.2.

Performance metrics were calculated separately for each test individual and their mean and standard deviation across test individuals are shown here. NB: Naïve-Bayes; KNN: K-Nearest Neighbours; RF: Random Forest; SVM: Support Vector Machine.

Appendix S3.4: Results with the SVM-SVM-SVM hybrid model during LOIO

Though resting behaviour performance during LOIO was the poorest of the four behaviours, a sensitivity of 85.7% for resting detection was observed with the SVM-SVM-SVM hybrid model when we aggregated confusion matrices for all five individuals. Table 1 (in the main text) shows that more than half of all resting bouts came from a single individual (dataset #1), and we did obtain good resting detection (90.7% sensitivity, 90.8% precision) for that particular individual. In fact, resting detection was poor only for the two individuals (datasets #2 and #7) for which few resting bouts were collected. Further investigation revealed that meerkats rested in different postures – some individuals curled up into a ball during resting rather than lying flat on the ground. This, and slight collar fixation variations across individuals, led to varying mean surge acceleration values that compromised resting detection.

The aggregate confusion matrix, and the individual-wise behaviour recognition performance and confusion matrices obtained with the SVM-SVM-SVM hybrid model during LOIO are provided in Table S3.7, and Table S3.8 and Table S3.9 respectively.

		<i>Predicted</i>			
		Vigilance	Resting	Foraging	Running
<i>Actual</i>	<i>Vigilance</i>	16069	279	340	8
	<i>Resting</i>	337	3252	203	3
	<i>Foraging</i>	80	102	20002	27
	<i>Running</i>	0	0	47	266

Table S3.7. **Aggregate confusion matrix obtained with the SVM-SVM-SVM hybrid model for LOIO.** Predictions are arranged along the columns and true labels along the rows; correct predictions are shown in bold.

Individual No.	Vigilance			Resting			Foraging			Running			Overall accuracy (%)
	Sen (%)	Spec (%)	Prec (%)	Sen (%)	Spec (%)	Prec (%)	Sen (%)	Spec (%)	Prec (%)	Sen (%)	Spec (%)	Prec (%)	
1	93.8	99.1	99.2	90.7	96.9	90.8	97.1	95.4	82.9	72.5	100	92.6	93.5
2	98.4	98.5	97.9	36.7	99.0	31.4	98.8	99.4	99.5	93.1	100	96.4	97.8
6	92.1	99.3	97.5	94.5	98.8	90.1	99.4	97.3	98.5	96.9	99.7	84.1	97.1
7	98.1	98.9	99.1	60.9	99.9	95.0	99.6	97.9	97.1	97.5	99.8	72.2	98
11	96.6	96.4	88.5	74.4	99.9	98.2	98.9	96.9	98.6	71.4	100	100	96.1

Table S3.8. **Individual-wise behaviour recognition performance with the SVM-SVM-SVM hybrid model for LOIO.** Note that all five recording sessions listed here were done on distinct individuals.

Recording Session No.	Confusion Matrix					
		<i>Predicted</i>				
		Vig.	Rest	Frg.	Run	
1	<i>Actual</i>	Vig.	4310	167	115	2
		Rest	16	1918	179	1
		Frg.	18	27	1516	1
		Run	0	0	19	50
2	<i>Actual</i>	Vig.	3834	37	24	1
		Rest	76	44	0	0
		Frg.	5	59	5251	0
		Run	0	0	2	27
6	<i>Actual</i>	Vig.	1509	69	61	0
		Rest	37	703	4	0
		Frg.	1	8	4411	18
		Run	0	0	3	95
7	<i>Actual</i>	Vig.	4694	0	86	5
		Rest	41	95	18	2
		Frg.	0	5	3485	8
		Run	0	0	1	39
11	<i>Actual</i>	Vig.	1722	6	54	0
		Rest	167	492	2	0
		Frg.	56	3	5339	0
		Run	0	0	22	55

Table S3.9. **Individual-wise confusion matrices with the SVM-SVM-SVM hybrid model.**

Notice that for recording session #2, #6, #7 and #11, the major source of erroneous resting detection is due to the misclassification of vigilance as resting. Note that all five recording sessions listed here were done on distinct individuals.

References

- Bidder, O. R., Campbell, H. A., Gómez-Laich, A., Urgé, P., Walker, J., Cai, Y., ... & Wilson, R. P. (2014). Love thy neighbour: automatic animal behavioural classification of acceleration data using the k-nearest neighbour algorithm. *PloS one*, 9(2), e88609.
- Boyd, J., Sundaram, H., & Shrivastava, A. (2010, March). Power-accuracy tradeoffs in human activity transition detection. In *Proceedings of the Conference on Design, Automation and Test in Europe* (pp. 1524-1529). European Design and Automation Association.
- Chakravarty P., Cozzi G., Ozgul A., Aminian K. (2019). Data from: A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods in Ecology and Evolution*. Dryad Digital Repository: doi:10.5061/dryad.7q294p8.
- Clutton-Brock, T. H., Gaynor, D., McIlrath, G. M. et al. (1999). Predation, group size and mortality in a cooperative mongoose, *Suricata suricatta*. *Journal of Animal Ecology*, 68(4), 672-683.
- Cozzi, G., Broekhuis, F., McNutt, J. W., Turnbull, L. A., Macdonald, D. W., & Schmid, B. (2012). Fear of the dark or dinner by moonlight? Reduced temporal partitioning among Africa's large carnivores. *Ecology*, 93(12), 2590-2599.
- Ferraris, F., Grimaldi, U., & Parvis, M. (1995). Procedure for effortless in-field calibration of three-axial rate gyro and accelerometers. *Sensors and Materials*, 7(5), 311-330.
- Gao, L., Campbell, H. A., Bidder, O. R., & Hunter, J. (2013). A Web-based semantic tagging and activity recognition system for species' accelerometry data. *Ecological Informatics*, 13, 47-56.
- Gleiss, A. C., Wilson, R. P., & Shepard, E. L. (2011). Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure. *Methods in Ecology and Evolution*, 2(1), 23-33.
- Golabek, K. A., Jordan, N. R., & Clutton-Brock, T. H. (2008). Radiocollars do not affect the survival or foraging behaviour of wild meerkats. *Journal of Zoology*, 274(3), 248-253.
- Goldbogen, J. A., Calambokidis, J., Oleson, E., Potvin, J., Pyenson, N. D., Schorr, G., & Shadwick, R. E. (2011). Mechanics, hydrodynamics and energetics of

- blue whale lunge feeding: efficiency dependence on krill density. *Journal of Experimental Biology*, 214(1), 131-146.
- Grünewälder, S., Broekhuis, F., Macdonald, D. W., Wilson, A. M., McNutt, J. W., Shawe-Taylor, J., & Hailes, S. (2012). Movement activity based classification of animal behaviour with an application to data from cheetah (*Acinonyx jubatus*). *PLoS One*, 7(11), e49120.
- Hammond, T. T., Springthorpe, D., Walsh, R. E., & Berg-Kirkpatrick, T. (2016). Using accelerometers to remotely and automatically characterize behavior in small animals. *Journal of Experimental Biology*, 219(11), 1618-1624.
- He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, 21(9), 1263-1284.
- Hodge, S. J., Manica, A., Flower, T. P., & Clutton-Brock, T. H. (2008). Determinants of reproductive success in dominant female meerkats. *Journal of Animal Ecology*, 77(1), 92-102.
- Kawabe, R., Kawano, T., Nakano, N., Yamashita, N., Hiraishi, T., & Naito, Y. (2003). Simultaneous measurement of swimming speed and tail beat activity of free-swimming rainbow trout *Oncorhynchus mykiss* using an acceleration data-logger. *Fisheries science*, 69(5), 959-965.
- Laich, A. G., Wilson, R. P., Quintana, F., & Shepard, E. L. (2008). Identification of imperial cormorant *Phalacrocorax atriceps* behaviour using accelerometers. *Endangered species research*, 10, 29-37.
- Lemaitre, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. *Journal of Machine Learning Research*, 18(17), 1-5.
- McClune, D. W., Marks, N. J., Wilson, R. P., Houghton, J. D., Montgomery, I. W., McGowan, N. E., ... & Scantlebury, M. (2014). Tri-axial accelerometers quantify behaviour in the Eurasian badger (*Meles meles*): towards an automated interpretation of field data. *Animal Biotelemetry*, 2(1), 5.
- McGregor, S. J., Busa, M. A., Yaggie, J. A., & Bollt, E. M. (2009). High resolution MEMS accelerometers to estimate VO₂ and compare running mechanics between highly trained inter-collegiate and untrained runners. *PloS one*, 4(10), e7355.

- Naito, Y., Bornemann, H., Takahashi, A., McIntyre, T., & Plötz, J. (2010). Fine-scale feeding behavior of Weddell seals revealed by a mandible accelerometer. *Polar Science*, 4(2), 309-316.
- 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.
- Pedregosa et al. (2011), Scikit-learn: Machine Learning in Python, JMLR 12, pp. 2825-2830.
- Resheff, Y. S., Rotics, S., Harel, R., Spiegel, O., & Nathan, R. (2014). AcceleRater: a web application for supervised learning of behavioral modes from acceleration measurements. *Movement ecology*, 2(1), 27.
- 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.
- Ropert-Coudert, Y., Kato, A., Wilson, R. P., & Cannell, B. (2006). Foraging strategies and prey encounter rate of free-ranging Little Penguins. *Marine Biology*, 149(2), 139.
- Sellers, W. I., Varley, J. S., & Waters, S. S. (1998). Remote monitoring of locomotion using accelerometers: a pilot study. *Folia Primatologica*, 69(Suppl. 1), 82-85.
- Shepard, E. L., Wilson, R. P., Quintana, F., Laich, A. G., Liebsch, N., Albareda, D. A., ... & Newman, C. (2008). Identification of animal movement patterns using tri-axial accelerometry. *Endangered Species Research*, 10, 47-60.
- Soltis, J., Wilson, R. P., Douglas-Hamilton, I., Vollrath, F., King, L. E., & Savage, A. (2012). Accelerometers in collars identify behavioral states in captive African elephants *Loxodonta africana*. *Endangered Species Research*, 18(3), 255-263.
- 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), 2.
- Watanabe, S., Izawa, M., Kato, A., Ropert-Coudert, Y., & Naito, Y. (2005). A new technique for monitoring the detailed behaviour of terrestrial animals: a case

study with the domestic cat. *Applied Animal Behaviour Science*, 94(1), 117-131.

- Watanabe, Y. Y., & Takahashi, A. (2013). Linking animal-borne video to accelerometers reveals prey capture variability. *Proceedings of the National Academy of Sciences*, 110(6), 2199-2204.
- Yoda, K., Naito, Y., Sato, K., Takahashi, A., Nishikawa, J., Ropert-Coudert, Y., ... & Le Maho, Y. (2001). A new technique for monitoring the behaviour of free-ranging Adelie penguins. *Journal of Experimental Biology*, 204(4), 685-690.
- Wilson, R. P., White, C. R., Quintana, F., Halsey, L. G., Liebsch, N., Martin, G. R., & Butler, P. J. (2006). Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *Journal of Animal Ecology*, 75(5), 1081-1090.

CHAPTER 4

Coarse-Scale Behaviour Recognition Using Magnetometers⁴

⁴ This chapter was published as, “Chakravarty, P., Maalberg, M., Cozzi, G., Ozgul, A., & Aminian, K. (2019). Behavioural compass: animal behaviour recognition using magnetometers. *Movement Ecology*, 7(1), 28.”

Abstract

1. Animal-borne data loggers today often house several sensors recording simultaneously at high frequency. This offers opportunities to gain fine-scale insights into behaviour from individual-sensor as well as integrated multi-sensor data. In the context of behaviour recognition, even though accelerometers have been used extensively, magnetometers have recently been shown to detect specific behaviours that accelerometers miss. The prevalent constraint of limited training data necessitates the importance of identifying behaviours with high robustness to data from new individuals, and may require fusing data from both these sensors. However, no study yet has developed an end-to-end approach to recognise common animal behaviours such as foraging, locomotion, and resting from magnetometer data in a common classification framework capable of accommodating and comparing data from both sensors.
2. We address this by first leveraging magnetometers' similarity to accelerometers to develop biomechanical descriptors of movement: we use the static component given by sensor tilt with respect to Earth's local magnetic field to estimate posture, and the dynamic component given by change in sensor tilt with time to characterise movement intensity and periodicity. We use these descriptors within an existing hybrid scheme that combines biomechanics and machine learning to recognise behaviour. We showcase the utility of our method on triaxial magnetometer data collected on ten wild Kalahari meerkats (*Suricata suricatta*), with annotated video recordings of each individual serving as groundtruth. Finally, we compare our results with accelerometer-based behaviour recognition.
3. The overall recognition accuracy of >94% obtained with magnetometer data was found to be comparable to that achieved using accelerometer data. Interestingly, higher robustness to inter-individual variability in dynamic behaviour was achieved with the magnetometer, while the accelerometer was better at estimating posture.
4. Magnetometers were found to accurately identify common behaviours, and were particularly robust to dynamic behaviour recognition. The use of biomechanical considerations to summarise magnetometer data makes the

hybrid scheme capable of accommodating data from either or both sensors within the same framework according to each sensor's strengths. This provides future studies with a method to assess the added benefit of using magnetometers for behaviour recognition.

4.1 | Background

Behaviour is a central component of any animal's life and the result of important biotic and abiotic interactions. Its accurate description is therefore crucial for a full appreciation of an animal's biology. Small, light-weight animal-borne data loggers have proved to be indispensable as they bypass the logistical difficulties of directly observing animals and enable data to be collected on an animal as it goes about its daily life in its natural environment [1]. Data loggers today can often house several sensors (e.g. [2], [3]), each measuring different physical quantities such as acceleration, magnetic field intensity, angular velocity, light level, and depth. Simultaneously recorded high-frequency multi-sensor data offer the opportunity to gain fine-scale insights into behaviour by leveraging information not only from individual data streams, but also by fusing data from multiple sensors.

In the context of animal behaviour recognition, triaxial accelerometers [4] and magnetometers [5] have both been used to identify movement patterns in animals. Though accelerometers have by far been used more extensively (e.g. [6], [7], [8], [9], [10]), it has recently been shown that magnetometers can better resolve certain low-acceleration behaviours of biological importance, e.g. thermal soaring in Andean condors (*Vultur gryphus*) [11]. In fact, a recent comparison of accelerometers and magnetometers has demonstrated that there can be quantifiably large differences in recognition capability between the two sensors for certain specific behaviours [5]. Future behaviour recognition algorithms may thus seek to leverage the complementarity of these two sensors by fusing data from both sensors within a single classification framework. However, it is not known how recognition capability differs between the two sensors for the case of common animal behaviours such as foraging, locomotion, and resting.

One of the reasons for the success of accelerometers in recognising animal behaviour may be their ability to measure both static tilt with respect to Earth's gravity vector as well as dynamic acceleration resulting from animal motion. Despite the numerous advantages of the accelerometer, however, the sensor has some inherent limitations that may render it unsuitable for use in certain situations. Firstly, during dynamic movements, the accelerometer is sensitive to both body segment tilt and dynamic acceleration due to motion. Dynamic acceleration interferes with the

change of tilt, and the two cannot be separated. In extreme cases, such as when an animal is ‘pulling g’ [5] or in freefall, the accelerometer cannot be used to measure tilt because the total measured acceleration approaches zero. Secondly, for the same activity, signal magnitudes vary greatly depending upon sensor location on the body [12]. This may be problematic for fine-scale estimation of behavioural parameters. For instance, in human accelerometer-based pedometer applications, the accuracy of step counting changes if the pedometer is attached to any location other than the waist [13]. Thirdly, accelerometers may not be well-suited to detection and characterisation of dynamic behaviours involving slow, especially rotation-based, movement [11].

Magnetometers bear surprising similarities to accelerometers: they can measure a static component through inclination with respect to Earth’s magnetic field as well as a dynamic component corresponding to changes in sensor inclination over time. The static component has been used extensively to obtain animal heading and perform dead-reckoning (e.g. [14], [15], [16]). The resulting movement paths have been used to, for instance, quantify differences between straight-line and tortuous-path travel to infer underlying behaviour ([17], [18]), and understand animals’ sense of orientation [19]. The dynamic component of the magnetometer has been used to extract metrics describing angular velocity for human wearable sensing applications ([20], [21]). In spite of these similarities, magnetometers are not prone to the problems highlighted above for accelerometers. Firstly, the magnetometer directly measures sensor tilt. The dynamic component is not mixed with the static component of the signal, and may be obtained by differentiating the signal with respect to time [20]. Note, however, that when the axis of rotation happens to align exactly with the local magnetic field line – an unlikely scenario over extended time – the dynamic component will be zero [5]. The equivalent operation for the accelerometer (i.e. integration with respect to time) does not directly provide velocity because of the need to resolve the constant of integration through knowledge of initial or final velocity from a different source. Secondly, since it is likely that a wild animal’s natural habitat will be far from man-made sources of magnetic field disturbances, the signal magnitude will be the same regardless of activity type or sensor location on the body. Note, however, that the presence of magnetic field disturbances might preclude comparison of signal-derived metrics between different locations. Thirdly,

magnetometers have been shown to be capable of resolving behaviours that are not easily discerned using accelerometers, such as thermal soaring in Himalayan griffon vultures (*Gyps himalayensis*) [5]. Despite the magnetometer's potential for behaviour telemetry, there is a lack of an end-to-end method for identifying common animal behaviours from magnetometer data.

Here, we demonstrate that biomechanically relevant features describing posture, movement intensity, and periodicity can be derived from static and dynamic components of recorded magnetometer data. These can be combined with an existing framework (based on acceleration data, [10]) that combines biomechanics and machine learning to assign accelerometer signals into behavioural categories. We showcase the application of these principles for data collected on wild meerkats (*Suricata suricatta*), a social foraging, <1kg carnivore inhabiting the Kalahari and Namib deserts of Southern Africa [22], where the classification of their main activities such as vigilance, foraging, resting, and running, is essential for characterising their individual and social behaviour. We provide a comparison of magnetometer-based behaviour recognition performance with the accelerometer-based one, discuss the strengths and weaknesses of the magnetometer as a standalone sensor for behaviour recognition, and discuss possibilities for fusing data from both sensors to achieve more accurate and robust behaviour recognition.

4.2 | Methods

4.2.1 | Deriving biomechanical descriptors of movement using magnetometer data

In a recent study on behaviour recognition using accelerometers [10], posture, movement intensity, and periodicity were used as biomechanical descriptors of static and dynamic behaviours.

Behaviour separation using posture estimated from magnetometer data can be achieved when a given axis of the sensor aligns in two opposite directions along the vertical axis for the two static behaviours to be separated (Figure 4.1). [10] used the

accelerometer's surge axis, which corresponds to the same direction as that of the magnetometer's roll axis in the present study, for quantifying posture since values along this axis were least susceptible to changes caused by possible rotations of the collar around the axis of the meerkat's cylindrical neck. Let \vec{B}_E be the local magnetic field vector with dip angle δ at the sensor location. During meerkat vigilance (Figure 4.1a), in an idealised case, the roll axis would point directly upwards, perpendicular to the horizontal plane (the latter shown as a salmon-pink disk), and the sensor's roll axis would measure $|\vec{B}_E| \sin \delta$. During curled-up resting (Figure 4.1b), on the other hand, the roll axis would point downwards, perpendicular to the horizontal plane, and the sensor's roll axis would measure $-|\vec{B}_E| \sin \delta$. We hypothesised that this polarity (positive and negative value of $|\vec{B}_E| \sin \delta$) would enable discrimination of the two static behaviours, vigilance and curled-up resting. When the roll axis lies in the horizontal plane (Figure 4.1c), however, the measurement along the roll axis of the projection of \vec{B}_E onto the horizontal plane, $|\vec{B}_E| \cos \delta$, would be affected by the azimuthal orientation of the animal (angle α between the direction faced with respect to magnetic North in the horizontal plane), and the measured value would now be $|\vec{B}_E| \cos \delta \cos \alpha$. Since values of α may vary arbitrarily between 0° and 360° , the roll axis would record measurements in the range of $[-|\vec{B}_E| \cos \delta, |\vec{B}_E| \cos \delta]$ when it lies in the horizontal plane. Thus, static behaviours such as belly-flat resting as well as dynamic behaviours such as foraging and running may be difficult to separate only on the basis of posture, since the sensor's roll axis can be oriented arbitrarily with respect to the North direction.

In carrying out dynamic activities such as running, the collar sensor would follow in the wake of movements made by the torso and neck as the animal heaves and sways, and rotate around the animal's neck. Thus, the magnetometer axes would continuously change orientation with respect to the local magnetic field lines. The more intense the bodily movement is, the faster the sensor would change orientation with respect to the local field lines. For instance, large oscillations in triaxial magnetometer signals recorded during cheetah (*Acinonyx jubatus*) running behaviour have previously been reported [2]. If this motion is periodic, the change in sensor orientation will also be periodic. Thus, the magnitude of change in recorded signal values could be used as an indicator for the intensity of movement, and help

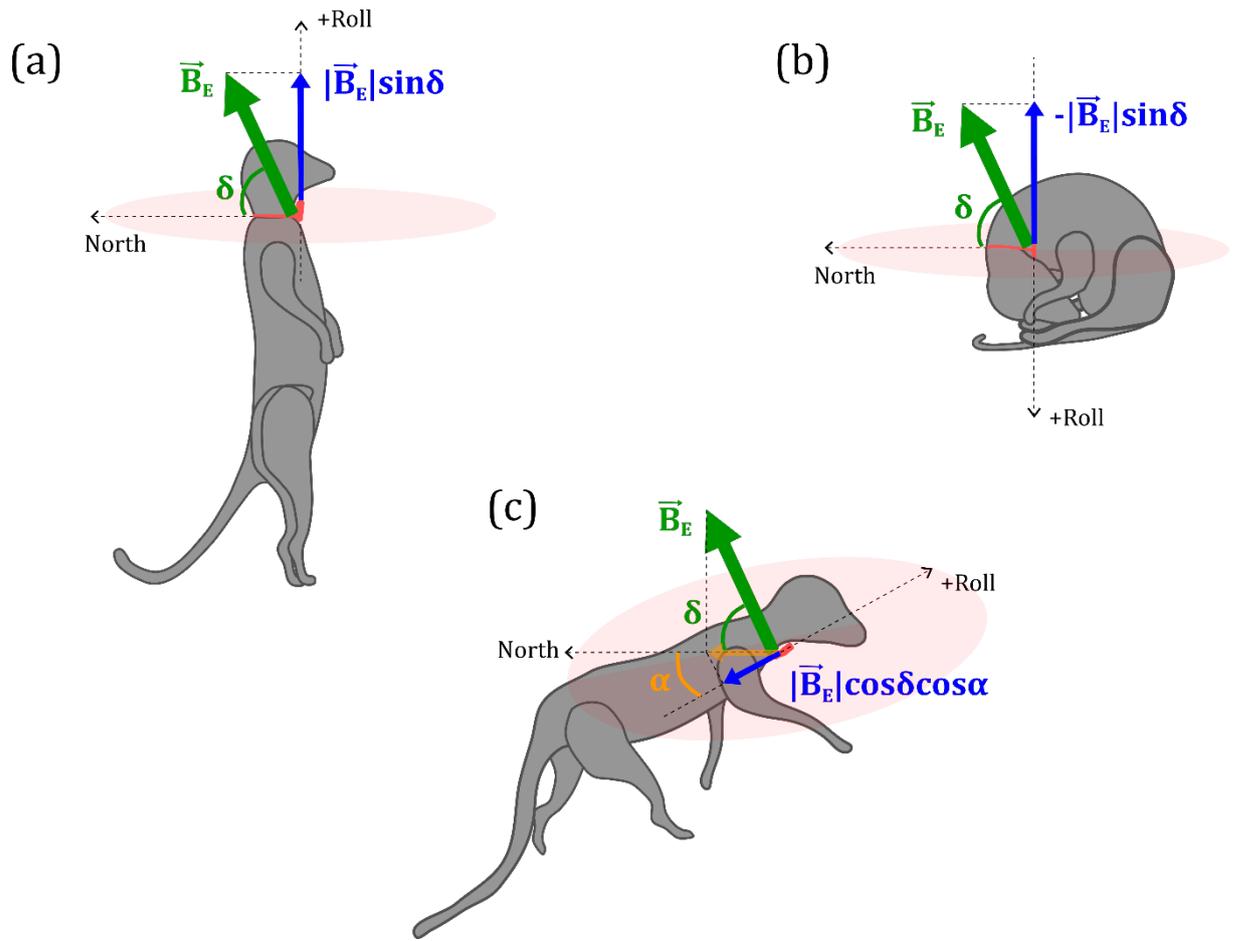


Figure 4.1. Using magnetometer data to distinguish between different meerkat postures. The Earth's magnetic field \vec{B}_E (green arrows) inclined at a dip angle of δ with respect to the horizontal plane (salmon-pink disk) subtends components (blue arrows) equal in magnitude and opposite in sign along the collar sensor's (in red) roll axis during (a) vigilance, and (b) curled-up resting, demonstrated in the simplified case when the roll axis is perfectly aligned with the local vertical direction. When the roll axis lies in the horizontal plane, as shown in (c), the measured component of \vec{B}_E is further affected by the possibly arbitrary azimuthal angle α .

distinguish between static and dynamic behaviours. Further, the periodicity of the rate of change in recorded signal values could be used to distinguish between the dynamic behaviours: for meerkats, running has been shown to be highly periodic, and foraging to be relatively aperiodic [10]. Measures of both intensity and periodicity may be characterised either by computing the amount of variation in the recorded signal itself, through measures such as standard deviation, or by computing the amount of variation in the time-differentiated signal.

4.2.2 | Data collection and groundtruthing

Data from eleven recording sessions of three hours each were collected on ten adult meerkats at the Kalahari Meerkat Project, as described in [10]; one of the individuals was recorded twice. The individuals bore collars equipped with an inertial measurement unit (IMU) (adapted version of Physilog IV, GaitUp SA, Switzerland) containing a triaxial accelerometer (recording at 100 Hz/axis) and triaxial magnetometer [23], the latter recording at a sampling frequency of 50 Hz/axis with a range of $\pm 1000 \mu\text{T}$ and 16-bit resolution. The size of the collar case (IMU and battery) was 35 mm x 29 mm x 19 mm, and overall weight was <25 grams. The total geomagnetic field intensity at the study site was $27.3 \mu\text{T}$, with a declination angle of 17.9° pointing westwards and a dip (or inclination) angle of 65° pointing upwards, according to the International Geomagnetic Reference Field ([24]; values calculated from www.ngdc.noaa.gov/geomag/calculators/magcalc.shtml#igrfwmm). Collars were positioned on the animals so that the axes of the magnetometer were oriented as shown in Figure 4.2. The magnetometer was calibrated prior to each recording session according to the method by [25]. The software used to read magnetometer data resampled the data to 100 Hz/axis using linear interpolation (with the `interp1` function in MATLAB R2016b) to match the sampling frequency of the accelerometer also present on board the recording device.

After the captured animal was collared and released, it was filmed using a handheld video camera recording at 25 frames/second that was synchronised with the collar sensor (see Appendix S4.1 for more details). All videos were annotated using Solomon Coder (version: beta 17.03.22). This video annotation served as the groundtruthing data for our behaviour recognition scheme. Archetypal behaviours observed across a wide range of species [10] – foraging, running, and resting – were considered for the ethogram. In addition, we also considered vigilance, a behaviour typical for meerkats, where the individual is stationary and lifts its head and torso to survey its surroundings. Biologically significant information may be derived from these four behaviours [10]: (1) general stress or alertness level through vigilance, (2) periods of inactivity, mainly due to fatigue or excessive heat, through resting (3) proxies for body condition through foraging, and (4) high energy expenditure and possible important events such as aggressive interactions with rival groups through

running. Any behaviour dependent upon contextual information, such as territory marking or dyadic social interactions, was excluded from the ethogram.

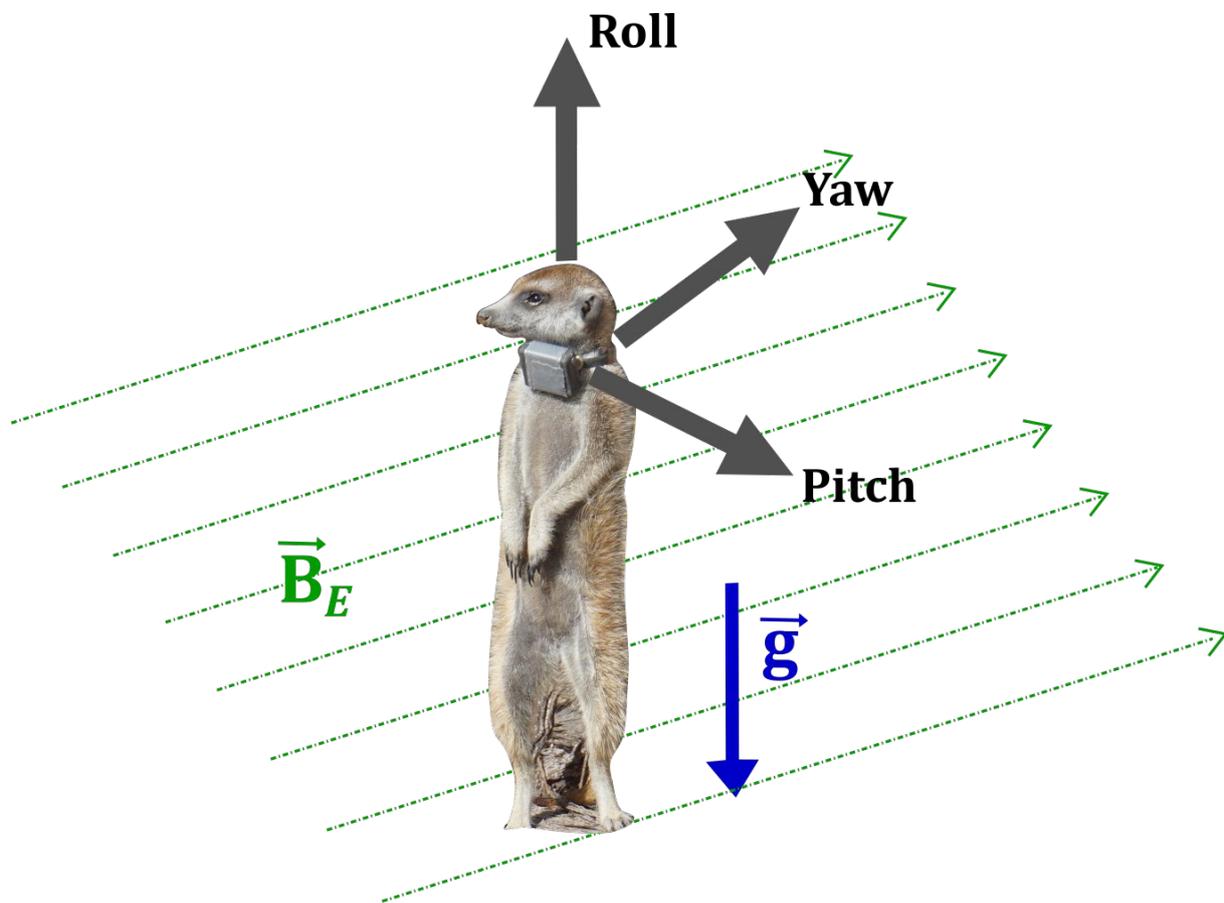


Figure 4.2. **Meerkat with collar, axes, and Earth's fields.** The orientation of the axes of the triaxial magnetometer fixed to a collar on the meerkat along with the directions of two of Earth's naturally occurring fields: Earth's magnetic field \vec{B}_E pointing towards the magnetic North Pole, and Earth's gravity vector \vec{g} pointing vertically downwards.

4.2.3 | Developing candidate features to quantify biomechanical descriptors of movement

Raw triaxial magnetic field intensity data was calibrated and summarised in the form of features quantifying the three biomechanical descriptors of movement discussed in Section 2.1, posture, movement intensity, and periodicity. Feature development followed from previous work done with accelerometers [10]. We computed features on a sliding window w of size two seconds with an overlap of

50% between successive windows. Windows containing data from exactly one video-labelled behaviour were retained, and those containing transitions between different behaviours were excluded. For each biomechanical descriptor, the candidate features (Table 4.1) were computed on each two-second window w containing $N = 200$ calibrated triaxial magnetic field intensity values recorded along the roll (m_{roll}), pitch (m_{pitch}), and yaw (m_{yaw}) axes.

Posture

We obtained a measure of neck inclination with respect to the local magnetic field vector by computing the mean of calibrated magnetic field intensity data recorded in each window w along the roll axis ($meanRoll$, equation 4.1, Table 4.1).

Intensity

We developed four candidate features (#2 to #5, Table 4.1) to quantify movement intensity: one ($stdRoll_w$, equation 4.2, Table 4.1) was aimed at characterising the extent to which m_{roll} varied in window w , whereas the three others aimed to quantify the rate of change of sensor orientation through metrics based on the time-differentiated signal ($meanAbsDiffRoll_w$, equation 4.3; $axMaxMeanAbsDiff_w$, equation 4.4; $avgMeanAbsDiff_w$, equation 4.5, Table 4.1). Since the differentiation operation results in amplification of sensor- and analog-to-digital signal quantization-generated noise at higher frequencies [26], the raw calibrated magnetometer signal was first low-pass filtered using a Butterworth filter of order 4 and cut-off frequency 10 Hz. MATLAB's (version R2016b) `diff` function was used to compute differences between successive signal samples, and each resulting difference was multiplied by the sampling frequency (since, in d/dt , $dt = 1/\text{sampling frequency}$ for discrete signals) to complete the time differentiation operation. To quantify the amount of rate of change in features computed from the time-differentiated signal (features #3, #4, #5 in Table 4.1), we took the absolute values of each differentiated sample and then computed the mean.

Periodicity

We quantified movement periodicity through the use of the Fourier transform (FT). As done in [10], for each window w , before computation of the FT, each input signal was filtered with a Butterworth low-pass filter of order 4 and cut-off frequency

10 Hz, normalised, zero-padded to smooth the frequency spectrum [27] by adding 100 zeroes before and after each two-second input signal, and windowed using the Blackman-Harris windowing function. This processed signal was then transformed with a frequency resolution of $U = 0.01$ Hz (corresponding to FT computation at $L = F_s/U = 10000$ frequencies), and the squared magnitude of each Fourier coefficient ($c_{f_i}^2, i \in 1 \dots L$), corresponding to the power of the signal at frequency f_i , was computed. Triaxial signals yielded three sets of coefficients, one for each axis: $\{c_{f_i,roll}, c_{f_i,pitch}, c_{f_i,yaw}\}$ in the case of the raw calibrated triaxial signal, and $\{\delta_{f_i,roll}, \delta_{f_i,pitch}, \delta_{f_i,yaw}\}$ in the case of the time-differentiated signal. For a triaxial signal, the resulting FT was averaged across the three axes. From the final FT, the maximum power obtained across all frequencies f_i ($i \in 1 \dots L$) was chosen as a measure of the signal periodicity. This FT-based operation was applied to four different input signals in order to develop four candidate features characterising movement periodicity: (1) roll component of the local magnetic field (*rollFftPeakPower*, equation 4.6, Table 4.1), (2) triaxial magnetometer signal (*avgFftPeakPower*, equation 4.7, Table 4.1), (3) time-differentiated roll signal (*rollDiffFftPeakPower*, equation 4.8, Table 4.1), and (4) time-differentiated triaxial signal (*avgDiffFftPeakPower*, equation 4.9, Table 4.1). All feature computation was done using MATLAB R2016b.

4.2.4 | Feature selection

To enable direct comparison with the three-feature accelerometer-based model in [10], we selected one feature for each of the three biomechanical descriptors of posture, movement intensity, and periodicity. We tested features quantifying movement intensity (feature # 2 to #5, Table 4.1) for their efficacy in separating static and dynamic behaviours, and foraging and running. We tested features quantifying movement periodicity (feature #6 to #9, Table 4.1) for their efficacy in separating foraging and running. We tested five different feature selection methods based on the filter method (using the `rankfeatures` function in MATLAB R2016b, © 2003-2016 The MathWorks, Inc. See Appendix S4.3 for more details) to select one feature to quantify movement intensity, and one to quantify periodicity. *meanRoll* (feature #1, Table 4.1), being the only candidate developed to

#	Biomechanical descriptor	Feature name	Feature description	Computation
1.	Posture	<i>meanRoll</i>	Mean of data from roll axis	$\frac{\sum_N m_{roll,w}}{N}$ (4.1)
2.		<i>stdRoll</i>	Standard deviation of data from roll axis	$std(m_{roll,w})$ (4.2)
3.		<i>meanAbsDiffRoll</i>	Mean of absolute values of time-differentiated roll data	$\frac{\sum_N \left \frac{d}{dt}(m_{roll,w}) \right }{N}$ (4.3)
4.	Intensity	<i>axMaxMeanAbsDiff</i>	Maximum, across axes, of mean of absolute values of time-differentiated data from each axis	$\max_{A \in \text{roll,pitch,yaw}} \left(\frac{\sum_N \left \frac{d}{dt}(m_{A,w}) \right }{N} \right)$ (4.4)
5.		<i>avgMeanAbsDiff</i>	Mean, across axes, of mean of absolute values of time-differentiated data from each axis	$\sum_{A \in \text{roll,pitch,yaw}} \frac{\sum_N \left \frac{d}{dt}(m_{A,w}) \right }{3N}$ (4.5)
6.		<i>rollFftPeakPower</i>	Maximum squared coefficient of Fourier transform of data from roll axis	$\max_{i \in 1 \dots L} (c_{f_i,roll,w}^2)$ (4.6)
7.	Periodicity	<i>avgFftPeakPower</i>	Mean, across axes, of maximum squared coefficient of Fourier transform of data from each axis	$\max_{i \in 1 \dots L} \left(\frac{c_{f_i,roll,w}^2 + c_{f_i,pitch,w}^2 + c_{f_i,yaw,w}^2}{3} \right)$ (4.7)
8.		<i>rollDiffFftPeakPower</i>	Maximum squared coefficient of Fourier transform of time-differentiated roll data	$\max_{i \in 1 \dots L} (\delta_{f_i,roll,w}^2)$ (4.8)
9.		<i>avgDiffFftPeakPower</i>	Mean, across axes, of maximum squared coefficient of Fourier transform of time-differentiated data from each axis	$\max_{i \in 1 \dots L} \left(\frac{\delta_{f_i,roll,w}^2 + \delta_{f_i,pitch,w}^2 + \delta_{f_i,yaw,w}^2}{3} \right)$ (4.9)

Table 4.1. **Feature development.** Candidate features developed to describe the three biomechanical descriptors used in this study: posture (#1), movement intensity (#2 to #5), and movement periodicity (#6 to #9). Features were computed on each two-second window w containing $N = 200$ calibrated tri-axial magnetic field intensity values recorded along the roll (m_{roll}), pitch (m_{pitch}), and yaw (m_{yaw}) axes. Equation numbers are indicated on the right.

describe posture, was chosen by default.

4.2.5 | Behaviour recognition scheme and cross-validation

The behaviour recognition scheme had the same hierarchical tree-like structure and hybrid form as the one found for meerkat behaviour recognition using accelerometers [10]. The scheme consisted of three nodes, each dividing a parent behavioural category (static or dynamic) into two daughter behavioural types (vigilance/resting or foraging/running, respectively). A Support Vector Machine (SVM) was used at each node to obtain optimal feature-value thresholds in a completely automated fashion. At the first node, features encoding information on posture and movement intensity were used to separate static and dynamic behaviours. At the second node, static behaviours were separated into vigilance and resting using postural information. At the third node, dynamic behaviours were separated into foraging and running using information on movement intensity and periodicity. The `svm` learner in MATLAB R2016b's `fitclinear` function (© 2015-2016 The MathWorks, Inc.) was used to train the SVM at each node.

To validate the predictions of the SVM-SVM-SVM hybrid model with the chosen features against groundtruth video-annotated behaviours, two cross-validation methods were tested: (1) stratified ten-fold cross-validation (STRAT), which evaluates model performance when the frequency and duration of different behaviours may be skewed, and (2) leave-one-individual-out cross-validation (LOIO), which evaluates model performance when inter-individual variability is taken into account [10]. We used standard confusion matrix-based metrics to evaluate and compare model performance. These performance statistics included three behaviour-specific metrics (sensitivity, precision, and specificity), and overall model accuracy (see Appendix S4.2 for mathematical definitions, computation and interpretation). Custom software was written in MATLAB R2016b to perform cross-validation.

4.3 | Results

4.3.1 | Collected data

A total of 82550 two-second bouts of video-labelled behaviour were collected for the four behaviours of interest (Table 4.2). The number of bouts collected per animal was 8255 ± 3229 (mean \pm SE). The frequency and duration of different behaviours were skewed: foraging (56.2%) was the most common behaviour while running was the rarest (1%). No resting behaviour was observed during six out of the eleven recording sessions; the number of resting bouts collected during the first recording session (55.7% of all resting bouts) far outnumbered those collected during the other recording sessions. Typical signals recorded for the four behaviours (Figure 4.3) were found to be in line with our biomechanical hypotheses (Section 2.1): static behaviours (bipedal vigilance and curled-up resting) showed little change (Figure 4.3, left), while dynamic behaviours (foraging and running) produced greater change in the signals with large, periodic oscillations during running (Figure 4.3, right).

#	Vigilance	Resting	Foraging	Running	Bouts per recording
1	4594	2114	1562	69	8339
2	3896	120	5315	29	9360
3	1453	0	6278	38	7769
4	5221	0	2823	161	8205
5	1890	0	6134	169	8193
6	1639	744	4438	98	6919
7	4785	156	3498	40	8479
8	71	0	4841	20	4932
9	4283	0	1713	43	6039
10	1906	0	4407	84	6397
11	1782	661	5398	77	7918
Bouts per Activity	31520	3795	46407	828	82550 (total bouts)

Table 4.2. **Summary of data collected.** Table adapted from [10]. Triaxial magnetometer data were collected on ten unique individuals; data from recording session #4 and #7 were collected on the same individual. A bout refers to a two-second window w containing one video-labelled behaviour.

4.3.2 | Features to quantify biomechanical descriptors from triaxial magnetometer data

Measures of posture (*meanRoll*, equation 4.1, Table 4.1) and movement intensity (*meanAbsDiffRoll*, equation 4.3, Table 4.1) were inputs to the first node to separate static behaviours from dynamic ones (Figure 4.4). Posture (*meanRoll*) was used to distinguish vigilance from resting in the second node, and finally, movement intensity (*meanAbsDiffRoll*) and periodicity (*avgDiffFftPeakPower*, equation 4.9, (Table 4.1) were used to distinguish foraging from running in the third node (Figure 4.4).

The use of *meanRoll* to quantify posture produced high separability between bipedal vigilance and curled-up resting (Figure 4.3 & 4.5). During the dynamic behaviours (foraging and running), where the orientation of the animal's body caused the roll axis of the magnetometer to lie approximately in the horizontal plane, the values recorded along the roll axis (Figure 4.3) were in an intermediate range between the extreme positive and extreme negative values recorded during bipedal vigilance (Figure 4.1a) and curled-up resting (Figure 4.1b), respectively.

Among the features developed to quantify movement intensity, *meanAbsDiffRoll* outperformed the other three candidates with regard to separating both static from dynamic behaviours (Table S4.1, Appendix S4.3), and foraging from running (Table S4.2, Appendix S4.3). Among the features developed to quantify movement periodicity, *avgDiffFftPeakPower* outperformed the other three candidates for the separation of foraging from running (Table S4.3, Appendix S4.3).

4.3.3 | Performance evaluation, and comparison with accelerometer-based behaviour recognition

Magnetometer-based behaviour recognition performance is presented and compared with that achieved with accelerometer data in [10] for STRAT (Table 4.3) and LOIO (Table 4.4), and through visual depiction of feature distributions and resulting decision boundaries (Figure 4.5).

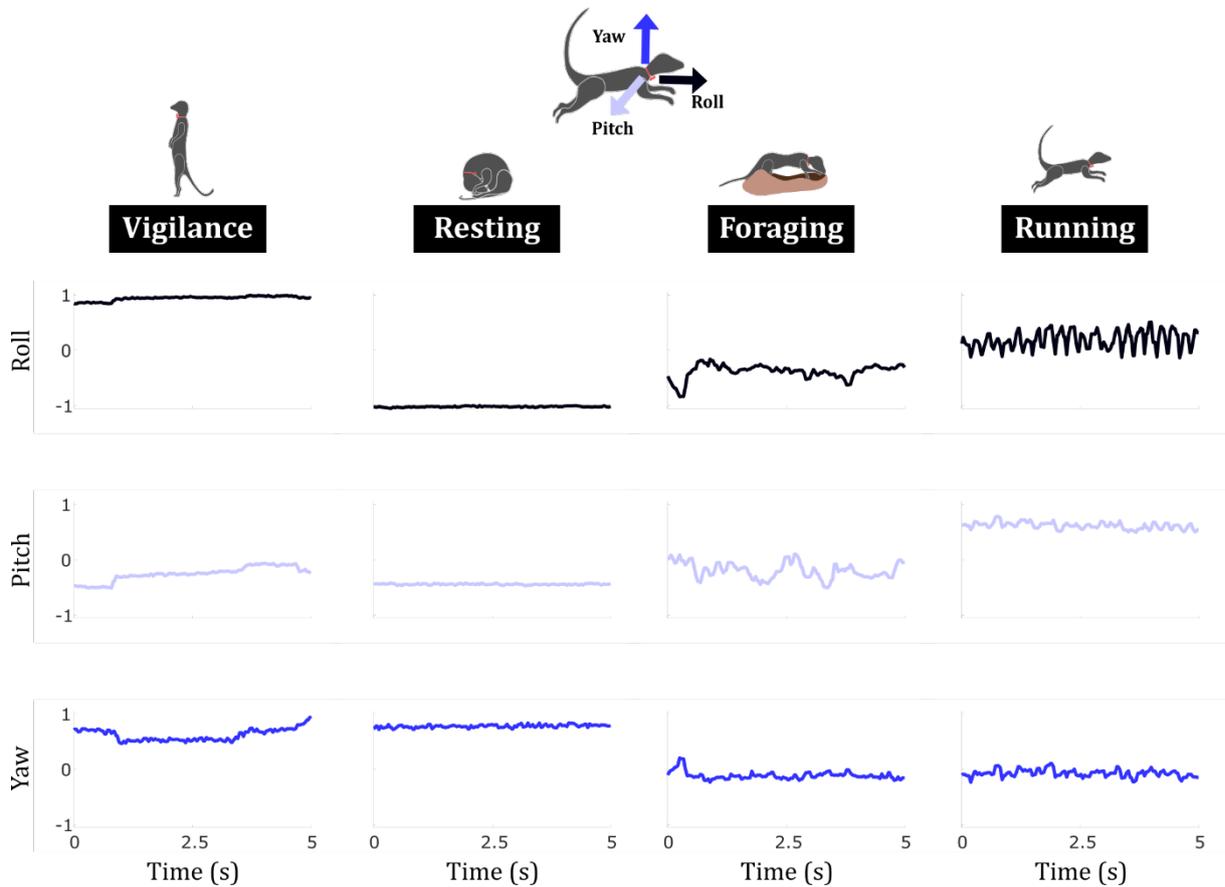
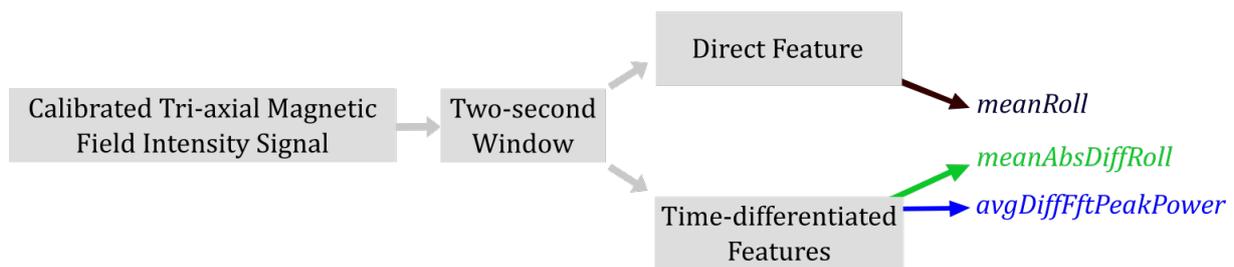


Figure 4.3. **Five-second snapshots of calibrated triaxial magnetometer data** for the four behaviours of interest for a typical individual (recording session #1). The horizontal axis shows time in seconds, and the vertical axis represents calibrated, normalised magnetic field intensity measured along the three axes of the sensor in each graph. The signals correspond, from left to right, to bipedal vigilance, curled-up resting, foraging, and running.

For STRAT, all performance metrics for the most common behaviours (foraging: 56.2% of dataset; vigilance: 38.2% of dataset), and overall model accuracy, were >95% (Table 4.3). Good performance was obtained even for the rarer behaviours, resting (4.6% of dataset) and running (1% of dataset), where all behaviour-specific metrics remained >83%. Further, overall as well as behaviour-wise recognition performance with the magnetometer were similar to that with the accelerometer (Table 4.3).

For LOIO, data from recording sessions numbers 3, 4, 5, 8, 9 and 10 were discarded since they did not contain any resting behaviour (Table 4.2). Once again,

(a)



(b)

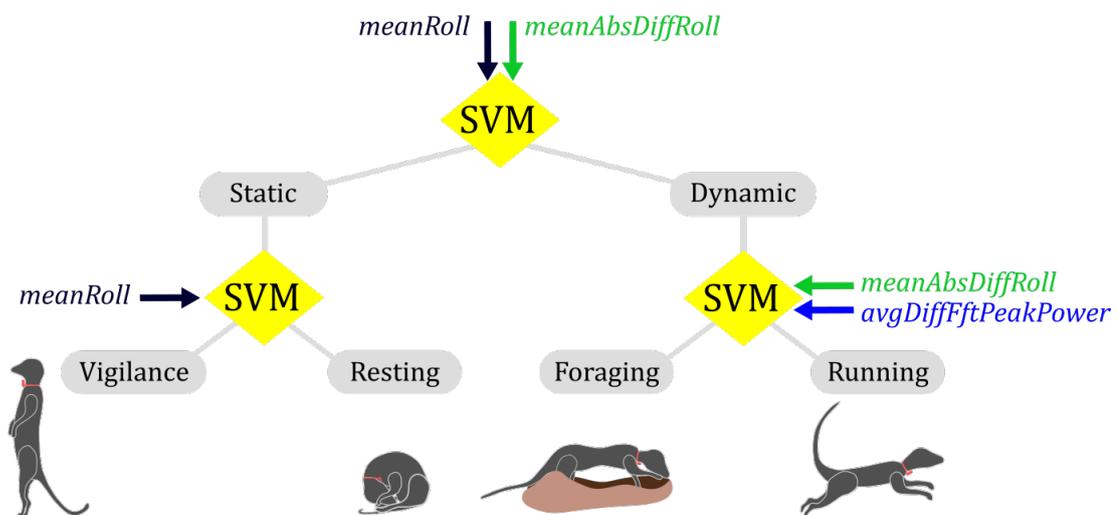


Figure 4.4. **Behaviour Recognition Scheme.** (a) Flowchart showing feature computation: *meanRoll* quantifies posture, *meanAbsDiffRoll* movement intensity, and *avgDiffFftPeakPower* periodicity. (b) Hierarchical classification scheme classifying behaviours as being either static or dynamic, then static behaviours as being either vigilance or resting, and finally dynamic behaviours as being either foraging or running.

even when inter-individual variation was taken into account, mean values of all performance metrics for the most common behaviours (foraging and vigilance), and overall model accuracy, were >95%, and were similar to those obtained with accelerometer-based behaviour recognition (Table 4.4).

4.4 | Discussion

We presented an end-to-end framework to identify common animal behaviours from magnetometer data. Using data collected on 10 wild meerkats, we demonstrated that accurate behaviour recognition can be achieved with a magnetometer alone with performance comparable to that with an accelerometer. Our results shed further light on the magnetometer's strengths and weaknesses in the context of behaviour telemetry, and suggest possibilities for leveraging the complementary merits of accelerometers and magnetometers within a single classification framework for more robust behaviour recognition.

Distinguishing dynamic behaviour using magnetometer-derived angular velocity

Differentiating magnetic field intensity with respect to time corresponds to quantifying changes in angles subtended by the Earth's magnetic field vector onto the three sensor axes with time, and provides an estimate of angular velocity [20]. To separate behaviours based on movement intensity, quantifying change in magnetometer-derived angular velocity was more effective than quantifying change in magnetic field values. This may be because even when the change in sensor inclination angle is small, the rate at which the angle changes may be high. *meanAbsDiffRoll* (equation 4.3, Table 4.1) was best at separating static and dynamic, and the two dynamic behaviours. The superior class separability of *meanAbsDiffRoll* implied that using only the roll axis was more effective than when contributions from the other two axes, pitch and yaw, were included. This may have been a consequence of the fact that the roll axis succeeded in capturing both up-and-down, and side-to-side bodily movements made by the meerkat's neck and torso during dynamic behaviours. The roll axis was also more robust than the other two axes to collar rotations. Magnetic field lines have, in general, a horizontal as well as vertical component – the inclination angle of Earth's magnetic field at the study site was 65° pointing upwards. The pitch axis would have been insensitive to the up-and-down movements, and the yaw axis insensitive to the side-to-side movements. Further, collar rotations around the meerkat's cylindrical neck could have confounded class separation through noisy variability in pitch- and yaw-axis contributions for the same

Sensor	Vigilance			Resting			Foraging			Running			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
Magnetometer	97	98.6	97.8	84.4	99.4	87.1	98.8	97.2	97.8	83.1	99.9	93.9	97.3
Accelerometer	97.1	98.8	98.1	85	99.4	87.1	99.3	97.8	98.3	85.9	99.9	92.1	97.7

Table 4.3. **STRAT cross-validation results.** The performance of the SVM-SVM-SVM hybrid model with magnetometer data is benchmarked against that obtained with accelerometer data reported in [10]. SVM: Support Vector Machine.

Sensor	Vigilance			Resting			Foraging			Running			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
Magnetometer	95.2	97.9	95.2	65.4	98.9	77.3	98.4	97.0	95.5	86.5	100	96.4	96.0
	± 2.4	± 1.9	± 6.2	± 25.9	± 0.9	± 31.1	± 0.9	± 1.2	± 0.5	± 3.7	± 0.0	± 3.4	± 1.5
Accelerometer	95.8	98.4	96.4	71.4	98.9	81.1	98.8	97.4	95.3	86.3	99.9	89.1	96.5
	± 2.8	± 1.2	± 4.5	± 23.6	± 1.2	± 28.0	± 1.0	± 1.5	± 7.0	± 13.2	± 0.1	± 11.1	± 1.8

Table 4.4. **LOIO cross-validation results.** The performance of the SVM-SVM-SVM hybrid model with magnetometer data is benchmarked against that obtained with accelerometer data reported in [10]. Performance metrics were calculated separately for each test individual, and their mean and standard deviation across test individuals are shown here. SVM: Support Vector Machine.

With Accelerometer

With Magnetometer

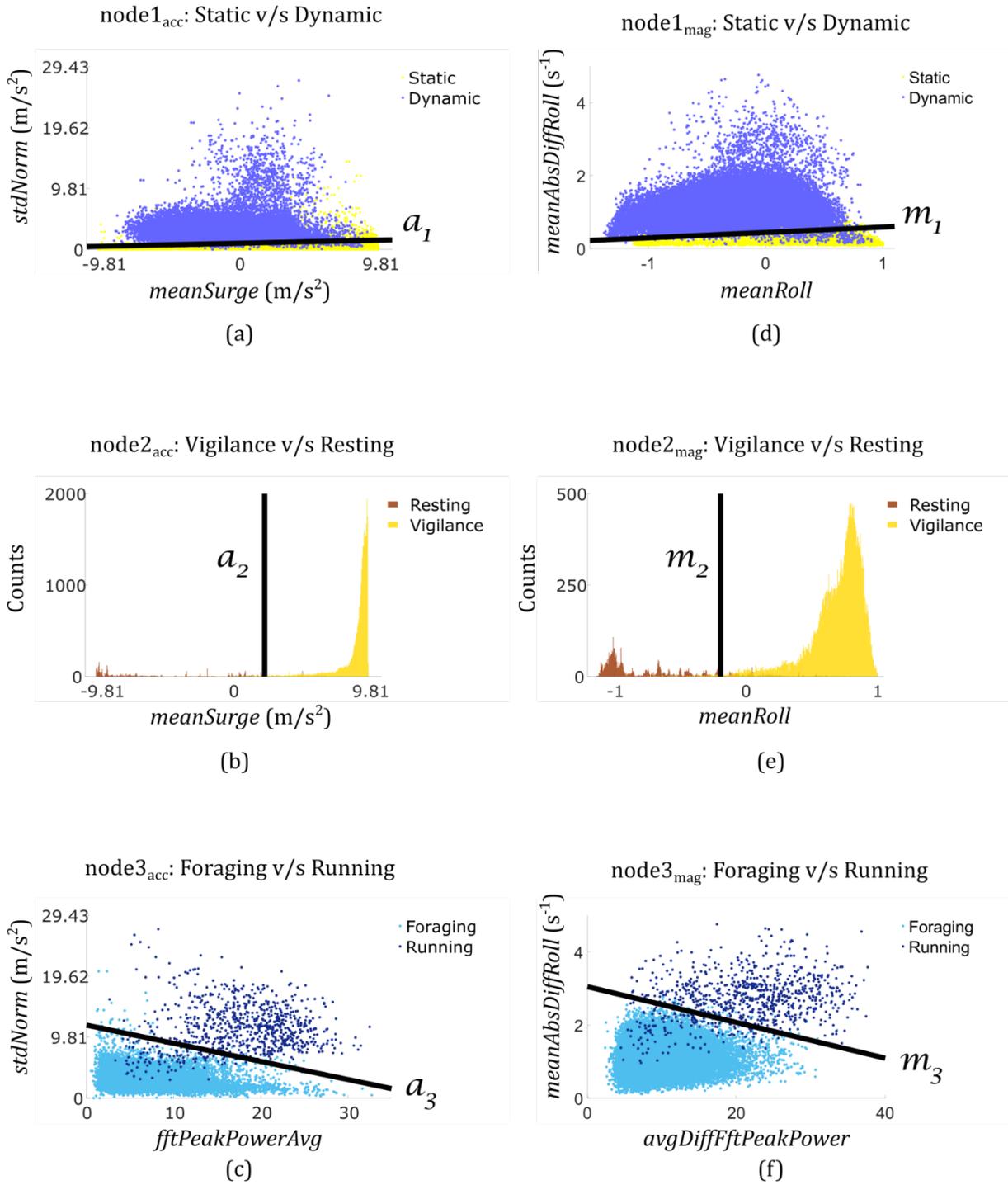


Figure 4.5. **Decision boundaries and feature distributions** obtained with accelerometer- (left) and magnetometer-based (right) behaviour recognition with Support Vector Machines trained on the entire dataset for each of the three nodes of the hierarchical behaviour recognition scheme. m_i and a_i refer to decision boundaries obtained with the magnetometer and accelerometer, respectively, with the subscript i indicating the node index.

activity. Note, however, that the precise choice of the feature describing movement intensity may change when dynamic behaviours of interest involve rotations about the roll axis, such as washing at sea by a Magellanic penguin (*Spheniscus magellanicus*) [5], or fast turning in cheetahs where the weight of the tag causes the collar to rotate around the neck due to centripetal acceleration [28].

Metrics based on magnetometer-derived angular velocity may be better suited than accelerometry for filtering out signal artefacts caused by sensor impacts. Compared to foraging versus running classification using accelerometer data (a_3 in Figure 4.5c), with the magnetometer there were fewer foraging bouts with low periodicity and high intensity that crossed the decision boundary m_3 (Figure 4.5f). While exploring the ground for prospective hunting locations, the meerkat's collar would often bump against vegetation or the ground. Additionally, while digging, the meerkat's pectoral muscles would hit against the collar. These impacts produced high, transient translational acceleration that led to a higher estimation of bout intensity. However, the magnetometer, being insensitive to translational acceleration [5], provided a lower estimate for bout intensity due to relatively slow collar orientation changes. Thus, such bouts were correctly classified as foraging with the magnetometer since their intensity placed them below the decision boundary m_3 (Figure 4.5f). This led to higher precision in the detection of running (7.3% higher mean precision and similar mean sensitivity with LOIO) with much lesser inter-individual variability in performance (9.5% lower standard deviation for sensitivity and 7.7% lower standard deviation for precision) compared to accelerometer-based classification. This was achieved despite running being the rarest behaviour (outnumbered 1:56 by foraging in terms of number of recorded bouts). In a similar fashion, fewer vigilance bouts yielded high enough magnetometer-based intensity to cross over m_1 (Figure 4.5d) and get misclassified as dynamic behaviour as compared to when the accelerometer (Figure 4.5a) was used (Table S4.4 and Table S4.5).

The stringency of the magnetometer in assigning high intensity to a bout of activity was not without its costs. Comparing aggregate confusion matrices observed with the magnetometer (Table S4.4) and accelerometer (Table S4.5), we observed a

higher number of relatively low-intensity foraging bouts getting misclassified as being static, thereby reducing foraging detection sensitivity as compared to accelerometer-based classification (especially for recording sessions #6, #7 and #11, Table S4.8). This may have been because the amplitude and rate of body movement-generated change in collar orientation during low-intensity foraging behaviour (for instance during slow ground scratching while keeping the head and torso in the same orientation) may not have been sufficient to generate a large-enough signal detectable above the noise floor introduced by the differentiation operation [26] during computation of *meanAbsDiffRoll*.

Finally, it has been reported that a combination of accelerometers and gyroscopes can lead to better activity recognition in human wearable sensor applications than when each sensor is used alone [20]. In animal studies, the magnetometer may be a viable alternative to the gyroscope for obtaining estimates of angular velocity due to the former's lower power consumption [29]. This could be important for facilitating long-duration recordings on small animals.

Estimating posture using magnetometer data

While it was possible to estimate posture using the magnetometer, the accelerometer-based posture measure was nevertheless found to be better at separating static behaviours. In our observations of static behaviour, a number of bouts of quadrupedal vigilance and belly-flat resting were also recorded apart from bipedal or sitting vigilance (Figure 4.1a), and curled-up resting (Figure 4.1b). In these postures, a significant component of the roll axis lay in the horizontal plane. Possibly arbitrary azimuthal orientation of the animal during these postures (Figure 4.1c) confounded the distinction between quadrupedal vigilance and belly-flat resting. This additional constraint degraded the accuracy of resting detection compared to that with the accelerometer (6% lower mean sensitivity, 3.8% lower mean precision. See also Figure 4.5, middle panel). In static behaviours, where the animal's body retains similar orientation with respect to the horizontal plane, such as during standing and lying in cows (*cf.* [30]), the confounding effect of possibly arbitrary azimuthal orientation may be particularly severe. Further, our implicit assumption that the calibration parameters computed at the beginning of each recording would be valid throughout the recording was found to be only partially true (see Appendix S4.5).

Magnetometer versus accelerometer: similarity and complementarity

Similar behaviour recognition performance with the two sensors suggests that it may not be necessary to make separate considerations for the choice of ethogram when working with magnetometers when archetypal behaviours such as foraging, fast locomotion, and resting are to be identified.

Our results reveal the selectivity of the magnetometer for bodily movement, and relative immunity to signal artefacts arising due to sensor impacts. This may offer the opportunity to study movement energetics using metrics based on magnetometer-derived angular velocity [5], which would be similar but complementary to the acceleration-based metrics ODBA [31] and VeDBA [32]. One advantage of the magnetometer that could be exploited in future studies is the weaker dependence of signal magnitude on sensor location on the animal's body. When a body segment rotates about a joint, the magnitude of acceleration is higher for distal compared to proximal parts, and this dependence on the location of accelerometer attachment might be especially important to take into account for larger animals. The magnitude of the magnetometer signal during segment rotation, however, would always be the same along a body segment regardless of body size or sensor placement. The apparent pitfall of the accelerometer in confounding bodily movement-produced signals with artefacts arising from sensor impacts could nevertheless be turned to an advantage for other applications where the detection of specific events is desirable. Impact-generated acceleration characteristics have, for instance, been used in the detection of falls in humans [33].

Combining magnetometer and accelerometer data to identify behaviour has been previously suggested [2]. In this study, we develop this idea further and suggest specific aspects of these two sensors to combine for better behavioural identification. Features derived from data from one or both sensors may be chosen according to their specific strengths as inputs for each node of the hierarchical classification scheme (Figure 4.4b). For instance, at the first node tasked with separating static behaviours from dynamic ones, the more reliable accelerometer-based posture measure (*meanSurge*) [10] may be combined with the more selective magnetometer-based movement intensity measure (*meanAbsDiffRoll*). Then, *meanSurge* could be used at the second node tasked with separating vigilance from resting on the basis of posture. At the third node, the magnetometer-based intensity (*meanAbsDiffRoll*) and

periodicity (*avgDiffFftPeakPower*) metrics may be used for higher-precision distinction between foraging and running. Finally, as has been done for some human movement studies [34], accelerometer and magnetometer data may be combined to give a more accurate and robust three-dimensional estimation of posture in such fused systems than either sensor alone.

4.5 | Conclusion

Our findings demonstrate that magnetometers can be used alone to achieve accurate and robust animal behaviour recognition. We showed that sensor tilt with respect to Earth's magnetic field, and metrics based on magnetometer-derived angular velocity may be used to extract biomechanically significant features to describe posture, movement intensity, and periodicity. Through the directed use of these features in a recently developed hybrid hierarchical behaviour recognition framework combining movement biomechanics and machine learning [10], we found that magnetometer-based behaviour recognition (i) produced similar results to those obtained with the accelerometer, (ii) was robust to inter-behaviour differences in duration and frequency of occurrence, and (iii) exceeded the accelerometer's resilience to inter-individual variability for dynamic behaviours.

Movements performed by free-living animals, broadly speaking, generate both acceleration as well as angular velocity. Our results reveal that, as long as a sensor can measure a static and dynamic component of movement, key biomechanical descriptors of motion can be quantified and used to recognise common animal behaviours with high accuracy. The generality afforded by the usage of biomechanical considerations to direct inertial sensor data processing, and the simple structure and implementation of the hybrid behaviour recognition framework make it possible to accommodate, compare, and leverage data from accelerometers, magnetometers, and gyroscopes within a single behaviour recognition scheme.

Appendices

Appendix S4.1: Synchronisation of the animal-borne IMU with the hand-held camera

To synchronise videos with the recorded sensor data, a hand-held Physilog IV system (GaitUp SA, Switzerland) connected to an ‘event marker’ was used. The event marker contained a button and an LED that switched on when the button was pressed. Prior to each deployment, the hand-held system was configured to communicate and synchronise with the animal-borne inertial measurement unit (IMU) over a dedicated radio frequency channel. At the beginning and end of each video, we pressed the button on the event marker and filmed the resulting LED. During post-processing, for each video, we matched the moment when the synchronisation radio pulse was received on the animal-borne system with the frame number of the video that showed the event marker’s LED lighting up. Knowing the starting and ending points of the recording in both the IMU and video data enabled us to synchronise both data streams and correct the camera clock’s drift with respect to the IMU micro-processor clock.

Appendix S4.2: Metrics for performance evaluation of classification models

As done in [10], we used a confusion matrix $[N_{jk}]$ to summarise model performance, where $j, k \in \{1: \text{vigilance}, 2: \text{resting}, 3: \text{foraging}, 4: \text{running}\}$ and N_{jk} denotes the number of samples of behaviour ‘ j ’ that were predicted by the model to be behaviour ‘ k ’. From this confusion matrix, we computed three behaviour-specific metrics - sensitivity (often referred to as recall in the literature), precision, and specificity, as well as overall model accuracy. These measures are computed as shown in Table S3.1.

Appendix S4.3: Feature Selection

Feature selection was done using the MATLAB R2016b function `rankfeatures` (© 2003-2016 The MathWorks, Inc.). Five standard filter methods were used for ranking candidate features developed to quantify movement intensity and periodicity, and the results are listed in the sections below.

1 | Intensity

This section gives details on feature selection that show how the best feature to quantify movement intensity was chosen from the candidate features for distinguishing between: (i) static versus dynamic behaviour (top node in Figure 4.4b in manuscript), and (ii) foraging versus running (bottom right node in Figure 4.4b in manuscript).

1.1 | Static versus Dynamic Behaviour Classification

This section compares the candidate features quantifying movement intensity for their efficacy in distinguishing between static and dynamic behaviour.

Rank	t-test		Entropy		Bhattacharyya		Roc		Wilcoxon (or Mann-Whitney)	
	Feature	Score	Feature	Score	Feature	Score	Feature	Score	Feature	Score
1	f4	336.42	f2	9.71	f2	0.08	f2	0.48	f1	1.41
2	f2	489.03	f4	60.88	f4	0.58	f4	0.50	f2	0.75
3	f3	390.88	f3	28.13	f3	0.38	f3	0.49	f4	0.01
4	f1	521.41	f1	37.88	f1	0.41	f1	0.50	f3	0.67

Table S4.1. Ranking features describing movement intensity for static versus dynamic behaviour classification. Serial numbers assigned to features: {f1: *stdRoll*; f2: *meanAbsDiffRoll*; f3: *axMaxMeanAbsDiff*; f4: *avgMeanAbsDiff*}.

1.2 | Foraging versus Running Classification

This section compares the candidate features quantifying movement intensity for their efficacy in distinguishing between the two dynamic behaviours, foraging and running.

Rank	t-test		Entropy		Bhatacharyya		Roc		Wilcoxon (or Mann-Whitney)	
	Feature	Score	Feature	Score	Feature	Score	Feature	Score	Feature	Score
1	f2	4.55	f2	0.55	f2	0.02	f2	0.00	f1	1.50
2	f3	73.24	f3	14.47	f3	0.68	f3	0.49	f2	0.52
3	f4	68.34	f4	12.74	f4	0.61	f4	0.48	f3	0.51
4	f1	62.14	f1	8.93	f1	0.29	f1	0.47	f4	0.47

Table S4.2. Ranking features describing movement intensity for dynamic behaviour classification (foraging versus running). Serial numbers assigned to features: {f1: *stdRoll*; f2: *meanAbsDiffRoll*; f3: *axMaxMeanAbsDiff*; f4: *avgMeanAbsDiff*}.

Conclusion: *meanAbsDiffRoll* consistently outperforms the other three candidate features as a metric to separate both static from dynamic behaviour as well as foraging from running on the basis of movement intensity.

2 | Periodicity: Foraging versus Running Classification

This section compares the candidate features quantifying movement periodicity for their efficacy in distinguishing between the two dynamic behaviours, foraging and running.

Rank	t-test		Entropy		Bhatacharyya		Roc		Wilcoxon (or Mann-Whitney)	
	Feature	Score	Feature	Score	Feature	Score	Feature	Score	Feature	Score
1	f4	5.30	f4	0.06	f4	0.03	f4	0.04	f4	0.81
2	f3	7.85	f3	0.09	f3	0.06	f3	0.07	f2	0.91
3	f2	42.07	f2	5.04	f2	2.94	f2	0.38	f1	0.80
4	f1	44.55	f1	9.59	f1	3.16	f1	0.41	f3	0.93

Table S4.3. Ranking features describing movement periodicity for dynamic behaviour classification (foraging versus running). Serial numbers assigned to features: {f1: *rollFftPeakPower*; f2: *avgFftPeakPower*; f3: *rollDiffFftPeakPower*; f4: *avgDiffFftPeakPower*}.

Conclusion: *avgDiffFftPeakPower* consistently outperforms the other three candidate features as a metric for separating foraging from running on the basis of movement periodicity.

Appendix S4.4: LOIO Results

The aggregate and individual-wise confusion matrices obtained with the SVM-SVM-SVM hybrid model during LOIO are provided here.

		<i>Predicted</i>			
		Vigilance	Resting	Foraging	Running
<i>Actual</i>	Vigilance	16148	227	320	1
	Resting	481	3208	105	1
	Foraging	188	107	19903	13
	Running	0	0	47	266

Table S4.4. Aggregate confusion matrix obtained during LOIO with magnetometer data. Predictions are arranged along the columns and true labels along the rows; correct predictions are shown in bold.

		<i>Predicted</i>			
		Vigilance	Resting	Foraging	Running
<i>Actual</i>	Vigilance	16069	279	340	8
	Resting	337	3252	203	3
	Foraging	80	102	20002	27
	Running	0	0	47	266

Table S4.5. Aggregate confusion matrix obtained with the SVM-SVM-SVM hybrid model for LOIO with accelerometer data, from [10]. Predictions are arranged along the columns and true labels along the rows; correct predictions are shown in bold.

Vigilance			Resting			Foraging			Running			Overall accuracy (%)
Sen (%)	Spec (%)	Prec (%)	Sen (%)	Spec (%)	Prec (%)	Sen (%)	Spec (%)	Prec (%)	Sen (%)	Spec (%)	Prec (%)	
96.7	97.2	96	84.5	99.1	90.6	98.5	97.7	97.7	85	100	94.7	96.4

Table S4.6. Aggregate performance metrics for magnetometer-based behaviour recognition for LOIO using the SVM-SVM-SVM hybrid model.

Recording Session No.	Vigilance			Resting			Foraging			Running			Overall accuracy (%)
	Sen (%)	Spec (%)	Prec (%)	Sen (%)	Spec (%)	Prec (%)	Sen (%)	Spec (%)	Prec (%)	Sen (%)	Spec (%)	Prec (%)	
1	95.6	98.6	98.9	96.3	99.0	97.1	97.1	96.8	87.3	87	100	96.8	96
2	96.2	98.5	97.9	36.7	98.1	20.2	98.8	99.2	99.4	89.7	100	100	96.9
6	92.3	99	96.5	96	98.9	91.2	98.8	96.2	97.9	85.7	99.9	95.5	96.8
7	99.5	97.6	98.1	44.9	100	97.2	98.9	98.7	98.2	82.5	100	97.1	98.2
11	97.3	93.6	81.6	52	99.6	92.2	97.9	97.5	98.8	81.8	99.9	88.7	93.8

Table S4.7. Individual-wise results for magnetometer-based behaviour recognition with the SVM-SVM-SVM hybrid model for LOIO.

#	Confusion Matrix with Accelerometer					Confusion Matrix with Magnetometer						
1	<i>Actual</i>	<i>Predicted</i>				<i>Actual</i>	<i>Predicted</i>					
			Vig.	Rest	Frg.		Run		Vig.	Rest	Frg.	Run
		Vig.	4310	167	115		2	Vig.	4394	39	161	0
		Rest	16	1918	179		1	Rest	28	2036	50	0
		Frg.	18	27	1516		1	Frg.	23	21	1516	2
Run	0	0	19	50	Run	0	0	9	60			
2	<i>Actual</i>	<i>Predicted</i>				<i>Actual</i>	<i>Predicted</i>					
			Vig.	Rest	Frg.		Run		Vig.	Rest	Frg.	Run
		Vig.	3834	37	24		1	Vig.	3748	118	30	0
		Rest	76	44	0		0	Rest	76	44	0	0
		Frg.	5	59	5251		0	Frg.	6	56	5253	0
Run	0	0	2	27	Run	0	0	3	26			
6	<i>Actual</i>	<i>Predicted</i>				<i>Actual</i>	<i>Predicted</i>					
			Vig.	Rest	Frg.		Run		Vig.	Rest	Frg.	Run
		Vig.	1509	69	61		0	Vig.	1513	60	66	0
		Rest	37	703	4		0	Rest	16	714	14	0
		Frg.	1	8	4411		18	Frg.	39	9	4386	4
Run	0	0	3	95	Run	0	0	14	84			
7	<i>Actual</i>	<i>Predicted</i>				<i>Actual</i>	<i>Predicted</i>					
			Vig.	Rest	Frg.		Run		Vig.	Rest	Frg.	Run
		Vig.	4694	0	86		5	Vig.	4759	0	26	0
		Rest	41	95	18		2	Rest	55	70	30	1
		Frg.	0	5	3485		8	Frg.	35	2	3461	0
Run	0	0	1	39	Run	0	0	7	33			
11	<i>Actual</i>	<i>Predicted</i>				<i>Actual</i>	<i>Predicted</i>					
			Vig.	Rest	Frg.		Run		Vig.	Rest	Frg.	Run
		Vig.	1722	6	54		0	Vig.	1734	10	37	1
		Rest	167	492	2		0	Rest	306	344	11	0
		Frg.	56	3	5339		0	Frg.	85	19	5287	7
Run	0	0	22	55	Run	0	0	14	63			

Table S4.8. Individual-wise confusion matrices for accelerometer- ([10]) and magnetometer-based (right) behaviour recognition. The entries in each confusion matrix here are arranged according to the row and column headers in Table S4.4).

Appendix S4.5: Variation in the Norm of Magnetometer Data

This section reports the variation in the norm of magnetic field intensity recorded by the animal-borne magnetometer. While the norm of calibrated tri-axial magnetometer data was equal to 1 towards the start of the recording (Figure S1), norm values did not remain at 1 throughout, and varied irregularly during the duration of the recording. This variation in magnetic field intensity is different from the one due to projection of a Cartesian grid on a three-dimensional Earth described by [35], since it was observed over the scale of a few metres. We speculate that variation across such short length-scales could only have been caused due to unexpected local magnetic interferences, such as metallic instruments being carried by the person following the animal being recorded (e.g. video-camera, radio receiver, VHF tracking device), the occasional proximity of the animal to fences and power-lines, and the proximity of our study site to the world's largest land-based manganese field, the base of which is reported to be an iron formation bed formed from ferromagnetic magnetite crystals ([36], <http://wwwu.edu.uni-klu.ac.at/mmessner/sites/rsa/kuruman/kuruman.htm#f2a>). While this seemingly arbitrarily changing magnetic field intensity could have aided the detection of dynamic behaviour because of the coupled effect of changing recorded values due to both movement and external magnetic field intensity change, it could potentially render comparison between postural estimations at different locations difficult. This is because different ranges of recorded values would be obtained depending on the extent of soft-iron and hard-iron effects [37].

Analysis based on all 10,818,151 triaxial magnetometer data points recorded in this study indicates that the overall calibrated magnetometer norm was 1.18 ± 0.13 (mean \pm SE) (Table S4.9).

We next demonstrate in Figure S4.1 the effect of calibration based on [25] by plotting in three-dimensional space the point-clouds obtained pre- and post-calibration. Figure S4.2 shows a histogram (computed with 200 bins) of calibrated magnetic field intensity norm values for data from all eleven recording sessions, and Figure S4.3 presents histograms (computed with 200 bins) of recording session-wise data.

Recording Session Number	Mean \pm std of calibrated magnetometer norm	Number of tri-axial magnetometer samples
1	0.93 \pm 0.35	1127349
2	0.82 \pm 0.17	1111204
3	0.96 \pm 0.18	1056316
4	0.93 \pm 0.14	1071018
5	1.26 \pm 0.12	1039226
6	1.66 \pm 0.37	1051605
7	1.30 \pm 0.33	1099111
8	2.09 \pm 0.30	573149
9	0.81 \pm 0.20	776537
10	1.50 \pm 0.23	881126
11	0.72 \pm 0.15	1031510
Mean \pm standard error of calibrated magnetometer norm across recording sessions: 1.18 \pm 0.13		Total Samples: 10,818,151

Table S4.9. Characterising the variation in calibrated magnetometer norm across recording sessions.

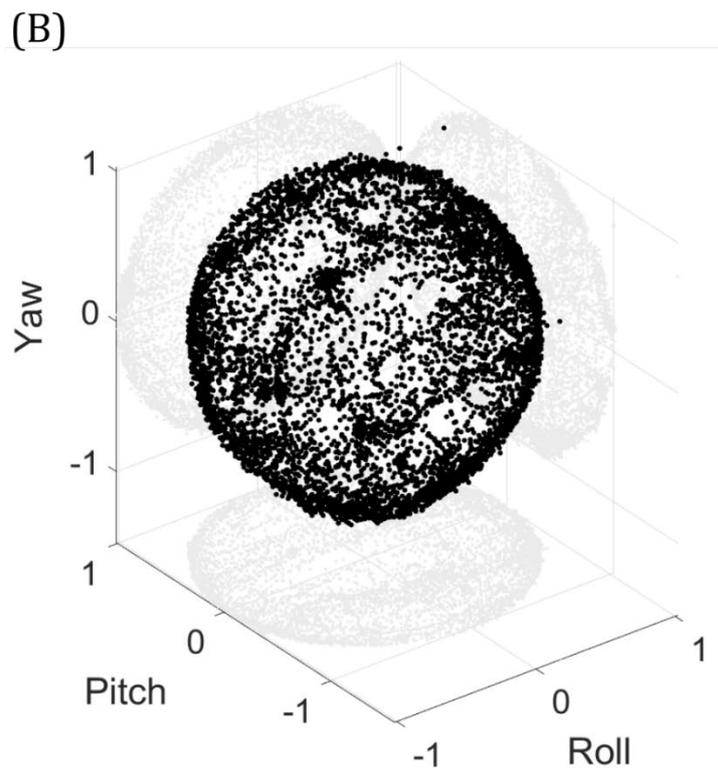
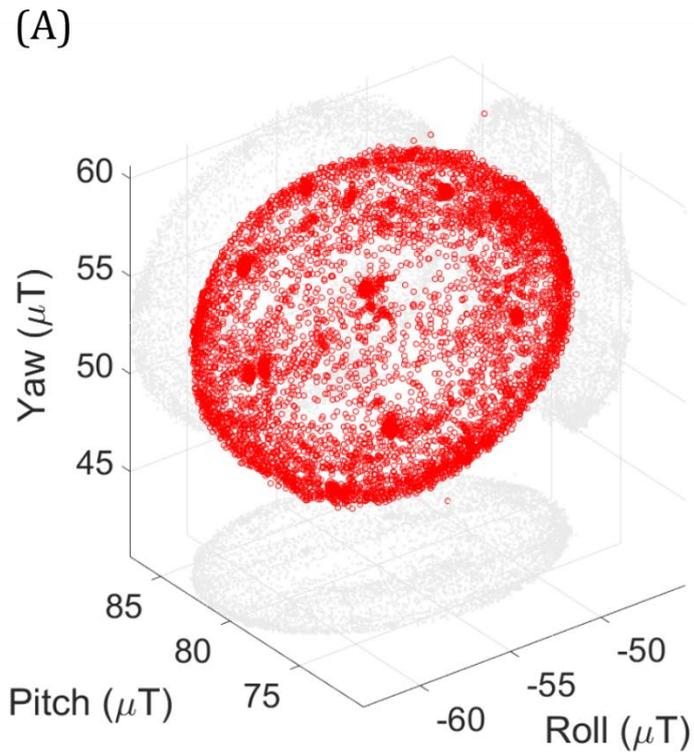


Figure S4.1. **Magnetometer calibration.** (A) Uncalibrated data in units of micro-Tesla recorded during the calibration procedure for recording session #10, affected by hard- and soft-iron effects, shown in red hollow circles. (B) The same data after calibration shown in normalised units of magnetic field strength, shown in black solid circles. In both cases, the two-dimensional (2D) projections of the 3D data cluster upon the three mutually perpendicular axis planes are shown in gray circles to permit plane-wise visualisation.

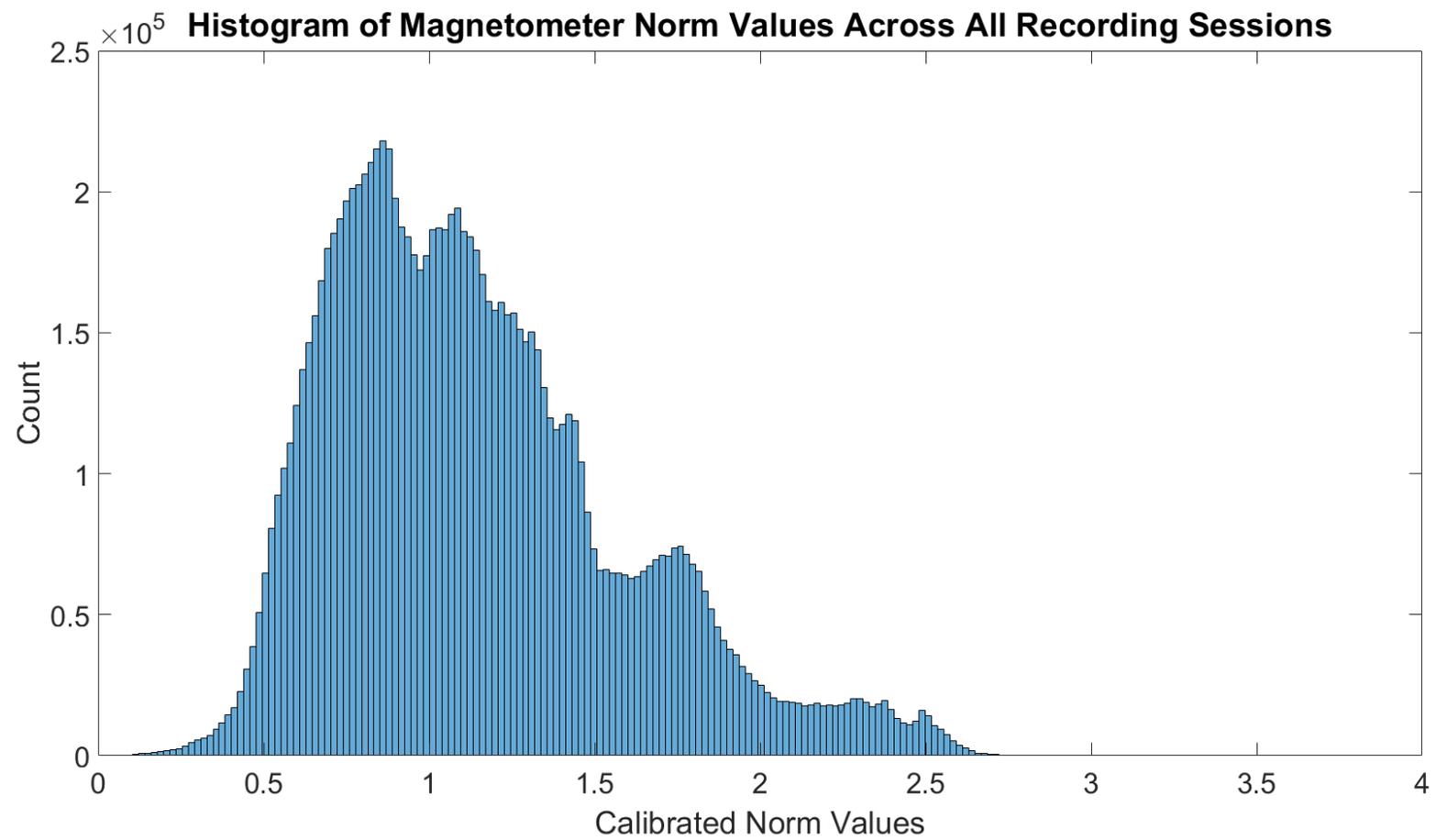


Figure S4.2. Variation in calibrated magnetometer norm computed across all recording sessions.

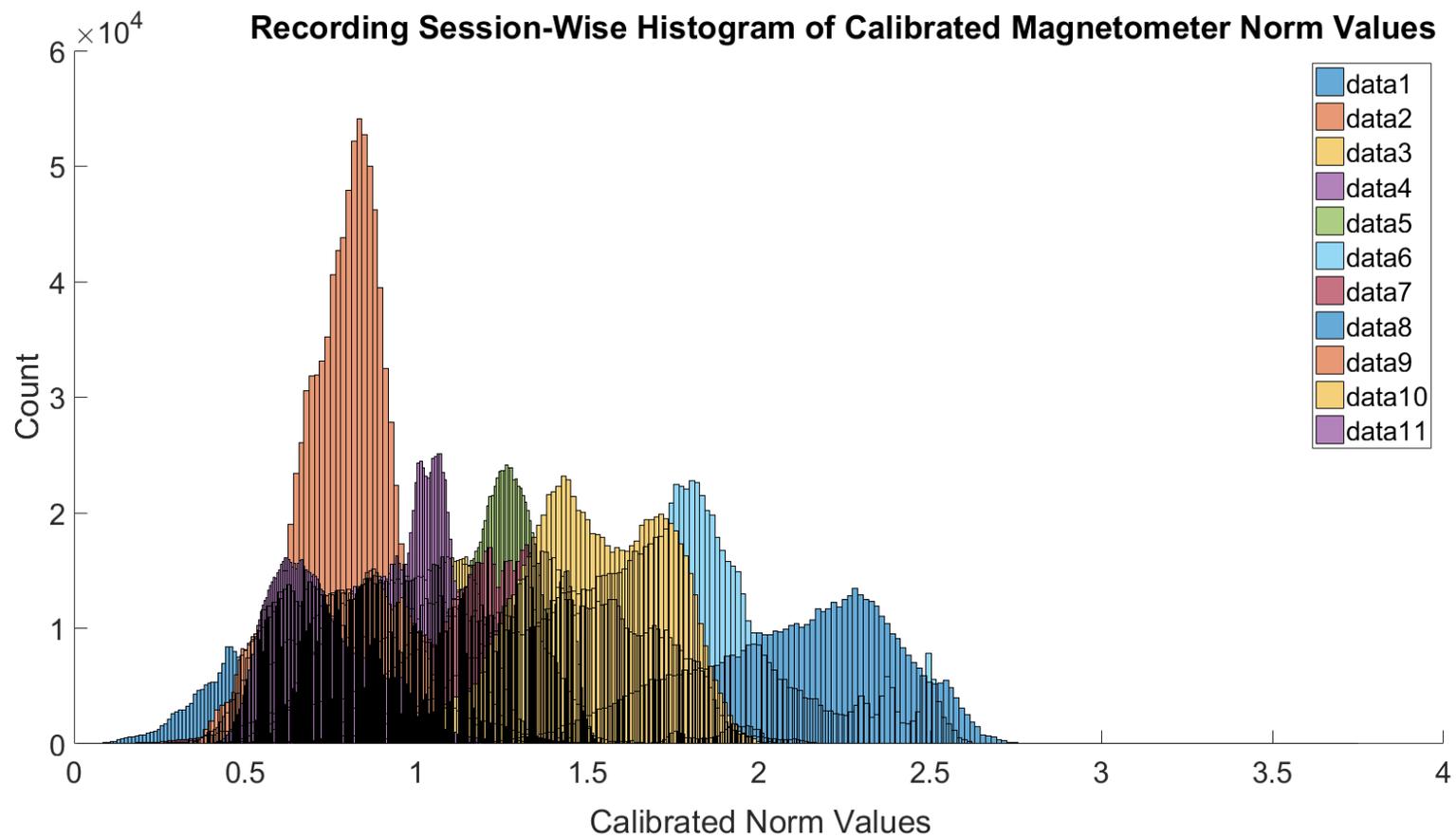


Figure S4.3. Variation in calibrated magnetometer norm for each of the eleven recording sessions.

References

1. Ropert-Coudert Y, Wilson RP. Trends and perspectives in animal-attached remote sensing. *Frontiers in Ecology and the Environment*. 2005 Oct;3(8):437-44.
2. Wilson RP, Shepard EL, Liebsch N. Prying into the intimate details of animal lives: use of a daily diary on animals. *Endangered Species Research*. 2008 Jan 18;4(1-2):123-37.
3. Noda T, Kawabata Y, Arai N, Mitamura H, Watanabe S. Animal-mounted gyroscope/accelerometer/magnetometer: In situ measurement of the movement performance of fast-start behaviour in fish. *Journal of experimental marine biology and ecology*. 2014 Feb 1;451:55-68.
4. Shepard EL, Wilson RP, Quintana F, Laich AG, Liebsch N, Albareda DA, Halsey LG, Gleiss A, Morgan DT, Myers AE, Newman C. Identification of animal movement patterns using tri-axial accelerometry. *Endangered Species Research*. 2008 Mar 31;10:47-60.
5. Williams HJ, Holton MD, Shepard EL, Largey N, Norman B, Ryan PG, Duriez O, Scantlebury M, Quintana F, Magowan EA, Marks NJ. Identification of animal movement patterns using tri-axial magnetometry. *Movement ecology*. 2017 Dec;5(1):6.
6. Naito Y, Bornemann H, Takahashi A, McIntyre T, Plötz J. Fine-scale feeding behavior of Weddell seals revealed by a mandible accelerometer. *Polar Science*. 2010 Aug 1;4(2):309-16.
7. Nathan R, Spiegel O, Fortmann-Roe S, Harel R, Wikelski M, Getz WM. 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*. 2012 Mar 15;215(6):986-96.
8. Cozzi G, Broekhuis F, McNutt JW, Turnbull LA, Macdonald DW, Schmid B. Fear of the dark or dinner by moonlight? Reduced temporal partitioning among Africa's large carnivores. *Ecology*. 2012 Dec 1;93(12):2590-9.
9. Hammond TT, Springthorpe D, Walsh RE, Berg-Kirkpatrick T. Using accelerometers to remotely and automatically characterize behavior in small animals. *Journal of Experimental Biology*. 2016 Jun 1;219(11):1618-24.

10. Chakravarty, P., Cozzi, G., Ozgul, A., & Aminian, K. (2019). A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods in Ecology and Evolution*, *10*(6), 802-814.
11. Williams HJ, Shepard EL, Duriez O, Lambertucci SA. Can accelerometry be used to distinguish between flight types in soaring birds?. *Animal Biotelemetry*. 2015 Dec;3(1):45.
12. Bouten CV, Sauren AA, Verduin M, Janssen JD. Effects of placement and orientation of body-fixed accelerometers on the assessment of energy expenditure during walking. *Medical and biological engineering and computing*. 1997 Jan 1;35(1):50-6.
13. Amini N, Sarrafzadeh M, Vahdatpour A, Xu W. Accelerometer-based on-body sensor localization for health and medical monitoring applications. *Pervasive and mobile computing*. 2011 Dec 1;7(6):746-60.
14. Wilson RP. Dead reckoning: a new technique for determining penguin movements at sea. *Meeresforsch..* 1988;32:155-8.
15. Wilson RP, Wilson MP, Link R, Mempel H, Adams NJ. Determination of movements of African penguins *Spheniscus demersus* using a compass system: dead reckoning may be an alternative to telemetry. *Journal of Experimental Biology*. 1991 May 1;157(1):557-64.
16. Mitani Y, Sato K, Ito S, Cameron MF, Siniff DB, Naito Y. A method for reconstructing three-dimensional dive profiles of marine mammals using geomagnetic intensity data: results from two lactating Weddell seals. *Polar Biology*. 2003 May 1;26(5):311-7.
17. Wilson RP. Movements in Adélie Penguins foraging for chicks at Ardley Island, Antarctica: circles within spirals, wheels within wheels. *Polar Bioscience*. 2002(15):75-87.
18. Wilson RP, Liebsch N, Davies IM, Quintana F, Weimerskirch H, Storch S, Lucke K, Siebert U, Zankl S, Müller G, Zimmer I. All at sea with animal tracks; methodological and analytical solutions for the resolution of movement. *Deep Sea Research Part II: Topical Studies in Oceanography*. 2007 Feb 1;54(3-4):193-210.
19. Matsumura M, Watanabe YY, Robinson PW, Miller PJ, Costa DP, Miyazaki N. Underwater and surface behavior of homing juvenile northern elephant seals. *Journal of Experimental Biology*. 2011 Feb 15;214(4):629-36.

20. Kunze K, Bahle G, Lukowicz P, Partridge K. Can magnetic field sensors replace gyroscopes in wearable sensing applications?. In International Symposium on Wearable Computers (ISWC) 2010 2010 Oct 10 (pp. 1-4). IEEE.
21. Abdulla UA, Taylor K, Barlow M, Naqshbandi KZ. Measuring walking and running cadence using magnetometers. In 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications 2013 Jul 16 (pp. 1458-1462). IEEE.
22. Clutton-Brock TH, Manser M. Meerkats: cooperative breeding in the Kalahari. *Cooperative breeding in vertebrates*. 2016 Jan 7:294-317.
23. Semiconductor F. Xtrinsic MAG3110 Three-Axis, Digital Magnetometer. MAG3110 Datasheet. 2013 Feb.
24. Thébault E, Finlay CC, Beggan CD, Alken P, Aubert J, Barrois O, Bertrand F, Bondar T, Boness A, Brocco L, Canet E. International geomagnetic reference field: the 12th generation. *Earth, Planets and Space*. 2015 Dec;67(1):79.
25. Bonnet S, Bassompierre C, Godin C, Lesecq S, Barraud A. Calibration methods for inertial and magnetic sensors. *Sensors and Actuators A: Physical*. 2009 Dec 1;156(2):302-11.
26. Usui S, Amidror I. Digital low-pass differentiation for biological signal processing. *IEEE Transactions on Biomedical Engineering*. 1982 Oct(10):686-93.
27. Zivanovic M, Carlosena A. On asymmetric analysis windows for detection of closely spaced signal components. *Mechanical systems and signal processing*. 2006 Apr 1;20(3):702-17. [references numbers starting from old26 to be updated to new]
28. Wilson JW, Mills MG, Wilson RP, Peters G, Mills ME, Speakman JR, Durant SM, Bennett NC, Marks NJ, Scantlebury M. Cheetahs, *Acinonyx jubatus*, balance turn capacity with pace when chasing prey. *Biology letters*. 2013 Oct 23;9(5):20130620.
29. Katevas K, Haddadi H, Tokarchuk L. Sensingkit: Evaluating the sensor power consumption in ios devices. In 2016 12th International Conference on Intelligent Environments (IE) 2016 Sep 14 (pp. 222-225). IEEE.

30. Martiskainen P, Järvinen M, Skön JP, Tiirikainen J, Kolehmainen M, Mononen J. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Applied animal behaviour science*. 2009 Jun 1;119(1-2):32-8.
31. Wilson RP, White CR, Quintana F, Halsey LG, Liebsch N, Martin GR, Butler PJ. Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *Journal of Animal Ecology*. 2006 Sep 1;75(5):1081-90.
32. McGregor SJ, Busa MA, Yaggie JA, Bollt EM. High resolution MEMS accelerometers to estimate VO₂ and compare running mechanics between highly trained inter-collegiate and untrained runners. *PloS one*. 2009 Oct 6;4(10):e7355.
33. Bourke AK, O'brien JV, Lyons GM. Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait & posture*. 2007 Jul 1;26(2):194-9.
34. Kemp B, Janssen AJ, Van Der Kamp B. Body position can be monitored in 3D using miniature accelerometers and earth-magnetic field sensors. *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control*. 1998 Dec 1;109(6):484-8.
35. Bidder, O. R., Walker, J. S., Jones, M. W., Holton, M. D., Urge, P., Scantlebury, D. M., ... & Wilson, R. P. (2015). Step by step: reconstruction of terrestrial animal movement paths by dead-reckoning. *Movement ecology*, 3(1), 23.
36. Cairncross, B., & Dixon, R. (accessed in March, 2018). Kalahari Manganese Field (Kuruman area) (RSA). <http://www.wu.edu.uni-klu.ac.at/mmessner/sites/rsa/kuruman/kuruman.htm#f2a>.
37. Caruso, M. J. (2000, March). Applications of magnetic sensors for low cost compass systems. In *Position Location and Navigation Symposium* (pp. 177-184).

CHAPTER 5

Fine-Scale Behaviour Recognition Using Accelerometers⁵

⁵ This chapter is currently undergoing minor revisions at *Methods in Ecology and Evolution*, and was submitted as, “Chakravarty P., Cozzi G., Dejnabadi H., Léziart P.A., Manser M., Ozgul A., Aminian K. *Seek and learn: Automated identification of microevents in animal behaviour using envelopes of acceleration data and machine learning.*”

Abstract

1. Animal-borne accelerometers have been used across more than 120 species to infer biologically significant information such as energy expenditure and broad behavioural categories. While the accelerometer's high sensitivity to movement and fast response times present the unprecedented opportunity to resolve fine-scale behaviour, leveraging this opportunity will require overcoming the challenge of developing general, automated methods to analyse the nonstationary signals generated by nonlinear processes governing erratic, impulsive movement characteristic of complex behaviour.
2. We address this issue by conceptualising complex behaviour in terms of characteristic microevents: impulsive movements producing brief (<1 s) shock signals in accelerometer data. We propose a novel 'seek-and-learn' approach: a microevent detection step first locates where shock signals occur ('seek') by searching for peaks in envelopes of acceleration data. Robust machine learning ('learn') employing meaningful features then separates microevents. We showcase the application of our method on triaxial accelerometer data collected on ten free-living meerkats (*Suricata suricatta*) for four fine-scale foraging behaviours – searching, one-armed digging, two-armed digging, and head jerks during prey ingestion. Annotated videos served as groundtruth, and performance was benchmarked against that of a variety of classical machine-learning approaches.
3. Microevent identification (μEvId) with eight features in a three-node hierarchical classification scheme employing logistic regression at each node achieved a mean overall accuracy of >85% during leave-one-individual-out cross-validation, and exceeded that of the best classical machine learning approach by 9%. μEvId was found to be robust not only to inter-individual variation but also to large changes in model parameters.
4. Our results show that microevents can be modelled as impulse responses of the animal body-and-sensor system. The microevent detection step retains only informative regions of the signal, which results in the selection of discriminative features that reflect biomechanical differences between microevents. Moving-window-based classical machine learning approaches lack this prefiltering step, and were found to be suboptimal for capturing the

nonstationary dynamics of the recorded signals. The general, automated technique of μEvid , together with existing models that can identify broad behavioural categories, provides future studies with a powerful toolkit to exploit the full potential of acceleration data for animal behaviour recognition.

5.1 | Introduction

Fine-scale information obtained through animal-attached sensors recording at high frequency has proven to be invaluable for understanding behaviour, the environment, and the relationship between the two (Ropert-Coudert & Wilson 2005; Kays et al. 2015) across more than 120 species (Brown et al. 2013). By measuring both a dynamic component of movement indicative of activity intensity (Müller & Schrader 2003; Wilson et al. 2006), and a static component due to Earth's gravity indicative of posture (Hansson et al. 2001), accelerometers have enabled inference of broad behavioural categories such as locomotion, resting, and foraging (e.g. Nathan et al. 2012; Chakravarty et al. 2019a). Accelerometers have high sensitivity to movement, measuring acceleration magnitudes ranging from small fractions of Earth's gravity ($g = 9.81 \text{ m/s}^2$) to several thousand times g , and fast response times, enabling sampling at several thousand Hertz (Béliveau et al. 1999). These features present the unprecedented opportunity to resolve fine movement, and thus fine-scale behaviour. However, leveraging this opportunity will require overcoming the challenge of developing general, automated methods to analyse the complex signals generated by fine-scale behaviour.

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

Sudden muscle activation by the animal or induced sensor impact during a microevent is akin to an impulse given to the animal body-and-sensor system, and the brief (<1 s) shock signal recorded by the accelerometer akin to the system's impulse response. The impulse response of a dynamical system measures the time profile of the effect of shocks at a given point in time on the expected future values

of variables in the system (Pesaran & Shin 1998). Knowing the impulse response is invaluable since it completely characterises a linear, time-invariant system (Phillips, Parr & Riskin 2003). Sensor-bearing animal bodies, however, are nonlinear, time-varying systems: the impulse response varies unpredictably with time. This is in part due to body tissue viscoelasticity (Hill 1938), but perhaps mainly due to nonlinear movement artefacts arising from sensors that are attached to the skin (e.g. for tape-, clamp- or glue-on tags) or loosely suspended (e.g. collar, leg bracelet) (Brown et al. 2013). The local mobility of skin-attached markers in human biomechanics studies, termed soft tissue artefact, is known to lead to errors in recorded trajectories that can be of the same order of magnitude as the underlying joint motions (Camomilla, Dumas & Cappozzo 2017). Contact forces arising from collisions of loosely suspended sensors with the body during dynamic motion vary nonlinearly with the shape, surface conditions, and mechanical properties of the contacting bodies (Koshy, Flores & Lankarani 2013). These nonlinearities yield nonstationary signals that are difficult to analyse using common techniques such as statistical or Fourier analysis.

Existing methods to recognise animal behaviour from acceleration data fall under two broad categories: automated and semiautomated. In automated recognition, descriptive features are first computed from a moving window of fixed size that is slid across acceleration data. Different behaviours are then separated using thresholds (e.g. Viviant et al. 2010; Watanabe et al. 2013) or machine-learning algorithms (e.g. Martiskainen et al. 2009; Nathan et al. 2012). Features are commonly statistical summaries of acceleration data, such as mean, standard deviation, and skewness, but have also been derived from signal spectra computed through wavelet transformation (Sakamoto et al. 2009). The preset window size for feature computation in these methods precludes identification of individual microevents that may have occurred within a window. Semiautomated recognition involves manually characterising signal patterns followed by classifying using a decision tree (Wilson et al. 2018b). While this approach does not require a fixed-size moving window to operate, the authors report that it requires appreciable investment in time and understanding (Wilson et al. 2018b). Further, time-varying sensor impact characteristics may make it difficult to avoid subjectivity in the process of manual characterisation of waveforms recorded for the same type of microevent.

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

In this study, we present a novel method for automatically identifying behavioural microevents that combines considerations of the physics of animal body-and-sensor systems, the technique of signal enveloping, feature engineering, and robust machine learning. In our method, we: (1) use enveloping to automatically detect microevents, (2) engineer meaningful features to characterise microevents, and (3) use robust machine-learning algorithms to classify microevents. We showcase the application of our method on data collected on 10 wild-living meerkats (*Suricata suricatta*), where fine-scale foraging behaviour such as searching, one-armed and two-armed digging, and prey ingestion were groundtruthed against annotated videos. We compare the performance of microevent identification (μEvId) against that of a variety of classical machine-learning approaches, and present an analysis of the robustness of μEvId performance to changes in model parameters.

5.2 | Material and Methods

5.2.1 | Microevent identification

5.2.1.1 | Detecting microevents in recorded acceleration

Our microevent (μE) detection method involves: (i) selecting the accelerometer axis on which μE 's will be detected, (ii) computing signal amplitude, and (iii) segmenting the amplitude to find locations and spans of μE 's.

Axis selection and pre-filtering

The first step is to choose one axis of the accelerometer that best records μE 's. There is general consensus on names for acceleration directions with respect to the animal's body: surge (anterior-posterior), sway (medial-lateral), and heave (inferior-superior) (Nathan et al. 2012). Surge acceleration is a natural choice to study animal movement dynamics, and has been preferred when deploying uniaxial accelerometers on terrestrial (Watanabe et al. 2005), avian (Sakamoto et al. 2009) and marine (Iwata et al. 2012) animals in the past. Further, it has been shown to be sensitive to fine-scale behaviours such as head movements during prey capture (e.g. Suzuki et al. 2009; Viviant et al. 2010). This step is followed by pre-filtering data from this axis. Low-pass filtering to remove unwanted sensor- and analog-to-digital signal quantization-generated noise (Usui & Amidror 1982) should suffice for most applications, for instance when bursts of activity associated with a μE of interest are separated by short periods of relative inactivity. If, however, the μE of interest is performed during other dynamic activity of similar intensity, for instance, prey-capture attempts during swimming in marine predators (e.g. Viviant et al. 2010), pre-filtering should involve band-pass or band-stop filtering to suppress the signal component corresponding to locomotion. The latter works because the impulse-like μE is composed of several frequencies whereas locomotion is carried out in a relatively tight frequency range for a given species.

Computing signal amplitude

Raw upper and lower envelopes of the filtered acceleration signal s_{acc} are first computed using the maximum and minimum of s_{acc} in a moving window of size M_{env} samples centred at each sample of s_{acc} (Figure 5.1A). A low-pass filter of cut-

off frequency f_c is then applied to obtain smoothed upper (U_{acc}) and lower (L_{acc}) envelopes. Finally, signal amplitude A_{acc} is computed according to equation (1):

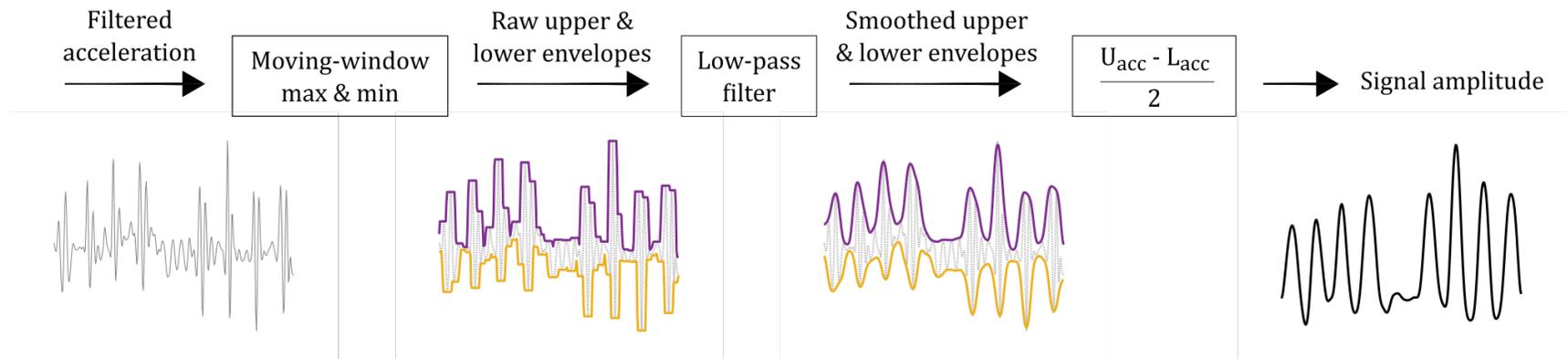
$$A_{acc} = \frac{U_{acc} - L_{acc}}{2} \quad (1)$$

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

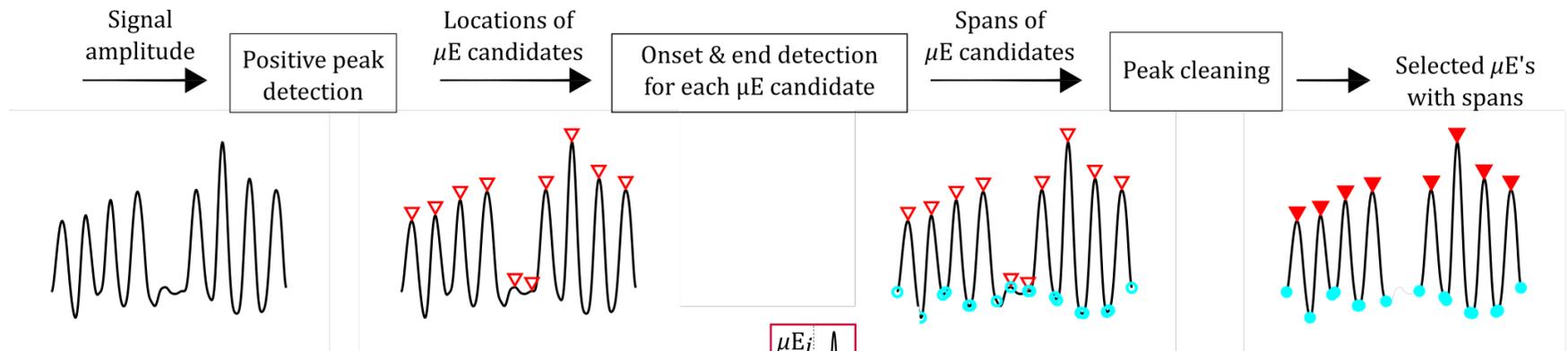
Segmenting amplitude to find microevents

Positive peaks in A_{acc} are first found and designated as possible μE candidates (Figure 5.1B). The span of each μE candidate is defined as being the portion of A_{acc} starting from μE 's onset to its end. To find estimates of the onset and end of each μE candidate, two thresholds are used – an amplitude threshold th_{amp} , and a standard-deviation threshold th_{std} computed on a moving window of size M_{seg} (can be chosen to be equal to M_{env} by default) samples – in conjunction with negative peak detection (Appendix S5.1). The two thresholds are used to filter out regions of A_{acc} that belong to ‘non-event zones’, i.e. zones which cannot contain μE 's of interest because A_{acc} is either nearly constant, slowly rising or slowly decaying in such regions. Custom MATLAB code (`segmentAmp`) to carry out amplitude segmentation is supplied as Supplementary Information.

(A) Computing signal amplitude from filtered acceleration



(B) Segmenting amplitude to detect microevents (μE)



(C) Defining μE context for feature computation

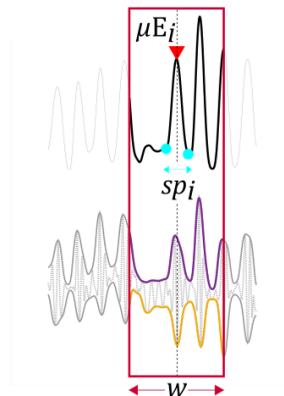


Figure 5.1. **Microevent detection and characterisation.** (A) Smoothed upper (U_{acc} , purple) and lower (L_{acc} , yellow) envelopes of filtered acceleration (grey) are first computed, and half of their sample-wise difference defined as signal amplitude (black), (B) signal amplitude (black) is segmented into individual microevents (solid red triangles) with spans given by each microevent's onset and end (solid cyan discs) after selecting from the pool of microevent candidates (locations: unfilled red triangles; spans: unfilled cyan circles) and removing non-event zones (grey), and (C) two neighbourhoods were considered for characterising microevents through feature computation: local, given by microevent μE_i 's span (solid cyan discs), and contextual, given by a window (dark red box) of length w seconds centred at μE_i 's location (solid red triangle). Features were computed from four signals per accelerometer axis: acceleration signal, its upper and lower envelopes, and amplitude.

5.2.1.2 | Feature development and selection

Both local and contextual data from each axis are used in this step to characterise each microevent μE_i (Figure 5.1C) through feature computation. Local data are defined as being delimited by the span sp_i of μE_i , and contextual data by a fixed window of length w seconds centred at μE_i 's amplitude peak. Local data capture short time scale manoeuvres performed by the animal during μE_i , and contextual data capture the longer time scale, background movement during which μE_i occurred. We drove the feature development process by first identifying four general feature categories: biomechanics (e.g. posture, intensity), frequency (e.g. periodicity, dominant frequency), coordination (e.g. coefficient of correlation between data from different axes), and pattern (e.g. signal tortuosity, number of zero-crossings, microevent's peak shape, statistical features). We created a non-exhaustive list of 200 candidate features across these four broad categories (Appendix S5.2) based on data from filtered acceleration, upper and lower envelopes, and signal amplitude for each accelerometer axis. Feature selection is then carried out by choosing the most discriminative feature from each sub-category (e.g. using `rankfeatures`, © MATLAB R2019a). This approach helps achieve an order-of-magnitude reduction in the number of features (e.g. from 200 to 10) while ensuring that meaningful features are selected.

5.2.1.3 | Classification

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

5.2.2 | Case study: Foraging in Kalahari meerkats

5.2.2.1 | Data collection and groundtruthing

Fieldwork was conducted as described in Chakravarty et al. 2019a. Data from ten adult meerkats foraging in their natural habitat were collected at the Kalahari Meerkat Project (Clutton-Brock & Manser 2016) from August 2016 to November 2017. The animals were habituated to a stage where it was possible to follow them within a distance of 1 m. Individuals bore collars equipped with an inertial measurement unit (IMU) (adapted version of Physilog IV, GaitUp SA, Switzerland) containing a triaxial accelerometer recording at a sampling frequency of $f_s = 100$ Hz/axis with a range of ± 16 g (± 156.96 m/s²) and 16-bit resolution. The size of the collar case (IMU and battery) was 35 mm x 29 mm x 19 mm, and overall weight was < 25 grams. Accelerometer calibration was conducted prior to each recording session according to Ferraris, Grimaldi & Parvis 1995. Each recording session lasted about 3 continuous hours in the morning during which each individual was filmed using an electronically synchronised handheld video camera recording at 25 frames/second (Chakravarty et al. 2019b). Videos were annotated by playing them in Solomon Coder (version: beta 17.03.22) and noting down start and end times of behaviours of interest (in Microsoft® Excel®, 2013).

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

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

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

Video annotation for these four behaviours served as groundtruth. We defined a three-node hierarchical classification scheme (Figure 5.2) to separate chewing from the other three activities at the first node, then digging (1AD and 2AD) from searching at the second node, and finally 1AD from 2AD at the third node. The rationale behind this classification structure was that firstly, chewing primarily involved head motion whereas the other three activities primarily involved limb motion. Secondly, 1AD and 2AD are both forms of ‘digging’ whereas searching primarily involves walking-turning-pausing.

5.2.2.3 | Applying the microevent identification algorithm

Axis selection and signal amplitude computation

We chose the surge axis for microevent detection. We filtered surge acceleration with a Butterworth low-pass filter of order 4 and cut-off frequency 25 Hz. In line with the considerations in Section 5.2.1.2, and given that meerkat running frequency was observed to be ~4 Hz (Chakravarty et al. 2019a), we computed and

smoothed envelopes with $M_{env} = 11$ samples, and a Butterworth low-pass filter of order 4 and cut-off frequency $f_c = 5$ Hz, respectively.

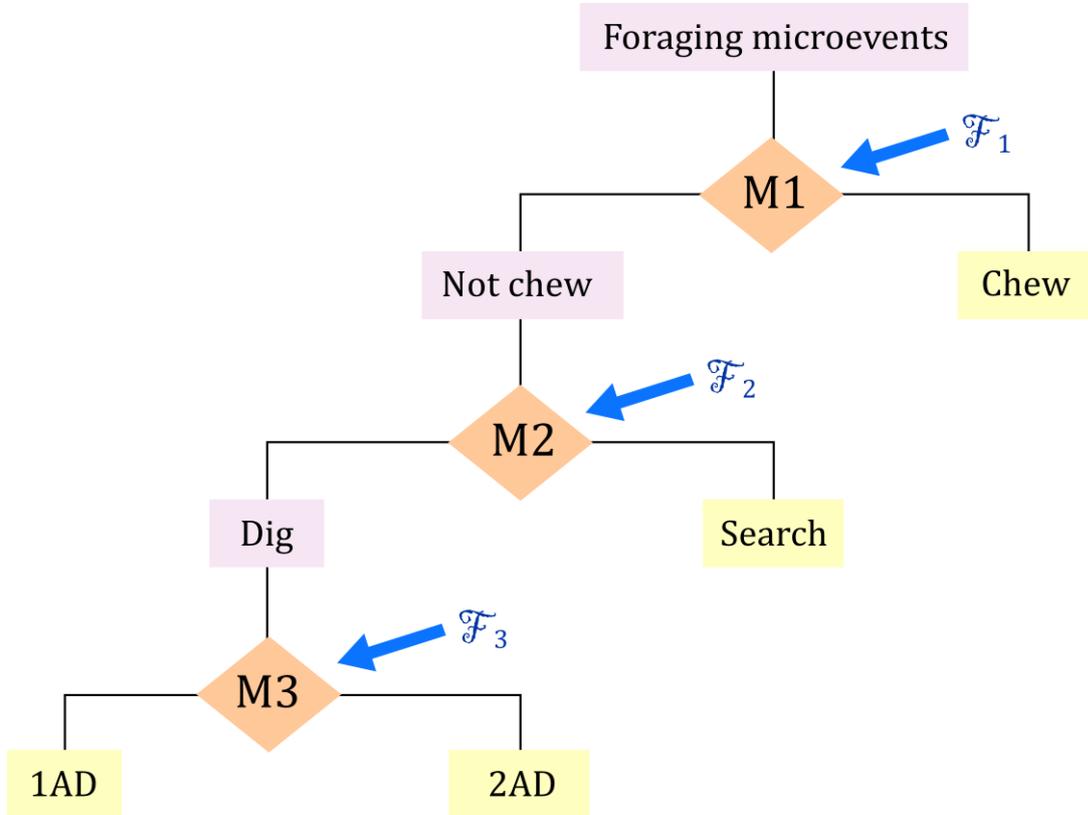


Figure 5.2. **Microevent classification scheme.** A hierarchical, tree-like classification scheme using one machine-learning algorithm (M_i , $i \in \{1,2,3\}$) and feature set (\mathcal{F}_i , $i \in \{1,2,3\}$) per node was used to identify microevents occurring during fine-scale meerkat foraging behaviours. 1AD: one-armed digging; 2AD: two-armed digging.

Amplitude segmentation

For amplitude segmentation, we chose an amplitude threshold of $th_{amp} = 0.1$ g (0.98 m/s^2), and standard deviation threshold of $th_{std} = 0.01$ g (0.098 m/s^2) with a moving window of size $M_{seg} = 11$ samples ($= M_{env}$ by default). The choices of th_{amp} and th_{std} are conservative ones, designed to select the maximum number of microevents from a signal while rejecting spurious peaks and zones in the signal where short durations of static activity occur, e.g. pauses between successive head jerks during chewing, or between successive bursts of 1AD.

Feature development and selection

We chose a window of size $w = 1.5$ seconds for feature computation to ensure that the window would be long enough to include sufficient context before and after the microevent, but short enough to capture enough data since meerkat behaviours are typically short-lived. We computed 200 features from triaxial acceleration across nine sub-categories (*cf.* Section 5.2.1.3): posture, intensity, periodicity, dominant frequency, band-specific frequency content, microevent amplitude-peak characteristics, coordination between surge and heave, tortuosity, and statistical. We selected the best features from each sub-category and node (Figure 5.2) using five filter feature-ranking methods (Appendix S5.3). Further, we computed nine sets of features and compared them by varying the microevent neighbourhood (local only, contextual only, both local and contextual), and signals from which features were computed (acceleration only, envelope-based only, both acceleration and envelope-based). Finally, to compare results with fewer features, we computed an additional set of results by using only the features selected for the first node at all three nodes.

Classification

At each node, we tested five machine learning algorithms: Naïve-Bayes (NB), Linear Discriminant Analysis (LDA), Logistic Regression (LR), Support Vector Machine with linear kernel (SVM), and Decision Tree (DT). Five model combinations, M1-M2-M3 (Figure 5.2) resulted from these upon application of the same algorithm at each node, e.g. LR-LR-LR.

5.2.2.4 | Cross-validation

To ensure that the inevitable imbalance in the amount of data collected per individual did not bias model training, we balanced the dataset so that each behavioural class had an equal amount of data from each individual (Appendix S5.4). For this, we resampled the dataset through a combination of oversampling rare classes using the Synthetic Majority Oversampling Technique (SMOTE) and randomly undersampling over-represented classes (Chawla et al. 2002) (Appendix S5.4). Thereafter, we performed leave-one-individual-out cross-validation (LOIO) on the balanced dataset. Models were trained using data pooled over all individuals

except one, and tested on data from the individual left out. This process was repeated until each individual had been the ‘test’ individual once. To prevent specific outputs of random number generation in SMOTE from potentially biasing results, we averaged results over an ensemble of ten balanced datasets for each cross-validation run. In this process, the ten resulting confusion matrices for each test individual (one confusion matrix per balanced dataset) were aggregated, and thirteen model performance metrics were computed as in Chakravarty et al. 2019a (Appendix S5.5): three behaviour-specific metrics (sensitivity, precision, and specificity) for each of the four behaviours and overall model accuracy. Finally, the effect of dataset balancing was assessed by comparing LOIO results obtained with the balanced dataset with those obtained with the imbalanced dataset.

5.2.2.5 | Alternative classification strategies: benchmarking against classical machine learning approaches

We benchmarked microevent identification performance against sliding window-based classification using classical machine learning (CML). We considered two machine learning algorithms – Random Forest (RF, 500 trees), and Support Vector Machine (SVM, linear kernel) – along with two feature families: (1) the 38 summary statistics presented in Nathan et al. 2012 (Table S5.4) – such as statistical moments (mean, standard deviation), autocorrelation and trend of data from triaxial acceleration and vectorial norm, and (2) 131 of the 200 features developed in this study (Table S5.2) that did not require microevent-related information. To ensure that models being compared had the same complexity in terms of number of features employed, we chose the top n_F features from each feature family (Appendix S5.6), where n_F would be the number of features employed by the best microevent identification model.

CML has previously been used with small window sizes (0.3 s) to differentiate fine-scale behaviour such as prey handling from swimming in little penguins (*Eudyptula minor*) (Carroll et al. 2014). We included this possibility by computing benchmarking results with two window sizes: 1.5 s with 1 s overlap, and 0.3 s with 50% overlap between successive windows. We performed feature selection separately for the two window sizes (Appendix S5.6). Finally, we computed LOIO classification results with balanced as well as imbalanced datasets.

5.3 | Results

5.3.1 | Collected data and detected microevents

Data for all four behaviours were collected for eight of ten individuals (Table 5.1). The dataset was imbalanced both across individuals and behaviours: (i) 55% of the total duration of recorded behaviours came from three individuals (individuals #3, #6, and #10), and (ii) 50% of the duration of chewing data came from one individual (#3), and 39% of all 2AD came from one individual (#6). Recorded acceleration patterns for the microevents (μE 's) (Figure 5.3, left column) showed, as expected, multiple peaks with varying amplitudes and inter-peak intervals. μE 's lasted for <1 s: span lengths were 0.26 ± 0.08 s for 1AD, 0.30 ± 0.09 s for 2AD, 0.28 ± 0.09 s for searching, and 0.30 ± 0.07 s for chewing. 1AD was the 'busiest' activity with 2.1 ± 0.3 μE /second (mean \pm s.d. across individuals) while the rate was lower for the other activities: 1.5 ± 0.4 μE /s for searching, 1.5 ± 0.6 μE /s for 2AD, and 1.5 ± 0.6 μE /s for chewing.

Two advantages of enveloping can be observed visually: (i) by effectively 'sandwiching' filtered acceleration between the upper and lower envelopes, enveloping makes μE detection robust to changes in signal mean over time, such as during chewing (Figure 5.3, last row), and (ii) smoothing during enveloping allows μE detection even when microevent-generated signals differ slightly in duration, highly (2x) in amplitude, and in number of constituent peaks, as in 2AD (Figure 5.3, third row from top).

5.3.2 | Feature selection and data resampling

A set of eight features (Table S5.3) selected sub-category wise for the first node (and repeated at the two remaining nodes) given as input to an LR-LR-LR model (*cf.* Section 5.2.2.3 under '*Classification*') performed the best (Table 5.2, first row) for microevent identification. These eight features were derived from both local and contextual acceleration-only signals (*cf.* Section 5.2.2.3 under '*Feature development and selection*').

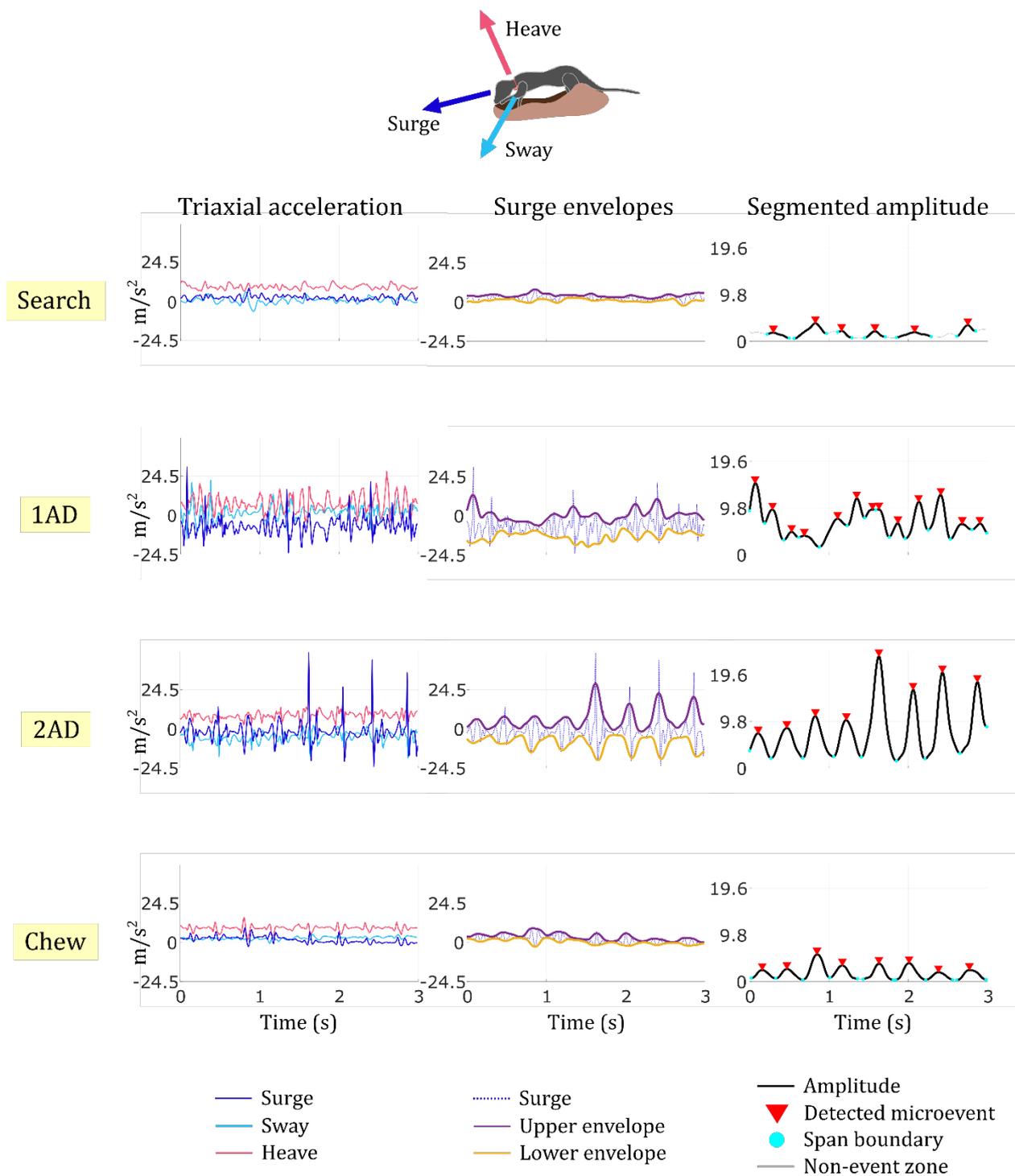


Figure 5.3. **Observed signals and detected microevents.** Each fine-scale meerkat foraging behaviour (rows) yielded nonstationary acceleration signals (left column). Smoothed upper and lower envelopes of surge acceleration (middle column) enabled computation of signal amplitude, which was segmented to detect microevents (right column). 1AD: one-armed digging; 2AD: two-armed digging.

For comparison with CML-based approaches, the top $n_F = 8$ features (*cf.* Section 5.2.2.5) from Nathan et al. 2012's feature family (Table S5.5 and Table S5.7) and those from the feature family developed in this study (Table S5.6 and Table S5.8) were selected. For LOIO, three individuals had to be left out because of insufficient data for at least one of the four behaviours (Table 5.1): individuals #2 and #7 did not have data from 1AD and 2AD, respectively, and individual #1 did not have enough chewing samples (SMOTE requires at least 6 samples from each class). Thereafter, the dataset with data from the retained individuals was resampled (*cf.* Section 5.2.2.4).

5.3.3 | Performance evaluation and comparison

For microevent identification (μEvId), means of all behaviour-specific performance metrics with μEvId were $\geq 80\%$, and mean overall accuracy across individuals was $>85\%$ (Table 5.2). Performance was robust across individuals, despite possible variations in digging style, sand compactness, digging site, and prey size. μEvId performed particularly well for chewing and 2AD, with high accuracy (means of all three behaviour-specific metrics $>89\%$ for chewing, and $>86\%$ for 2AD) and low standard deviation across individuals ($<6\%$ for chewing and $<9\%$ for 2AD across all three behaviour-specific performance metrics). Large ($>10\%$) standard deviations were observed for 1AD precision and Search sensitivity. This was primarily due to misclassification of Search as 1AD for a single individual (#4, Table S5.9). 1AD precision improved to $86.7 \pm 7.4\%$ and searching sensitivity to $87.6 \pm 7\%$ when this individual's metrics were excluded.

With balanced datasets, μEvId 's overall accuracy exceeded that of the best CML model by 9% (Table 5.2). For each CML model, there was at least one behaviour-specific performance metric that was $<70\%$. Further, μEvId was found to be more robust to inter-individual variation: standard deviations were $>10\%$ for only two of 13 performance metrics with μEvId , as opposed to eight of 13 performance metrics with the best CML model. For CML models, results using a longer window ($w = 1.5$ s) were consistently better than those with a short window ($w = 0.3$ s) regardless of feature family or machine learning algorithm employed (Table S5.11).

In the absence of data balancing, none of the models yielded a mean overall accuracy of >80% (Table S5.12).

Individual #	1AD sec ($n_{\mu E}$)	2AD sec ($n_{\mu E}$)	Search sec ($n_{\mu E}$)	Chew sec ($n_{\mu E}$)	Total/individual sec ($n_{\mu E}$)
1	21.7 (48)	11 (8)	30 (41)	6.5 (5)	69.2 (102)
2	0 (0)	34 (62)	37.2 (31)	20.4 (22)	91.6 (115)
3	60.4 (109)	61.1 (135)	64.8 (92)	327.7 (752)	514 (1088)
4	38.7 (70)	29.9 (23)	38.4 (73)	10.8 (11)	117.8 (177)
5	74.2 (151)	32.4 (58)	53.9 (92)	52.2 (85)	212.7 (386)
6	61.9 (136)	161.5 (379)	40.5 (63)	34.7 (44)	298.6 (622)
7	27.9 (53)	0 (0)	31.6 (54)	34 (26)	93.5 (133)
8	39.9 (94)	25.1 (39)	22.3 (33)	5.6 (14)	92.9 (180)
9	68.9 (157)	21.3 (29)	23.4 (19)	88.2 (171)	201.8 (376)
10	71.2 (184)	35.3 (22)	61.2 (120)	81 (120)	248.7 (446)
Total	464.8 (1002)	411.6 (755)	403.3 (618)	661.1 (1250)	1940.8 (3625)

Table 5.1. **Summary of groundtruth data and detected microevents.** The number of seconds and detected microevents for four fine-scale meerkat foraging behaviours are shown here. 1AD: one-armed digging; 2AD: two-armed digging; $n_{\mu E}$: number of detected microevents.

Finally, we tested $\mu Evid$'s sensitivity to changes in model parameter values (Table 5.3). The maximum drop Δ_{max} across mean values of the thirteen performance metrics during LOIO relative to the reference results (Table 5.2, top row) when each model parameter was varied through its sweeping range was used as

a measure of model sensitivity to parameter change. μEvId was found to be robust to large ($\geq \pm 33\%$) changes in each of the six model parameters (Table 5.3); detailed results are provided in Appendix S5.8. The largest sensitivity was to changes in M_{env} , with Δ_{max} occurring at $M_{env} = 5$ samples, and performance dropping again when $M_{env} > 11$ samples (Table S5.13). This justified the reasoning behind our choice of $M_{env} = 11$ samples (*cf.* Section 5.2.2.3). Sensitivity to changes in w was also relatively high, with results generally improving as w increased (Table S5.18), suggesting that features incorporating greater contextual information were more discriminative. This, along with low sensitivity to changes in amplitude segmentation-related parameters (th_{amp} , th_{std} and M_{seg}), suggests that the choices of window size for envelope computation (M_{env}) and context definition (w) are most critical to microevent identification performance.

5.4 | Discussion

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

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

Results from a previous study indicated that there was low separability between different meerkat foraging behaviours in terms of posture, intensity, and

periodicity (Chakravarty et al. 2019a). Results with CML in this study, too, indicate that even when the number of features is increased (from three in Chakravarty et al. 2019a to eight in the present study), separability between foraging behaviours remains relatively low. Yet, there is a large difference in performance between μE^{Id} and CML which is not explained by typical factors that influence behaviour recognition performance, such as feature selection, machine learning algorithm, class imbalance, and classification scheme.

Despite testing several combinations of selected features (two feature families, two window sizes) and machine-learning algorithms (Random Forest, Support Vector Machine), the overall accuracy of the best CML model was still lower than that of μE^{Id} by 9%. Balancing the amount of data from different individuals and classes prior to cross-validation removed differences arising from class imbalance in both cases (μE^{Id} , CML). Class balancing also nullified the greater ease with which hierarchical classification schemes (e.g. μE^{Id}) can deal with class imbalance compared to directly classifying all behaviours at once (CML) (Chakravarty et al. 2019a). Therefore, since no major differences are encountered during the classification process, the microevent detection step conducted prior to classification in μE^{Id} must be driving its improved class separability.

A fundamental methodological difference between the two approaches is that in μE^{Id} , windows (for feature computation) are defined only where microevents are detected ('informative' regions of the signal). Thus, the microevent detection step automatically rejects regions of the signal consisting of low-acceleration movements or smooth changes in acceleration uncharacteristic of microevents ('non-informative' regions). Further, centring the window at the detected microevent location provides the requisite signal context (before, after) within which the μE 's signal occurs. Features are thus computed only from informative regions of the signal since non-informative regions are filtered out. In contrast, in CML, a sliding window does not distinguish between informative and noninformative regions of the signal and, consequently, fails to choose the most discriminative features, leading to poorer class separation. For instance, *cxFFTpeakpowerHeave* (periodicity of heave acceleration) and *cxFFTpowerfracHeave* (fraction of heave signal in frequency range 2.5 – 5 Hz) (Table S5.3) were both higher for two-armed digging than for the other behaviours. Their selection reflects the visual observation that the trunk of a meerkat engaged in

Model	1AD			2AD			Search			Chew			Overall Accuracy (%)	
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)		
Microevent identification	83.5 ± 7.8	93.8 ± 4.8	83.0 ± 11.4	90.6 ± 8.7	95.0 ± 2.8	86.4 ± 6.2	80.0 ± 19.5	94.7 ± 4.0	84.3 ± 8.7	89.2 ± 5.7	97.6 ± 1.5	92.6 ± 4.3	85.8 ± 5.3	
Classical machine learning														
RF	$w = 1.5$ s, $fF2$	84.4 ± 9.6	92.1 ± 5.7	79.5 ± 11.2	77.8 ± 16.8	91.4 ± 13.6	82.9 ± 16.9	75.4 ± 21.6	92.7 ± 3.7	78.5 ± 8.6	66.7 ± 22.1	91.9 ± 6.7	75.2 ± 12.0	76.1 ± 10.0
SVM	$w = 1.5$ s, $fF1$	77.4 ± 14.2	93.7 ± 4.9	81.6 ± 10.4	68.1 ± 25.5	89.2 ± 7.9	73.4 ± 13.6	75.2 ± 17.8	90.6 ± 10.8	79.8 ± 18.1	86.4 ± 12.0	95.5 ± 3.6	87.7 ± 8.7	76.8 ± 6.2

Table 5.2. **Results with leave-one-individual-out cross-validation.** The performance of microevent identification (eight features with logistic regression) was compared with classical machine learning (CML) approaches employing the same number of features. For CML, results are shown for the best-performing combinations of moving-window size ($w = 1.5$ s and $w = 0.3$ s) and feature family ($fF1$ and $fF2$: sets of top eight features from feature family in Nathan et al. 2012 and this study, respectively) for two different machine learning algorithms (RF: Random Forest; SVM: Support Vector Machine, linear kernel).

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

Table 5.3. **Effect of varying microevent identification model parameters on classification performance.** The six parameters used in the microevent identification algorithm were varied and the maximum drop in mean values across thirteen performance metrics – sensitivity, specificity and precision for each of the four behaviours, and overall accuracy – during leave-one-individual-out cross validation used to provide a measure of model robustness to parameter change.

two-armed digging moves rhythmically – within a relatively tight frequency range – downward as the meerkat reaches for more sand to dig out, and upward as it flings dug-out sand backward. Neither of these features were selected by any of the CML models.

Limitations

Though $\mu\text{E}v\text{Id}$ does not impose restrictions on the duration of signal patterns of interest, it does require all patterns to have similar duration, e.g. <1 s for μE 's. Thus, it cannot be used to identify signal patterns across time scales, i.e. patterns whose durations differ by an order of magnitude, say 1 s – 10 s. This is because envelope computation in the μE detection step requires window size (M_{env} , Table 5.3) to be fixed: short windows would not be long enough to capture the long patterns of interest, whereas long windows would club together shorter patterns of interest within the same window. Available methods to identify patterns across scales involve signal characterisation by hand (Wilson et al. 2018b). Changes to sensor housing may affect $\mu\text{E}v\text{Id}$'s accuracy. This is because μE -generated impulse responses depend on the physical properties (geometry, material) of the animal body-and-sensor system, which would change depending on factors such as tag shape and weight, and stiffness and tightness of tag attachment. For meerkats, we prepared new housing for a cuboidal sensor case using duct tape, and a cable tie adjusted to each individual's neck diameter. Our results indicate that $\mu\text{E}v\text{Id}$ may nevertheless be robust to minor variations in sensor housing. Finally, $\mu\text{E}v\text{Id}$'s accuracy may be affected by changes in sensor orientation. This is because, unlike models where features are engineered to minimise the effect of changes in sensor orientation (Chakravarty et al. 2019a), features in $\mu\text{E}v\text{Id}$ include those computed from axes that can undergo large changes in orientation, e.g. the sensor's heave and sway axes can rotate around the axis of the meerkat's cylindrical neck. Nevertheless, $\mu\text{E}v\text{Id}$ was found to be robust to the minor collar rotations observed in our study.

Potential for other studies, and future work

The fact that a majority of $\mu\text{E}v\text{Id}$'s selected features (Table S5.3) were based on contextual acceleration (Figure 5.1C) suggests that including data from a time window longer than the signal pattern of interest gives more accurate and robust results than characterising the signal pattern alone. This finding could be leveraged to

define robust guidelines for manual waveform characterisation in semiautomated approaches (Wilson et al. 2018b). μEvId shows that algorithmic advances can simplify tag attachment, e.g. by possibly removing the need for multiple accelerometers (on the head and back) to measure rapid head movements during prey capture in Adélie penguins (*Pygoscelis adeliae*) (Watanabe et al. 2013). Our specificity analysis, i.e. the accuracy with which behaviours can be rejected when they are not present, showed that μEvId can accurately separate different types of confounding μE 's. High specificity is desirable since misclassifying confounding fine-scale behaviours can alter biological interpretation, e.g. detection of prey biting motions in Weddell seals (*Leptonychotes weddellii*) may be biased by head bobbing associated with vocalisation (Naito et al. 2010). Recent work investigated prey capture variability in Adélie penguins using recorded acceleration (Watanabe et al. 2013). The generality of μEvId will enable more studies to address such high-level biological questions. Series of identified microevents can be used as 'building blocks' to define bouts of higher-level behaviour. For instance, identification of successive chewing μE 's can be used to first define a bout of the higher-level behaviour, 'prey ingestion'. The duration of prey ingestion and feature values of its constituent μE 's can then be combined to predict prey size.

5.5 | Conclusion

We showed that robust, automated recognition of fine-scale animal behaviours from recorded acceleration can be achieved by conceptualising complex behaviour as being composed of characteristic microevents. Our findings demonstrate that transient acceleration patterns recorded during microevents can be modelled as impulse responses of the animal body-and-sensor system. Different microevents generate different impulse responses and are carried out in different signal contexts. In our microevent identification model, enveloping detects where impulse responses occur in the signal ('seeking'), and machine learning employing features that characterise both impulse response and signal context separates microevents ('learning'). The microevent detection step acts as a prefilter that selects informative, microevent-generated regions of the signal while rejecting

noninformative regions uncharacteristic of microevents. This prefiltering results in the selection of discriminative features that reflect biomechanical differences in the movements carried out by the animal during a microevent. The classical approach of sliding-window-based machine learning cannot separate informative and noninformative regions of the signal, and was found to be suboptimal for capturing the nonstationarity of signals recorded during complex animal behaviour. Microevent identification, together with existing models that can identify broad behavioural categories, provides future studies with a powerful toolkit to exploit the full potential of acceleration data for animal behaviour recognition.

Appendices

Appendix S5.1 | Further details on the amplitude segmentation algorithm

This section provides more details on the algorithm developed to segment signal amplitude to detect micro-events (Section 5.2.1.2 in main text). A sample in A_{acc} is said to be lying in a non-event zone if the amplitude at that location is less than th_{amp} and the local variation is less than th_{std} . The local variation for each sample of A_{acc} is estimated by the standard deviation of A_{acc} computed over a moving window of size M_{seg} samples centred at that sample. The onset of p_i is defined to be either the preceding negative peak in A_{acc} or the first sample after the preceding non-event zone, whichever comes later. A boundary condition is applied so that the first sample in A_{acc} is assigned as the onset for the earliest p_i in case there is neither a negative peak nor a non-event zone preceding it. Similarly, the end of p_i is defined to be either the following negative peak in A_{acc} or the last sample before the following non-event zone, whichever comes earlier. Again, a boundary condition is applied so that the last sample in A_{acc} is assigned as the end of the last p_i in case there is neither a negative peak nor a non-event zone following it. Finally, candidate μE 's undergo a selection process: if the height of the corresponding p_i fails to rise by more than th_{amp} above the heights of A_{acc} at both p_i 's onset and end, we reject the candidate. This is done so that we only select μE 's that involve an abrupt increase in (pre-filtered) acceleration amplitude. At the end of this step, we obtain the locations and spans of all μE 's found in A_{acc} .

Appendix S5.2 | Non-exhaustive, category-wise list of 200 features developed in this study

Let s_D^{surge} , s_D^{sway} , s_D^{heave} , and s_D^{norm} denote the surge, sway, heave, and norm signals respectively within the domain D , where D can either be sp (span of the micro-event) or cx (context of the micro-event defined by a window of length w centred at the micro-event location), and s_D^{norm} is given by:

$$s_D^{norm} = \sqrt{(s_D^{surge})^2 + (s_D^{sway})^2 + (s_D^{heave})^2}$$

Similarly, let U_D^{surge} , U_D^{sway} , U_D^{heave} , and U_D^{norm} denote the upper envelopes, L_D^{surge} , L_D^{sway} , L_D^{heave} , and U_D^{norm} the lower envelopes, and A_D^{surge} , A_D^{sway} , A_D^{heave} , and A_D^{norm} the signal amplitudes of surge, sway, heave, and norm of acceleration within domain D .

For features where a moving mean was calculated, a window of size 0.2 s (or 20 samples @100 Hz) was slid forward by 1 sample. ‘De-meaned’ signals are used for some features: these were obtained by first computing the moving mean of the signal, and then subtracting it from the signal.

Table S5.1 provides definitions for some terms used to describe the features (Table S5.2) developed in this study.

Term	Definition
Power fraction	sum of squared Fourier coefficients in a given frequency range (say 2.5 to 5 Hz) divided by the sum of squared Fourier coefficients for the entire frequency range for a given signal, e.g. s_{cx}^{heave} .
Number of zero-crossings	number of times the values of a signal cross the value of zero.
Peak prominence	‘The prominence of a peak is the minimum vertical distance that the signal must descend on either side of the peak before either climbing back to a level higher than the peak or reaching an endpoint.’ (from MATLAB R2019a’s online documentation for the findpeaks function).
Peak width	‘The width of each peak is computed as the distance between the points to the left and right of the peak where the signal intercepts a reference line whose height is’ given by half the peak’s height (from MATLAB R2019a’s online documentation for the findpeaks function).

Table S5.1. Definitions for terms used in the feature table in Table S5.2.

Table S5.2. (next page) The list of features developed in this study along with computation details. Features are grouped into four meaningful categories (biomechanical, frequency, coordination, pattern). For each feature, the microevent neighbourhood (local or contextual) and signal (acceleration only, envelope only, or both acceleration and envelope) used for computation are indicated. Custom MATLAB code `getpk2pkHeight` (to compute features 36-51, 60-75, 123-134, and 139-150), `getFFTpeakpower` (to compute features 80-92, and 95-109), and `getFFTpowerfrac` (to compute features 110-112) are included in Supplementary Information.

S.No.	Category	Sub-category	Name	Microevent neighbourhood (local/contextual)	Signal (acceleration/envelope/both)	Description
1.	Biomechanical	Posture	spMeanSurge	local	acceleration	Mean of s_{sp}^{surge} .
2.			spMeanSway	local	acceleration	Mean of s_{sp}^{sway} .
3.			spMeanHeave	local	acceleration	Mean of s_{sp}^{heave} .
4.			cxMeanSurge	contextual	acceleration	Mean of s_{cx}^{surge} .
5.			cxMeanSway	contextual	acceleration	Mean of s_{cx}^{sway} .
6.			cxMeanHeave	contextual	acceleration	Mean of s_{cx}^{heave} .
7.			cxMaxMovMeanSurge	contextual	acceleration	Maximum of moving mean of s_{cx}^{surge} .
8.			cxMinMovMeanSurge	contextual	acceleration	Minimum of moving mean of s_{cx}^{surge} .
9.			cxRangeMovMeanSurge	contextual	acceleration	Range (maximum minus minimum) of moving mean of s_{cx}^{surge} .
10.		Intensity	cxMeanSurgeAMP	contextual	envelope	Mean of A_{cx}^{surge} .
11.			cxMeanSwayAMP	contextual	envelope	Mean of A_{cx}^{sway} .
12.			cxMeanHeaveAMP	contextual	envelope	Mean of A_{cx}^{heave} .
13.			cxMeanNormAMP	contextual	envelope	Mean of A_{cx}^{norm} .
14.			cxMinSurgeAMP	contextual	envelope	Minimum of A_{cx}^{surge} .
15.			cxMinSwayAMP	contextual	envelope	Minimum of A_{cx}^{sway} .
16.			cxMinHeaveAMP	contextual	envelope	Minimum of A_{cx}^{heave} .
17.			cxMinNormAMP	contextual	envelope	Minimum of A_{cx}^{norm} .
18.			cxMaxSurgeAMP	contextual	envelope	Maximum of A_{cx}^{surge} .
19.	cxMaxSwayAMP	contextual	envelope	Maximum of A_{cx}^{sway} .		

20.	cxMaxHeaveAMP	contextual	envelope	Maximum of A_{cx}^{heave} .
21.	cxMaxNormAMP	contextual	envelope	Maximum of A_{cx}^{norm} .
22.	cxRangeSurgeAMP	contextual	envelope	Range (maximum minus minimum) of A_{cx}^{surge} .
23.	cxRangeSwayAMP	contextual	envelope	Range (maximum minus minimum) of A_{cx}^{sway} .
24.	cxRangeHeaveAMP	contextual	envelope	Range (maximum minus minimum) of A_{cx}^{heave} .
25.	cxRangeNormAMP	contextual	envelope	Range (maximum minus minimum) of A_{cx}^{norm} .
26.	cxStdSurgeUE	contextual	envelope	Standard deviation of U_{cx}^{surge} .
27.	cxStdSwayUE	contextual	envelope	Standard deviation of U_{cx}^{sway} .
28.	cxStdHeaveUE	contextual	envelope	Standard deviation of U_{cx}^{heave} .
29.	cxStdSurgeLE	contextual	envelope	Standard deviation of L_{cx}^{surge} .
30.	cxStdSwayLE	contextual	envelope	Standard deviation of L_{cx}^{sway} .
31.	cxStdHeaveLE	contextual	envelope	Standard deviation of L_{cx}^{heave} .
32.	spRangeSurge	local	acceleration	Range (maximum minus minimum) of s_{sp}^{surge} .
33.	spRangeSway	local	acceleration	Range (maximum minus minimum) of s_{sp}^{sway} .
34.	spRangeHeave	local	acceleration	Range (maximum minus minimum) of s_{sp}^{heave} .
35.	spRangeNorm	local	acceleration	Range (maximum minus minimum) of s_{sp}^{norm} .
36.	spMaxPk2PkDescentSurge	local	acceleration	Maximum descent (height difference between a positive peak and the following negative peak) in s_{sp}^{surge} .
37.	spMaxPk2PkDescentSway	local	acceleration	Maximum descent (height difference between a positive peak and the following negative peak) in s_{sp}^{sway} .
38.	spMaxPk2PkDescentHeave	local	acceleration	Maximum descent (height difference between a positive peak and the following negative peak) in s_{sp}^{heave} .
39.	spMaxPk2PkDescentNorm	local	acceleration	Maximum descent (height difference between a positive peak and the following negative peak) in s_{sp}^{norm} .
40.	spMaxPk2PkAscentSurge	local	acceleration	Maximum ascent (height difference between a negative peak and the following positive peak) in s_{sp}^{surge} .

41.		spMaxPk2PkAscentSway	local	acceleration	Maximum ascent (height difference between a negative peak and the following positive peak) in s_{sp}^{sway} .
42.		spMaxPk2PkAscentHeave	local	acceleration	Maximum ascent (height difference between a negative peak and the following positive peak) in s_{sp}^{heave} .
43.		spMaxPk2PkAscentNorm	local	acceleration	Maximum ascent (height difference between a negative peak and the following positive peak) in s_{sp}^{norm} .
44.		spMaxPk2PkHtSurge	local	acceleration	Maximum ascent or descent (height difference between a negative peak and the following positive peak, or vice versa) in s_{sp}^{surge} .
45.		spMaxPk2PkHtSway	local	acceleration	Maximum ascent or descent (height difference between a negative peak and the following positive peak, or vice versa) in s_{sp}^{sway} .
46.		spMaxPk2PkHtHeave	local	acceleration	Maximum ascent or descent (height difference between a negative peak and the following positive peak, or vice versa) in s_{sp}^{heave} .
47.		spMaxPk2PkHtNorm	local	acceleration	Maximum ascent or descent (height difference between a negative peak and the following positive peak, or vice versa) in s_{sp}^{norm} .
48.		spMeanPk2PkHtSurge	local	acceleration	Mean of ascents and descents (height difference between a negative peak and the following positive peak, and vice versa) in s_{sp}^{surge} .
49.		spMeanPk2PkHtSway	local	acceleration	Mean of ascents and descents (height difference between a negative peak and the following positive peak, and vice versa) in s_{sp}^{sway} .
50.		spMeanPk2PkHtHeave	local	acceleration	Mean of ascents and descents (height difference between a negative peak and the following positive peak, and vice versa) in s_{sp}^{heave} .
51.		spMeanPk2PkHtNorm	local	acceleration	Mean of ascents and descents (height difference between a negative peak and the following positive peak, and vice versa) in s_{sp}^{norm} .
52.		spMeanAbsJerkSurge	local	acceleration	Mean of absolute value of time-derivative of s_{sp}^{surge} .
53.		spMeanAbsJerkSway	local	acceleration	Mean of absolute value of time-derivative of s_{sp}^{sway} .
54.		spMeanAbsJerkHeave	local	acceleration	Mean of absolute value of time-derivative of s_{sp}^{heave} .
55.		spMeanAbsJerkNorm	local	acceleration	Mean of absolute value of time-derivative of s_{sp}^{norm} .
56.		cxStdSurge	contextual	acceleration	Standard deviation of s_{cx}^{surge} .
57.		cxStdSway	contextual	acceleration	Standard deviation of s_{cx}^{sway} .
58.		cxStdHeave	contextual	acceleration	Standard deviation of s_{cx}^{heave} .
59.		cxStdNorm	contextual	acceleration	Standard deviation of s_{cx}^{norm} .
60.		cxMaxPk2PkDescentSurge	contextual	acceleration	Maximum descent (height difference between a positive peak and the following negative peak) in s_{cx}^{surge} .

61.			cxMaxPk2PkDescentSway	contextual	acceleration	Maximum descent (height difference between a positive peak and the following negative peak) in s_{cx}^{sway} .
62.			cxMaxPk2PkDescentHeave	contextual	acceleration	Maximum descent (height difference between a positive peak and the following negative peak) in s_{cx}^{heave} .
63.			cxMaxPk2PkDescentNorm	contextual	acceleration	Maximum descent (height difference between a positive peak and the following negative peak) in s_{cx}^{norm} .
64.			cxMaxPk2PkAscentSurge	contextual	acceleration	Maximum ascent (height difference between a negative peak and the following positive peak) in s_{cx}^{surge} .
65.			cxMaxPk2PkAscentSway	contextual	acceleration	Maximum ascent (height difference between a negative peak and the following positive peak) in s_{cx}^{sway} .
66.			cxMaxPk2PkAscentHeave	contextual	acceleration	Maximum ascent (height difference between a negative peak and the following positive peak) in s_{cx}^{heave} .
67.			cxMaxPk2PkAscentNorm	contextual	acceleration	Maximum ascent (height difference between a negative peak and the following positive peak) in s_{cx}^{norm} .
68.			cxMaxPk2PkHtSurge	contextual	acceleration	Maximum ascent or descent (height difference between a negative peak and the following positive peak, or vice versa) in s_{cx}^{surge} .
69.			cxMaxPk2PkHtSway	contextual	acceleration	Maximum ascent or descent (height difference between a negative peak and the following positive peak, or vice versa) in s_{cx}^{sway} .
70.			cxMaxPk2PkHtHeave	contextual	acceleration	Maximum ascent or descent (height difference between a negative peak and the following positive peak, or vice versa) in s_{cx}^{heave} .
71.			cxMaxPk2PkHtNorm	contextual	acceleration	Maximum ascent or descent (height difference between a negative peak and the following positive peak, or vice versa) in s_{cx}^{norm} .
72.			cxMeanPk2PkHtSurge	contextual	acceleration	Mean of ascents and descents (height difference between a negative peak and the following positive peak, and vice versa) in s_{cx}^{surge} .
73.			cxMeanPk2PkHtSway	contextual	acceleration	Mean of ascents and descents (height difference between a negative peak and the following positive peak, and vice versa) in s_{cx}^{sway} .
74.			cxMeanPk2PkHtHeave	contextual	acceleration	Mean of ascents and descents (height difference between a negative peak and the following positive peak, and vice versa) in s_{cx}^{heave} .
75.			cxMeanPk2PkHtNorm	contextual	acceleration	Mean of ascents and descents (height difference between a negative peak and the following positive peak, and vice versa) in s_{cx}^{norm} .
76.			cxMeanAbsJerkSurge	contextual	acceleration	Mean of absolute value of time-derivative of s_{cx}^{surge} .
77.			cxMeanAbsJerkSway	contextual	acceleration	Mean of absolute value of time-derivative of s_{cx}^{sway} .
78.			cxMeanAbsJerkHeave	contextual	acceleration	Mean of absolute value of time-derivative of s_{cx}^{heave} .
79.			cxMeanAbsJerkNorm	contextual	acceleration	Mean of absolute value of time-derivative of s_{cx}^{norm} .
80.	Frequency	Periodicity	cxFFTpeakpowerSurgeUE	contextual	envelope	Maximum squared coefficient of Fourier transform of U_{cx}^{surge} .

81.		cxFFTpeakpowerSwayUE	contextual	envelope	Maximum squared coefficient of Fourier transform of U_{cx}^{sway} .
82.		cxFFTpeakpowerHeaveUE	contextual	envelope	Maximum squared coefficient of Fourier transform of U_{cx}^{heave} .
83.		cxFFTpeakpowerUEavg	contextual	envelope	Maximum of axis-averaged (term-wise sum across axes divided by number of axes) squared coefficients of Fourier transforms of U_{cx}^{surge} , U_{cx}^{sway} and U_{cx}^{heave} .
84.		cxFFTpeakpowerSurgeLE	contextual	envelope	Maximum squared coefficient of Fourier transform of L_{cx}^{surge} .
85.		cxFFTpeakpowerSwayLE	contextual	envelope	Maximum squared coefficient of Fourier transform of L_{cx}^{sway} .
86.		cxFFTpeakpowerHeaveLE	contextual	envelope	Maximum squared coefficient of Fourier transform of L_{cx}^{heave} .
87.		cxFFTpeakpowerLEavg	contextual	envelope	Maximum of axis-averaged (term-wise sum across axes divided by number of axes) squared coefficients of Fourier transforms of L_{cx}^{surge} , L_{cx}^{sway} and L_{cx}^{heave} .
88.		cxFFTpeakpowerSurge	contextual	acceleration	Maximum squared coefficient of Fourier transform of s_{cx}^{surge} .
89.		cxFFTpeakpowerSway	contextual	acceleration	Maximum squared coefficient of Fourier transform of s_{cx}^{sway} .
90.		cxFFTpeakpowerHeave	contextual	acceleration	Maximum squared coefficient of Fourier transform of s_{cx}^{heave} .
91.		cxFFTpeakpowerNorm	contextual	acceleration	Maximum squared coefficient of Fourier transform of s_{cx}^{norm} .
92.		cxFFTpeakpowerAvg	contextual	acceleration	Maximum of axis-averaged (term-wise sum across axes divided by number of axes) squared coefficients of Fourier transforms of s_{cx}^{surge} , s_{cx}^{sway} and s_{cx}^{heave} .
93.	Dominant frequency	cxNumEvents	contextual	envelope	Total number of micro-event peaks found in A_{cx}^{surge} .
94.		cxEventFreq	contextual	envelope	Reciprocal of mean distance (in seconds) between successive micro-event peaks in A_{cx}^{surge} .
95.		cxFFTdomfreqSurgeUE	contextual	envelope	Frequency at which maximum squared coefficient of Fourier transform of U_{cx}^{surge} is obtained.
96.		cxFFTdomfreqSwayUE	contextual	envelope	Frequency at which maximum squared coefficient of Fourier transform of U_{cx}^{sway} is obtained.
97.		cxFFTdomfreqHeaveUE	contextual	envelope	Frequency at which maximum squared coefficient of Fourier transform of U_{cx}^{heave} is obtained.
98.		cxFFTdomfreqNormUE	contextual	envelope	Frequency at which maximum squared coefficient of Fourier transform of U_{cx}^{norm} is obtained.
99.		cxFFTdomfreqAvgUE	contextual	envelope	Frequency at which maximum of axis-averaged squared coefficients of Fourier transforms of U_{cx}^{surge} , U_{cx}^{sway} and U_{cx}^{heave} is obtained.

100.		cxFFTdomfreqSurgeLE	contextual	envelope	Frequency at which maximum squared coefficient of Fourier transform of L_{cx}^{surge} is obtained.
101.		cxFFTdomfreqSwayLE	contextual	envelope	Frequency at which maximum squared coefficient of Fourier transform of L_{cx}^{sway} is obtained.
102.		cxFFTdomfreqHeaveLE	contextual	envelope	Frequency at which maximum squared coefficient of Fourier transform of L_{cx}^{heave} is obtained.
103.		cxFFTdomfreqNormLE	contextual	envelope	Frequency at which maximum squared coefficient of Fourier transform of L_{cx}^{norm} is obtained.
104.		cxFFTdomfreqAvgLE	contextual	envelope	Frequency at which maximum of axis-averaged squared coefficients of Fourier transforms of L_{cx}^{surge} , L_{cx}^{sway} and L_{cx}^{heave} is obtained.
105.		cxFFTdomfreqSurge	contextual	acceleration	Frequency at which maximum squared coefficient of Fourier transform of s_{cx}^{surge} is obtained.
106.		cxFFTdomfreqSway	contextual	acceleration	Frequency at which maximum squared coefficient of Fourier transform of s_{cx}^{sway} is obtained.
107.		cxFFTdomfreqHeave	contextual	acceleration	Frequency at which maximum squared coefficient of Fourier transform of s_{cx}^{heave} is obtained.
108.		cxFFTdomfreqNorm	contextual	acceleration	Frequency at which maximum squared coefficient of Fourier transform of s_{cx}^{norm} is obtained.
109.		cxFFTdomfreqAvg	contextual	acceleration	Frequency at which maximum of axis-averaged squared coefficients of Fourier transforms of s_{cx}^{surge} , s_{cx}^{sway} and s_{cx}^{heave} is obtained.
110.	Frequency distribution	cxFFTpowerfracHeaveUE	contextual	envelope	Power fraction in range 2.5 to 5 Hz for U_{cx}^{heave} .
111.		cxFFTpowerfracHeave	contextual	acceleration	Power fraction in range 2.5 to 5 Hz for s_{cx}^{heave} .
112.		othFFTpowerfracHeaveUEfacc	contextual	both	Power fraction in range 2.5 to 5 Hz for U_{cx}^{heave} divided by power fraction in range 2.5 to 5 Hz for s_{cx}^{heave} .
113.	Coordination	spCorrSurgeHeaveUE	local	envelope	Coefficient of correlation between U_{sp}^{surge} and U_{sp}^{heave} .
114.		cxCorrSurgeHeaveUE	contextual	envelope	Coefficient of correlation between U_{cx}^{surge} and U_{cx}^{heave} .
115.		cxCorrSurgeHeaveLE	contextual	envelope	Coefficient of correlation between L_{cx}^{surge} and L_{cx}^{heave} .
116.		cxCorrSurgeHeaveAMP	contextual	envelope	Coefficient of correlation between A_{cx}^{surge} and A_{cx}^{heave} .
117.		spCorrSurgeHeave	local	acceleration	Coefficient of correlation between s_{sp}^{surge} and s_{sp}^{heave} .
118.		cxCorrSurgeHeave	contextual	acceleration	Coefficient of correlation between s_{cx}^{surge} and s_{cx}^{heave} .

119.	Pattern	Tortuosity	spNumZeroXdmSurge	local	acceleration	Number of zero-crossings in de-meanded s_{sp}^{surge} .
120.			spNumZeroXdmSway	local	acceleration	Number of zero-crossings in de-meanded s_{sp}^{sway} .
121.			spNumZeroXdmHeave	local	acceleration	Number of zero-crossings in de-meanded s_{sp}^{heave} .
122.			spNumZeroXdmNorm	local	acceleration	Number of zero-crossings in de-meanded s_{sp}^{norm} .
123.			spNumAscentSurge	local	acceleration	Number of ascents (signal path from a negative peak to the following positive peak) in s_{sp}^{surge} .
124.			spNumAscentSway	local	acceleration	Number of ascents (signal path from a negative peak to the following positive peak) in s_{sp}^{sway} .
125.			spNumAscentHeave	local	acceleration	Number of ascents (signal path from a negative peak to the following positive peak) in s_{sp}^{heave} .
126.			spNumAscentNorm	local	acceleration	Number of ascents (signal path from a negative peak to the following positive peak) in s_{sp}^{norm} .
127.			spNumDescentSurge	local	acceleration	Number of descents (signal path from a positive peak to the following negative peak) in s_{sp}^{surge} .
128.			spNumDescentSway	local	acceleration	Number of descents (signal path from a positive peak to the following negative peak) in s_{sp}^{sway} .
129.			spNumDescentHeave	local	acceleration	Number of descents (signal path from a positive peak to the following negative peak) in s_{sp}^{heave} .
130.			spNumDescentNorm	local	acceleration	Number of descents (signal path from a positive peak to the following negative peak) in s_{sp}^{norm} .
131.			spNumAscentDescentSurge	local	acceleration	Number of ascents and descents (signal paths from a negative peak to the following positive peak, and vice versa) in s_{sp}^{surge} .
132.			spNumAscentDescentSway	local	acceleration	Number of ascents and descents (signal paths from a negative peak to the following positive peak, and vice versa) in s_{sp}^{sway} .
133.			spNumAscentDescentHeave	local	acceleration	Number of ascents and descents (signal paths from a negative peak to the following positive peak, and vice versa) in s_{sp}^{heave} .
134.			spNumAscentDescentNorm	local	acceleration	Number of ascents and descents (signal paths from a negative peak to the following positive peak, and vice versa) in s_{sp}^{norm} .
135.			cxNumZeroXdmSurge	contextual	acceleration	Number of zero-crossings in de-meanded s_{cx}^{surge} .
136.			cxNumZeroXdmSway	contextual	acceleration	Number of zero-crossings in de-meanded s_{cx}^{sway} .
137.			cxNumZeroXdmHeave	contextual	acceleration	Number of zero-crossings in de-meanded s_{cx}^{heave} .
138.			cxNumZeroXdmNorm	contextual	acceleration	Number of zero-crossings in de-meanded s_{cx}^{norm} .

139.		cxNumAscentSurge	contextual	acceleration	Number of ascents (signal path from a negative peak to the following positive peak) in s_{cx}^{surge} .
140.		cxNumAscentSway	contextual	acceleration	Number of ascents (signal path from a negative peak to the following positive peak) in s_{cx}^{sway} .
141.		cxNumAscentHeave	contextual	acceleration	Number of ascents (signal path from a negative peak to the following positive peak) in s_{cx}^{heave} .
142.		cxNumAscentNorm	contextual	acceleration	Number of ascents (signal path from a negative peak to the following positive peak) in s_{cx}^{norm} .
143.		cxNumDescentSurge	contextual	acceleration	Number of descents (signal path from a positive peak to the following negative peak) in s_{cx}^{surge} .
144.		cxNumDescentSway	contextual	acceleration	Number of descents (signal path from a positive peak to the following negative peak) in s_{cx}^{sway} .
145.		cxNumDescentHeave	contextual	acceleration	Number of descents (signal path from a positive peak to the following negative peak) in s_{cx}^{heave} .
146.		cxNumDescentNorm	contextual	acceleration	Number of descents (signal path from a positive peak to the following negative peak) in s_{cx}^{norm} .
147.		cxNumAscentDescentSurge	contextual	acceleration	Number of ascents and descents (signal paths from a negative peak to the following positive peak, and vice versa) in s_{cx}^{surge} .
148.		cxNumAscentDescentSway	contextual	acceleration	Number of ascents and descents (signal paths from a negative peak to the following positive peak, and vice versa) in s_{cx}^{sway} .
149.		cxNumAscentDescentHeave	contextual	acceleration	Number of ascents and descents (signal paths from a negative peak to the following positive peak, and vice versa) in s_{cx}^{heave} .
150.		cxNumAscentDescentNorm	contextual	acceleration	Number of ascents and descents (signal paths from a negative peak to the following positive peak, and vice versa) in s_{cx}^{norm} .
151.		othTortuositySurgeWrtUE	contextual	both	Mean of absolute values of term-wise differences between s_{cx}^{surge} and U_{cx}^{surge} .
152.		othTortuositySwayWrtUE	contextual	both	Mean of absolute values of term-wise differences between s_{cx}^{sway} and U_{cx}^{sway} .
153.		othTortuosityHeaveWrtUE	contextual	both	Mean of absolute values of term-wise differences between s_{cx}^{heave} and U_{cx}^{heave} .
154.		othTortuositySurgeWrtLE	contextual	both	Mean of absolute values of term-wise differences between s_{cx}^{surge} and L_{cx}^{surge} .
155.		othTortuositySwayWrtLE	contextual	both	Mean of absolute values of term-wise differences between s_{cx}^{sway} and L_{cx}^{sway} .
156.		othTortuosityHeaveWrtLE	contextual	both	Mean of absolute values of term-wise differences between s_{cx}^{heave} and L_{cx}^{heave} .
157.	Microevents in	spEventPkHeightAMP	local	envelope	Height of main micro-event's peak in A_{sp}^{surge} .
158.		spEventPkSpanAMP	local	envelope	Span of main micro-event's peak in A_{sp}^{surge} .

159.	amplitude signal	spEventPkHtBySpanAMP	local	envelope	Height of main micro-event's peak in A_{sp}^{surge} divided by its span.
160.		spEventPkWidthAMP	local	envelope	Width of main micro-event's peak in A_{sp}^{surge} .
161.		spEventPkPromAMP	local	envelope	Prominence of main micro-event's peak in A_{sp}^{surge} .
162.		spEventPkWbySpanAMP	local	envelope	Width of main micro-event's peak in A_{sp}^{surge} divided by its span.
163.		spEventPkPromByHtAmp	local	envelope	Prominence of main micro-event's peak in A_{sp}^{surge} divided by its height.
164.		spEventPkLocSymmAMP	local	envelope	Smaller of: number of samples between right span's edge from main micro-event's peak in A_{sp}^{surge} divided by that of left span's edge, and vice versa.
165.		spEventPkHtSymmAMP	local	envelope	Absolute value of difference in height of A_{sp}^{surge} at right span's and left span's edge.
166.		cxMeanPkHeightAMP	contextual	envelope	Mean height of all micro-event peaks found in A_{cx}^{surge} .
167.		cxMeanPkWidthAMP	contextual	envelope	Mean width of all micro-event peaks found in A_{cx}^{surge} .
168.		cxMeanPkPromAMP	contextual	envelope	Mean prominence of all micro-event peaks found in A_{cx}^{surge} .
169.	cxMeanPkPromByHtAmp	contextual	envelope	Mean of prominence divided by height of all micro-event peaks found in A_{cx}^{surge} .	
170.	cxFracNonEvent	contextual	envelope	Fraction of samples belonging to non-event zones in A_{cx}^{surge} .	
171.	Statistical	cxMinSurgeUE	contextual	envelope	Minimum of U_{cx}^{surge} .
172.		cxMinSwayUE	contextual	envelope	Minimum of U_{cx}^{sway} .
173.		cxMinHeaveUE	contextual	envelope	Minimum of U_{cx}^{heave} .
174.		cxMinSurgeLE	contextual	envelope	Minimum of L_{cx}^{surge} .
175.		cxMinSwayLE	contextual	envelope	Minimum of L_{cx}^{sway} .
176.		cxMinHeaveLE	contextual	envelope	Minimum of L_{cx}^{heave} .
177.		cxMaxSurgeUE	contextual	envelope	Maximum of U_{cx}^{surge} .
178.		cxMaxSwayUE	contextual	envelope	Maximum of U_{cx}^{sway} .

179.	cxMaxHeaveUE	contextual	envelope	Maximum of U_{cx}^{heave} .
180.	cxMaxSurgeLE	contextual	envelope	Maximum of L_{cx}^{surge} .
181.	cxMaxSwayLE	contextual	envelope	Maximum of L_{cx}^{sway} .
182.	cxMaxHeaveLE	contextual	envelope	Maximum of L_{cx}^{heave} .
183.	cxMeanSurgeUE	contextual	envelope	Mean of U_{cx}^{surge} .
184.	cxMeanSwayUE	contextual	envelope	Mean of U_{cx}^{sway} .
185.	cxMeanHeaveUE	contextual	envelope	Mean of U_{cx}^{heave} .
186.	cxMeanSurgeLE	contextual	envelope	Mean of L_{cx}^{surge} .
187.	cxMeanSwayLE	contextual	envelope	Mean of L_{cx}^{sway} .
188.	cxMeanHeaveLE	contextual	envelope	Mean of L_{cx}^{heave} .
189.	cxStdSurgeAMP	contextual	envelope	Standard deviation of A_{cx}^{surge} .
190.	cxStdSwayAMP	contextual	envelope	Standard deviation of A_{cx}^{sway} .
191.	cxStdHeaveAMP	contextual	envelope	Standard deviation of A_{cx}^{heave} .
192.	cxStdNormAMP	contextual	envelope	Standard deviation of A_{cx}^{norm} .
193.	spMaxSurge	local	acceleration	Maximum of s_{sp}^{surge} .
194.	spMaxSway	local	acceleration	Maximum of s_{sp}^{sway} .
195.	spMaxHeave	local	acceleration	Maximum of s_{sp}^{heave} .
196.	spMaxNorm	local	acceleration	Maximum of s_{sp}^{norm} .
197.	spMinSurge	local	acceleration	Minimum of s_{sp}^{surge} .
198.	spMinSway	local	acceleration	Minimum of s_{sp}^{sway} .
199.	spMinHeave	local	acceleration	Minimum of s_{sp}^{heave} .
200.	spMinNorm	local	acceleration	Minimum of s_{sp}^{norm} .

Appendix S5.3 | Feature selection for micro-event detection

At each node of the classification tree in Figure 5.2 in the main text, the best feature from each category was chosen by first ranking features in that category using five different filter-method feature selection methods in MATLAB R2019a, and then selecting the top-ranked feature. The five filter-method feature selection techniques used were: `ttest`, `entropy`, `bhattacharyya`, `roc`, and `wilcoxon`, which were chosen from the `rankfeatures` function in MATLAB R2019a. Ranks obtained using the five methods were combined by computing a ‘mock penalty score’ given by the sum of a feature’s ranks across the five methods. Thus, the lower the mock penalty score, the higher the feature’s overall rank. The feature with the lowest mock penalty score was chosen as the overall top-ranked feature for a given category and node.

The best results for micro-event detection were obtained with a set of eight features (Table S5.3) used with logistic regression at each node of the hierarchical classification scheme in Figure 5.2 of the main text.

S.No.	Type	Name
1.	<i>Posture</i>	<i>cxMinMovMeanSurge</i>
2.	<i>Intensity</i>	<i>cxStdSway</i>
3.	<i>Periodicity</i>	<i>cxFFTpeakpowerHeave</i>
4.	<i>Dom. freq.</i>	<i>cxFFTdomfreqNorm</i>
5.	<i>Freq. content</i>	<i>cxFFTpowerfracHeave</i>
6.	<i>Coordination</i>	<i>cxCorrSurgeHeave</i>
7.	<i>Tortuosity</i>	<i>cxNumAscDescHeave</i>
8.	<i>Statistical</i>	<i>spMinSurge</i>

Table S5.3. The set of eight features that gave the best results for micro-event detection when used with logistic regression at each of the three nodes in the hierarchical classification scheme in Figure 5.2 in the main text. These features were derived from both local (i.e. from micro-event span) and contextual (i.e. from window centred at detected micro-event) acceleration-only signals (*cf.* Section 2.2.3 under ‘*Feature development and selection*’).

Appendix S5.4 | A data resampling approach to address inter-individual data imbalance

Motivation

Most behaviour recognition studies aim to build models that can capture inter-individual variation in the behaviours of interest well enough to perform accurately on new, unseen individuals. This requires having sufficient instances of the same behaviour for several individuals in the groundtruthing dataset. This may be easier to ensure for broad behavioural categories, since it can in general be possible to plan data collection so that enough data from the behaviours of interest are captured on film. When one is interested in studying specific behaviours in wild animals, however, the unpredictability of free-living conditions means that data collection is inevitably opportunistic rather than systematic. This is bound to cause imbalance in the amounts of data collected for different individuals for the same behaviour. For instance, collecting groundtruth data on chewing from a predator will be contingent on the specific individual's foraging skill, the individual's luck in finding food in its current environment (Wilson et al. 2018), and whether prey capture occurred while filming the animal. Further, a given individual might capture a large prey item requiring several head movements to ingest, while other individuals may capture several smaller prey items requiring fewer head movements to ingest. The large amount of chewing data for one individual for a single prey item would bias the entire model to one specific prey type and individual whereas model robustness would require correct predictions across the maximum available range of prey types and individuals. Such dataset imbalances are problematic because classification algorithms typically seek to minimise overall error, and so will bias models towards data from individuals that have a majority share in the dataset for a given behaviour. The problem of inter-class data imbalance has received considerable attention (He & Garcia 2009), and several techniques have been developed to construct robust classifiers from imbalanced datasets (Branco, Torgo & Ribeiro 2016). Yet, efforts to study free-living animals in their natural environment, contrary to in-lab human studies where it is possible to control the type and duration of behaviours of interest, has meant that the problem of inter-individual data imbalance is new and, to our knowledge, has not been addressed before.

Proposed data resampling approach

One of the most popular approaches to address the problem of imbalanced representation of a two-class dataset involves balancing the dataset by undersampling the majority (more common) class and oversampling the minority (rarer) class to achieve the desired level of class parity in the dataset (Chawla et al. 2002).

Undersampling is achieved by choosing a random subset of the majority class data (henceforth referred to as Random Undersampling, RUS), and oversampling by the authors' proposed Synthetic Minority Oversampling Technique (SMOTE). In SMOTE, depending on the amount of oversampling required, synthetic data points for the minority class are generated at randomly chosen distances along the line segments joining each minority class sample to any/all of its 5 nearest neighbours in feature space. SMOTE effectively forces the decision region of the minority class to become more general as compared to that obtained after oversampling the minority class with replacement (Chawla et al. 2002).

Here, I extend this approach to tackle the new problem of balancing a dataset in the presence of inter-individual imbalance – i.e. when behavioural classes are not approximately equally represented across different individuals. Consider a multi-class, imbalanced dataset $\{\delta^{imb}\}$ containing instances from N_c classes ($N_c \geq 2$) collected for a total of N_i individuals. Let $\delta_{i,c}^{imb}$ represent the number of data points of class c collected for individual i . Next, let $M_{\{\delta^{imb}\}}$ be defined as follows:

$$M_{\{\delta^{imb}\}} = \text{median}(\{\{\delta_{i,c}^{imb}, \forall i \in 1 \dots N_i \text{ and } c \in 1 \dots N_c\}\})$$

Our data-balancing approach involves creating a new, balanced dataset $\{\delta^{bal}\}$ by resampling the original, imbalanced dataset $\{\delta^{imb}\}$ according to the following pseudo-code:

For each class c and individual i :

a. if $\delta_{i,c}^{imb} < M_{\{\delta^{imb}\}}$

oversample $\delta_{i,c}^{imb}$ using SMOTE such that the new number of data points for this individual and class is $\delta_{i,c}^{bal} = M_{\{\delta^{imb}\}}$.

b. else if $\delta_{i,c}^{imb} > M_{\{\delta^{imb}\}}$

randomly select a smaller number of samples $\delta_{i,c}^{bal} = M_{\{\delta^{imb}\}}$ from $\delta_{i,c}^{imb}$.

This ensures that not only will the representation of each class be equal in the resampled dataset (a total of $\sum_{i=1}^{N_i} \delta_{i,c}^{bal} = N_i \times M_{\{\delta^{imb}\}}$ data points for each class) but that our goal of balancing data from each individual for a given class will have been achieved, since the number of data points for a given class c from each individual will be equal to $\delta_{i,c}^{bal} = M_{\{\delta^{imb}\}} \forall i \in 1 \dots N_i$. Finally, note that the SMOTE algorithm as proposed by Chawla et al. 2002 only allows oversampling by $\lambda \times 100\%$, where λ is a positive integer. Since our application of SMOTE requires oversampling to a specific number $M_{\{\delta^{imb}\}}$, which may not be an integral multiple of the original number of underrepresented data points, I modify the SMOTE algorithm (custom MATLAB code `SMOTE` in Supplementary Information) so that it cycles through the original data points one by one and generates one new synthetic sample in each iteration and stops when the desired number of samples have been created. This offers precise control over the size of the oversampled dataset.

Appendix S5.5 | Metrics for performance evaluation of classification models

As done in Chakravarty et al. 2019, I used a confusion matrix $[N_{jk}]$ to summarise model performance, where $j, k \in \{1: 1AD, 2: 2AD, 3: search, 4: chew\}$ and N_{jk} denotes the number of samples of behaviour ‘ j ’ that were predicted by the model to be behaviour ‘ k ’. From this confusion matrix, I computed three behaviour-specific metrics – sensitivity (also referred to as recall in literature), precision, and specificity, as well as overall model accuracy. These measures are computed as shown in Table S3.1.

Appendix S5.6 | Feature selection for classical machine learning approaches

Two feature families were considered for computation of performance benchmarks with classical machine learning: (1) the set of 38 summary statistics

presented in Nathan et al. 2012 (Table S5.4), and (2) the set of 131 of 200 features (Table S5.2) that do not require micro-event detection-related information.

S.No.	Feature name	Feature description
1.	<i>meanSurge</i>	Mean of surge acceleration.
2.	<i>meanSway</i>	Mean of sway acceleration.
3.	<i>meanHeave</i>	Mean of heave acceleration.
4.	<i>meanNorm</i>	Mean of acceleration norm.
5.	<i>stdSurge</i>	Standard deviation of surge acceleration.
6.	<i>stdSway</i>	Standard deviation of sway acceleration.
7.	<i>stdHeave</i>	Standard deviation of heave acceleration.
8.	<i>stdNorm</i>	Standard deviation of acceleration norm.
9.	<i>skewSurge</i>	Skewness of surge acceleration.
10.	<i>skewSway</i>	Skewness of sway acceleration.
11.	<i>skewHeave</i>	Skewness of heave acceleration.
12.	<i>skewNorm</i>	Skewness of acceleration norm.
13.	<i>kurtSurge</i>	Kurtosis of surge acceleration.
14.	<i>kurtSway</i>	Kurtosis of sway acceleration.
15.	<i>kurtHeave</i>	Kurtosis of heave acceleration.
16.	<i>kurtNorm</i>	Kurtosis of acceleration norm.
17.	<i>maxSurge</i>	Maximum of surge acceleration.
18.	<i>maxSway</i>	Maximum of sway acceleration.
19.	<i>maxHeave</i>	Maximum of heave acceleration.
20.	<i>maxNorm</i>	Maximum of acceleration norm.
21.	<i>minSurge</i>	Minimum of surge acceleration.

22.	<i>minSway</i>	Minimum of sway acceleration.
23.	<i>minHeave</i>	Minimum of heave acceleration.
24.	<i>minNorm</i>	Minimum of acceleration norm.
25.	<i>acorrSurge</i>	Autocorrelation (for a displacement of one measurement) of surge acceleration.
26.	<i>acorrSway</i>	Autocorrelation (for a displacement of one measurement) of sway acceleration.
27.	<i>acorrHeave</i>	Autocorrelation (for a displacement of one measurement) of heave acceleration.
28.	<i>acorrNorm</i>	Autocorrelation (for a displacement of one measurement) of acceleration norm.
29.	<i>trendSurge</i>	The coefficient for a linear regression through the data for surge acceleration.
30.	<i>trendSway</i>	The coefficient for a linear regression through the data for sway acceleration.
31.	<i>trendHeave</i>	The coefficient for a linear regression through the data for heave acceleration.
32.	<i>trendNorm</i>	The coefficient for a linear regression through the data for acceleration norm.
33.	<i>corrSurgeSway</i>	Coefficient of correlation between surge and sway acceleration.
34.	<i>corrSwayHeave</i>	Coefficient of correlation between sway and heave acceleration.
35.	<i>corrHeaveSurge</i>	Coefficient of correlation between heave and surge acceleration.
36.	<i>odba</i>	Overall dynamic body acceleration (ODBA) as introduced in Wilson et al. 2006.
37.	<i>circVar_incli</i>	Circular variance of inclination ($\cos^{-1}[\textit{heave}/\textit{norm}]$).
38.	<i>circVar_azi</i>	Circular variance of azimuth ($\tan^{-1}[\textit{surge}/\textit{sway}]$).

Table S5.4. The family of 38 features presented in Nathan et al., 2012.

Four different feature selection methods in Weka (version 3.8.0) were run to select the top 8 features in each family of features. These methods were:

1. CfsSubset (filter method)
2. CorrelationAttributeEval (filter method)
3. InfoGainAttributeEval (filter method)
4. WrapperSubsetEval (wrapper method with Decision Tree J48)

S6.1 | Window size = 1.5 s, overlap = 1 s

S6.1.1 | Summary statistics in Nathan et al. 2012

CfsSubset (best first)	Filter Methods		Wrapper Method
	Correlation (ranked)	InfoGain (ranked)	WrapperSubsetEval (J48, best first)
<i>meanSurge</i>	<i>meanSurge</i>	<i>stdSway</i>	<i>meanSurge</i>
<i>meanSway</i>	<i>odba</i>	<i>odba</i>	<i>meanSway</i>
<i>meanNorm</i>	<i>meanNorm</i>	<i>meanSurge</i>	<i>meanHeave</i>
<i>stdSway</i>	<i>acorrHeave</i>	<i>meanNorm</i>	<i>stdSway</i>
<i>skewSurge</i>	<i>stdSway</i>	<i>kurtHeave</i>	<i>stdHeave</i>
<i>kurtHeave</i>	<i>corrHeaveSurge</i>	<i>acorrHeave</i>	<i>stdNorm</i>
<i>kurtNorm</i>	<i>acorrNorm</i>	<i>minSurge</i>	<i>maxSway</i>
<i>minSurge</i>	<i>minSurge</i>	<i>acorrSurge</i>	<i>maxHeave</i>
<i>acorrSurge</i>	<i>stdSurge</i>	<i>acorrNorm</i>	<i>maxNorm</i>
<i>acorrSway</i>	<i>kurtHeave</i>	<i>kurtNorm</i>	<i>minSway</i>
<i>acorrHeave</i>	<i>stdNorm</i>	<i>minSway</i>	<i>minNorm</i>
<i>acorrNorm</i>	<i>stdHeave</i>	<i>stdSurge</i>	<i>acorrHeave</i>
<i>trendSurge</i>	<i>minSway</i>	<i>corrHeaveSurge</i>	<i>acorrNorm</i>
<i>corrHeaveSurge</i>	<i>acorrSurge</i>	<i>stdNorm</i>	<i>trendHeave</i>
<i>odba</i>	<i>maxSway</i>	<i>acorrSway</i>	<i>corrHeaveSurge</i>

Table S5.5. Feature selection (window size 1.5 s) from the 38 summary statistics presented by Nathan et al. 2012 (Table S5.4). Four different feature selection techniques were used to select the top eight features (green). The features in the top fifteen in each method but not in the top eight are indicated in grey.

The following 8 features occurred the most often (Table S5.5), and were the ones selected for comparison with our model:

- *meanSurge* (all 4 methods)
- *stdSway* (all 4 methods)
- *acorrHeave* (all 4 methods)
- *acorrNorm* (all 4 methods)
- *corrHeaveSurge* (all 4 methods)
- *stdNorm* (2 filter methods, 1 wrapper method)
- *minSway* (2 filter methods, 1 wrapper method)
- *meanNorm* (3 filter methods; sum of ranks across 3 filter methods lowest compared to other features that appear in all three filter methods, e.g. *kurtHeave*, *minSurge*, *acorrSurge*, etc.).

S6.1.2 | Feature family from this study

CfsSubset (best first)	Filter Methods		Wrapper Method
	Correlation (ranked)	InfoGain (ranked)	WrapperSubsetEval (J48, best first)
<i>cxMeanSwayAMP</i>	<i>cxMeanSurgeLE</i>	<i>cxMeanSurgeLE</i>	<i>cxMinNormAMP</i>
<i>cxMinSurgeAMP</i>	<i>cxMaxSurgeLE</i>	<i>cxFFTdomfreqAvg</i>	<i>cxFFTpeakpowerHeaveUE</i>
<i>cxMinSwayAMP</i>	<i>cxMeanSurge</i>	<i>cxMaxSurgeLE</i>	<i>cxCorrSurgeHeaveLE</i>
<i>cxMinHeaveAMP</i>	<i>cxMinMovMeanSurge</i>	<i>cxFFTdomfreqHeave</i>	<i>cxMinSwayUE</i>
<i>cxMinNormAMP</i>	<i>cxMeanAbsJerkSway</i>	<i>othTortuositySwayWrtLE</i>	<i>cxMaxHeaveLE</i>
<i>cxFFTpeakpowerHeaveLE</i>	<i>cxFFTpowerfracHeave</i>	<i>cxMeanSwayAMP</i>	<i>cxMeanHeaveUE</i>
<i>cxFFTdomfreqAvgUE</i>	<i>cxMeanSwayAMP</i>	<i>othTortuositySwayWrtUE</i>	<i>cxMeanSurgeLE</i>
<i>cxFFTdomfreqHeaveLE</i>	<i>othTortuositySwayWrtUE</i>	<i>cxMeanAbsJerkSway</i>	<i>cxMeanHeave</i>
<i>cxFFTdomfreqAvgLE</i>	<i>othTortuositySwayWrtLE</i>	<i>cxMinNormAMP</i>	<i>cxMaxPk2PkDescentSurge</i>
<i>cxCorrSurgeHeaveUE</i>	<i>othTortuositySurgeWrtLE</i>	<i>cxMinMovMeanSurge</i>	<i>cxMaxPk2PkAscentHeave</i>
<i>cxCorrSurgeHeaveAMP</i>	<i>cxMaxMovMeanSurge</i>	<i>cxStdSway</i>	<i>cxMaxPk2PkHitHeave</i>
<i>cxMinSurgeLE</i>	<i>cxNumZeroXdmNorm</i>	<i>cxMeanSurge</i>	<i>cxMeanAbsJerkHeave</i>
<i>cxMaxSurgeLE</i>	<i>cxMinNormAMP</i>	<i>cxMinSurgeLE</i>	<i>cxFFTdomfreqSurge</i>
<i>cxMaxHeaveLE</i>	<i>cxMeanPk2PkHitSurge</i>	<i>cxMeanPk2PkHitSway</i>	<i>cxFFTdomfreqAvg</i>
<i>cxMeanSurgeLE</i>	<i>cxCorrSurgeHeaveAMP</i>	<i>cxMeanNormAMP</i>	<i>cxNumDescentHeave</i>

Table S5.6. Feature selection (window size 1.5 s) from the 131 of 200 features presented in this study that did not require micro-event detection related information (Table S5.2). Four different feature selection techniques were used to select the top eight features (green). The features in the top fifteen in each method but not in the top eight are indicated in grey.

From Table S5.6, one sees that the following 8 features occur the most often, and were the ones selected for comparison with our model:

- *cxMinNormAMP* (all 4 methods)
- *cxMeanSurgeLE* (all 4 methods)
- *cxMeanSwayAMP* (3 methods, all filter)
- *cxMaxSurgeLE* (3 methods, all filter)
- *cxMaxHeaveLE* (2 methods, 1 wrapper, 1 filter)
- *cxFFTdomfreqAvg* (2 methods, 1 wrapper, 1 filter)
- *cxMeanAbsJerkSway* (2 methods, 2 filter). Sum of ranks = 13.
- *cxMinMovMeanSurge* (2 methods, 2 filter). Sum of ranks = 14. (higher top rank of 4th compared to same-ranked *othTortuositySwayWrtLE* which has a top rank of 5th: this criterion was used to break the tie).

S6.2 | Window size = 0.3 s, overlap = 50%

S6.2.1 | Summary statistics in Nathan et al. 2012

CfsSubset (best first)	Filter Methods		Wrapper Method
	Correlation (ranked)	InfoGain (ranked)	WrapperSubsetEval (J48, best first)
<i>meanSurge</i>	<i>meanSurge</i>	<i>minSurge</i>	<i>meanSurge</i>
<i>meanNorm</i>	<i>minSurge</i>	<i>meanSurge</i>	<i>meanSway</i>
<i>stdSway</i>	<i>acorrHeave</i>	<i>odba</i>	<i>stdSway</i>
<i>skewSurge</i>	<i>corrHeaveSurge</i>	<i>stdSway</i>	<i>stdHeave</i>
<i>kurtHeave</i>	<i>odba</i>	<i>acorrHeave</i>	<i>kurtSway</i>
<i>kurtNorm</i>	<i>acorrNorm</i>	<i>acorrSurge</i>	<i>maxNorm</i>
<i>minSurge</i>	<i>stdSway</i>	<i>meanNorm</i>	<i>minSway</i>
<i>minSway</i>	<i>kurtHeave</i>	<i>kurtHeave</i>	<i>minHeave</i>
<i>acorrSurge</i>	<i>stdSurge</i>	<i>acorrNorm</i>	<i>acorrSurge</i>
<i>acorrSway</i>	<i>acorrSurge</i>	<i>circVar_incli</i>	<i>acorrSway</i>
<i>acorrHeave</i>	<i>meanNorm</i>	<i>stdSurge</i>	<i>acorrHeave</i>
<i>acorrNorm</i>	<i>stdNorm</i>	<i>corrHeaveSurge</i>	<i>trendHeave</i>
<i>trendHeave</i>	<i>minSway</i>	<i>minSway</i>	<i>corrSwayHeave</i>
<i>corrHeaveSurge</i>	<i>stdHeave</i>	<i>stdNorm</i>	<i>corrHeaveSurge</i>
<i>odba</i>	<i>maxSway</i>	<i>acorrSway</i>	<i>odba</i>

Table S5.7. Feature selection (window size 0.3 s) from the 38 summary statistics presented by Nathan et al. 2012 (Table S5.4). Four different feature selection techniques were used to select the top eight features (green). The features in the top fifteen in each method but not in the top eight are indicated in grey.

From Table S5.7, one sees that the following 8 features occur the most often, and were the ones selected for comparison with our model:

- *meanSurge* (all 4 methods)
- *stdSway* (all 4 methods)
- *acorrHeave* (all 4 methods)
- *acorrSurge* (all 4 methods)
- *minSway* (all 4 methods)
- *corrHeaveSurge* (all 4 methods)
- *odba* (all 4 methods)
- *acorrSway* (3 methods, 1 wrapper, 2 filter)

S6.2.2 | Feature family from this study

CfsSubset (best first)	Filter Methods		Wrapper Method
	Correlation (ranked)	InfoGain (ranked)	WrapperSubsetEval (J48, best first)
<i>cxMinSurgeAMP</i>	<i>cxMeanSurgeLE</i>	<i>cxMaxSurgeLE</i>	<i>cxMinSwayAMP</i>
<i>cxMinSwayAMP</i>	<i>cxMaxSurgeLE</i>	<i>cxMeanSurgeLE</i>	<i>cxMinNormAMP</i>
<i>cxMinHeaveAMP</i>	<i>cxMinMovMeanSurge</i>	<i>cxMinSurgeLE</i>	<i>cxMaxSurgeAMP</i>
<i>cxMinNormAMP</i>	<i>cxMinSurgeLE</i>	<i>cxMinMovMeanSurge</i>	<i>cxMaxSwayAMP</i>
<i>cxMaxSwayAMP</i>	<i>cxMeanSurge</i>	<i>cxMeanSurge</i>	<i>cxMaxHeaveAMP</i>
<i>cxStdHeaveLE</i>	<i>cxMaxMovMeanSurge</i>	<i>cxMinNormAMP</i>	<i>cxRangeNormAMP</i>
<i>cxFFTpeakpowerSwayLE</i>	<i>cxMinSwayAMP</i>	<i>cxMaxMovMeanSurge</i>	<i>cxStdSwayUE</i>
<i>cxFFTdomfreqSwayUE</i>	<i>cxFFTpowerfracHeave</i>	<i>cxMinSwayAMP</i>	<i>cxStdHeaveUE</i>
<i>cxCorrSurgeHeaveUE</i>	<i>cxFFTpeakpowerHeave</i>	<i>cxMeanSwayAMP</i>	<i>cxMinSwayUE</i>
<i>cxCorrSurgeHeaveAMP</i>	<i>cxMinNormAMP</i>	<i>cxMeanAbsJerkSway</i>	<i>cxMinSwayLE</i>
<i>cxMinSurgeLE</i>	<i>cxCorrSurgeHeave</i>	<i>othTortuositySwayWrtLE</i>	<i>cxMaxSurgeLE</i>
<i>cxMinSwayLE</i>	<i>cxMinSurgeUE</i>	<i>othTortuositySwayWrtUE</i>	<i>cxMaxSwayLE</i>
<i>cxMaxSurgeLE</i>	<i>othTortuositySurgeWrtLE</i>	<i>cxStdSway</i>	<i>cxMaxHeaveLE</i>
<i>cxMaxHeaveLE</i>	<i>cxMinSurgeAMP</i>	<i>cxMaxSwayAMP</i>	<i>cxStdHeaveAMP</i>
<i>cxMeanSurgeLE</i>	<i>cxFFTdomfreqAvg</i>	<i>cxMeanNormAMP</i>	<i>cxMeanSway</i>

Table S5.8. Feature selection (window size 0.3 s) from the 131 of 200 features presented in this study that did not require micro-event detection related information (Table S5.2). Four different feature selection techniques were used to select the top eight features (green). The features in the top fifteen in each method but not in the top eight are indicated in grey.

From Table S5.8, one sees that the following 8 features occur the most often, and were the ones selected for comparison with our model:

- *cxMinSwayAMP* (all 4 methods)
- *cxMinNormAMP* (all 4 methods)
- *cxMaxSurgeLE* (all 4 methods)
- *cxMaxSwayAMP* (3 methods; 2 filter, 1 wrapper)
- *cxMinSurgeLE* (3 methods; all filter)
- *cxMeanSurgeLE* (3 methods; all filter)
- *cxMinSwayLE* (2 methods; 1 wrapper, 1 filter)
- *cxMaxHeaveLE* (2 methods; 1 wrapper, 1 filter).

Appendix S5.7 | Individual-wise confusion matrices for the eight-feature LR micro-event identification model

Individual-wise confusion matrices for the micro-event identification model with eight features (Table S5.3) using logistic regression at each node of the hierarchical classification tree (Figure 5.2, main text) that achieved the best overall results are presented in Table S5.9.

Individual #	Confusion matrix ($w = 1.5$ s)					
3	<i>Predicted</i>					
		1AD	2AD	Search	Chew	
	<i>Actual</i>	1AD	661	79	107	43
		2AD	0	776	114	0
		Search	34	44	790	22
Chew		36	41	99	714	
4	<i>Predicted</i>					
		1AD	2AD	Search	Chew	
	<i>Actual</i>	1AD	682	103	69	36
		2AD	7	883	0	0
		Search	432	147	311	0
Chew		0	0	0	890	
5	<i>Predicted</i>					
		1AD	2AD	Search	Chew	
	<i>Actual</i>	1AD	796	31	58	5
		2AD	0	890	0	0
		Search	10	140	697	43
Chew		2	0	64	824	
6	<i>Predicted</i>					
		1AD	2AD	Search	Chew	
	<i>Actual</i>	1AD	654	21	173	42
		2AD	155	647	88	0
		Search	0	50	805	35
Chew		66	0	28	796	
8	<i>Predicted</i>					
		1AD	2AD	Search	Chew	
	<i>Actual</i>	1AD	764	82	28	16
		2AD	77	813	0	0
		Search	0	39	851	0
Chew		40	70	0	780	

		<i>Predicted</i>				
		1AD	2AD	Search	Chew	
9	<i>Actual</i>	1AD	814	34	11	31
		2AD	37	853	0	0
		Search	17	9	838	26
		Chew	52	0	75	763

		<i>Predicted</i>				
		1AD	2AD	Search	Chew	
10	<i>Actual</i>	1AD	829	13	37	11
		2AD	105	785	0	0
		Search	21	32	694	143
		Chew	61	6	34	789

Table S5.9. Individual-wise confusion matrices for the micro-event identification model for which LOIO results are presented in Table 5.2 of the main text).

Individual #	1AD			2AD			Search			Chew			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
3	74.3	97.4	90.4	87.2	93.9	82.6	88.8	88	71.2	80.2	97.6	91.7	82.6
4	76.6	83.6	60.8	99.2	90.6	77.9	34.9	97.4	81.8	100	98.7	96.1	77.7
5	89.4	99.6	98.5	100	93.6	83.9	78.3	95.4	85.1	92.6	98.2	94.5	90.1
6	73.5	91.7	74.7	72.7	97.3	90.1	90.4	89.2	73.6	89.4	97.1	91.2	81.5
8	85.8	95.6	86.7	91.3	92.8	81	95.6	99	96.8	87.6	99.4	98	90.1
9	91.5	96	88.5	95.8	98.4	95.2	94.2	96.8	90.7	85.7	97.9	93	91.8
10	93.1	93	81.6	88.2	98.1	93.9	78	97.3	90.7	88.7	94.2	83.7	87

Table S5.10. Individual-wise micro-event identification performance corresponding to the confusion matrices in Table S5.9.

Appendix S5.8 | Comparison of micro-event identification with classical machine learning

(next page)

Model		1AD			2AD			Search			Chew			Overall Accuracy (%)
		Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
Micro-event identification		83.5 ± 7.8	93.8 ± 4.8	83.0 ± 11.4	90.6 ± 8.7	95.0 ± 2.8	86.4 ± 6.2	80.0 ± 19.5	94.7 ± 4.0	84.3 ± 8.7	89.2 ± 5.7	97.6 ± 1.5	92.6 ± 4.3	85.8 ± 5.3
Classical machine learning														
RF	$w = 1.5$ s, $fF1$	80.4 ± 13.4	90.7 ± 3.9	74.8 ± 9.7	59.7 ± 27.4	91.7 ± 7.7	77.5 ± 13.3	75.0 ± 21.2	89.2 ± 8.1	73.1 ± 12.6	87.2 ± 6.5	95.8 ± 3.6	88.5 ± 8.0	75.6 ± 6.9
	$w = 1.5$ s, $fF2$	84.4 ± 9.6	92.1 ± 5.7	79.5 ± 11.2	77.8 ± 16.8	91.4 ± 13.6	82.9 ± 16.9	75.4 ± 21.6	92.7 ± 3.7	78.5 ± 8.6	66.7 ± 22.1	91.9 ± 6.7	75.2 ± 12.0	76.1 ± 10.0
	$w = 0.3$ s, $fF1'$	72.8 ± 9.5	88.2 ± 5.6	69.0 ± 8.9	61.7 ± 15.9	89.7 ± 5.7	69.2 ± 8.5	76.2 ± 13.1	92.2 ± 4.5	78.1 ± 7.9	80.3 ± 6.3	93.6 ± 3.0	81.5 ± 6.7	72.7 ± 3.8
	$w = 0.3$ s, $fF2'$	74.4 ± 8.6	89.8 ± 5.3	72.4 ± 10.6	65.9 ± 9.5	87.2 ± 6.9	64.9 ± 9.9	71.2 ± 14.5	85.8 ± 7.8	65.4 ± 14.3	44.0 ± 14.7	89.1 ± 2.9	56.2 ± 12.5	63.9 ± 5.9
SVM	$w = 1.5$ s, $fF1$	77.4 ± 14.2	93.7 ± 4.9	81.6 ± 10.4	68.1 ± 25.5	89.2 ± 7.9	73.4 ± 13.6	75.2 ± 17.8	90.6 ± 10.8	79.8 ± 18.1	86.4 ± 12.0	95.5 ± 3.6	87.7 ± 8.7	76.8 ± 6.2
	$w = 1.5$ s, $fF2$	82.3 ± 10.4	93.2 ± 4.0	81.6 ± 8.5	82.9 ± 13.4	90.9 ± 7.0	78.0 ± 13.3	76.8 ± 15.9	90.0 ± 5.0	72.8 ± 11.3	60.0 ± 11.1	93.2 ± 5.3	76.0 ± 13.5	75.5 ± 4.9
	$w = 0.3$ s, $fF1'$	66.4 ± 13.6	89.1 ± 5.5	68.4 ± 8.0	54.7 ± 19.3	88.6 ± 6.7	65.2 ± 9.5	79.0 ± 12.6	88.7 ± 8.1	73.2 ± 13.0	76.7 ± 9.4	92.6 ± 3.5	78.3 ± 7.3	69.2 ± 5.3
	$w = 0.3$ s, $fF2'$	73.7 ± 8.0	87.8 ± 7.2	69.2 ± 13.1	69.2 ± 10.4	83.4 ± 9.1	60.8 ± 12.9	66.4 ± 20.5	82.4 ± 12.6	62.7 ± 17.2	23.4 ± 18.7	90.6 ± 6.2	42.5 ± 19.0	58.2 ± 5.9

Table S5.11. **Results with the balanced dataset during leave-one-individual-out cross-validation.** The performance of micro-event identification (eight features with logistic regression) is compared with classical machine learning (CML) approaches employing the same number of features. For CML, to ensure that comparison was varied and rigorous, results were computed for two different machine learning algorithms (RF: Random Forest; SVM: Support Vector Machine with linear kernel), moving-window sizes ($w = 1.5$ s and $w = 0.3$ s), and feature families ($fF1$, $fF1'$: sets of top eight features from feature family in Nathan et al. 2012 for $w = 1.5$ s and $w = 0.3$ s, respectively; $fF2$, $fF2'$: sets of top eight features from feature family presented in this study for $w = 1.5$ s and $w = 0.3$ s, respectively).

Model		1AD			2AD			Search			Chew			Overall Accuracy (%)
		Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
Micro-event identification		86.9 ± 6.6	88.5 ± 12.5	78.5 ± 19.6	89.5 ± 11.8	92.8 ± 4.7	68.9 ± 18.3	61.3 ± 21.1	96.6 ± 3.6	77.3 ± 20.8	92.3 ± 7.9	94.3 ± 6.0	79.1 ± 15.7	79.2 ± 11.5
Classical machine learning														
RF	$w = 1.5$ s, $fg1$	84.3 ± 8.3	90.8 ± 7.0	79.5 ± 15.5	66.6 ± 25.4	92.9 ± 8.6	73.8 ± 20.6	67.7 ± 22.3	91.8 ± 10.2	70.3 ± 24.7	90.0 ± 9.6	92.9 ± 4.7	74.3 ± 20.0	75.8 ± 14.6
	$w = 1.5$ s, $fg2$	86.6 ± 6.8	89.8 ± 9.4	80.1 ± 13.2	72.2 ± 21.8	95.5 ± 5.8	80.4 ± 17.5	65.5 ± 26.7	96.6 ± 2.5	77.2 ± 18.9	78.7 ± 14.3	87.0 ± 6.5	59.6 ± 29.3	77.0 ± 10.6
	$w = 0.3$ s, $fg1'$	73.9 ± 8.9	88.1 ± 5.7	72.0 ± 11.9	60.3 ± 16.9	89.0 ± 7.5	60.0 ± 15.5	66.3 ± 17.1	94.3 ± 5.2	75.2 ± 19.0	85.9 ± 6.4	89.8 ± 4.7	68.9 ± 19.2	71.1 ± 8.4
	$w = 0.3$ s, $fg2'$	76.9 ± 6.3	88.5 ± 7.7	74.8 ± 11.0	60.3 ± 13.6	90.7 ± 3.6	61.2 ± 12.4	52.3 ± 18.2	92.7 ± 4.4	64.7 ± 22.6	64.1 ± 13.7	81.1 ± 4.9	49.6 ± 28.9	64.8 ± 8.5
SVM	$w = 1.5$ s, $fg1$	81.2 ± 10.5	92.7 ± 6.7	81.3 ± 15.6	75.3 ± 18.6	85.6 ± 13.0	57.5 ± 20.9	37.7 ± 28.8	96.6 ± 3.8	66.0 ± 33.1	92.0 ± 10.1	89.5 ± 7.9	67.5 ± 23.1	72.3 ± 12.8
	$w = 1.5$ s, $fg2$	86.3 ± 8.7	90.9 ± 7.0	80.3 ± 13.6	75.6 ± 16.9	93.3 ± 5.7	73.7 ± 18.0	40.5 ± 18.6	97.7 ± 1.0	74.3 ± 24.1	79.6 ± 12.7	80.1 ± 7.3	53.0 ± 27.2	72.1 ± 7.8
	$w = 0.3$ s, $fg1'$	75.0 ± 11.7	86.5 ± 7.1	69.9 ± 14.0	46.5 ± 23.5	87.6 ± 10.6	52.9 ± 19.3	52.4 ± 23.1	92.6 ± 9.6	73.9 ± 25.9	84.5 ± 7.6	86.8 ± 5.1	61.6 ± 23.1	64.4 ± 13.2
	$w = 0.3$ s, $fg2'$	76.5 ± 6.9	84.5 ± 10.2	68.8 ± 12.5	61.8 ± 18.8	89.1 ± 5.6	58.3 ± 14.8	15.3 ± 9.6	98.7 ± 1.2	71.9 ± 23.1	71.3 ± 15.2	71.5 ± 3.5	44.0 ± 26.0	59.2 ± 10.2

Table S5.12. **Results with the raw, imbalanced dataset during leave-one-individual-out cross-validation.** The performance of micro-event identification (eight features with logistic regression) is compared with classical machine learning (CML) approaches employing the same number of features. For CML, to ensure that comparison was varied and rigorous, results were computed for two different machine learning algorithms (RF: Random Forest; SVM: Support Vector Machine with linear kernel), moving-window sizes ($w = 1.5$ s and $w = 0.3$ s), and feature families ($fg1, fg1'$: sets of top eight features from feature family in Nathan et al. 2012 for $w = 1.5$ s and $w = 0.3$ s, respectively; $fg2, fg2'$: sets of top eight features from feature family presented in this study for $w = 1.5$ s and $w = 0.3$ s, respectively).

Appendix S5.9 | Robustness of micro-event identification model to change in model parameters

Detailed results with micro-event identifications obtained during LOIO when model parameters (Table 5.3, main text) were varied.

M_{env} (samples @100Hz)	1AD			2AD			Search			Chew			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
5	81.4 ± 8.0	91.1 ± 4.8	76.8 ± 10.4	84.6 ± 10.8	94.6 ± 3.3	85.0 ± 6.5	79.8 ± 21.5	94.7 ± 3.7	84.2 ± 7.7	89.2 ± 6.9	97.9 ± 1.2	93.4 ± 3.9	83.8 ± 4.6
7	83.2 ± 8.4	93.6 ± 4.8	82.3 ± 11.1	88.4 ± 8.5	93.8 ± 4.1	83.8 ± 8.5	78.3 ± 19.7	94.0 ± 4.3	82.7 ± 8.9	86.5 ± 7.8	97.4 ± 1.8	91.9 ± 5.3	84.1 ± 5.3
9	84.2 ± 8.6	94.0 ± 5.1	83.8 ± 12.0	89.2 ± 9.3	94.7 ± 3.2	85.7 ± 6.9	79.3 ± 19.8	94.2 ± 4.5	83.6 ± 8.8	88.5 ± 6.7	97.5 ± 1.4	92.4 ± 3.9	85.3 ± 5.5
11	83.3 ± 7.8	93.8 ± 4.7	82.9 ± 11.0	90.5 ± 8.9	95.1 ± 2.6	86.7 ± 5.9	80.1 ± 19.4	94.6 ± 3.9	83.9 ± 8.4	89.1 ± 5.6	97.5 ± 1.4	92.3 ± 4.1	85.8 ± 5.1
13	83.2 ± 8.8	93.2 ± 5.1	81.7 ± 11.3	90.8 ± 7.6	94.9 ± 2.8	86.2 ± 6.6	80.0 ± 20.5	94.5 ± 4.2	84.4 ± 8.8	86.1 ± 8.3	97.4 ± 1.7	92.0 ± 4.9	85.0 ± 4.7
15	82.3 ± 8.7	93.7 ± 4.5	82.4 ± 10.6	91.0 ± 7.2	94.4 ± 2.6	85.0 ± 5.9	81.3 ± 18.0	95.0 ± 3.7	85.5 ± 8.4	87.2 ± 8.6	97.6 ± 1.6	92.4 ± 4.9	85.5 ± 4.4
17	83.6 ± 10.3	93.2 ± 4.9	81.8 ± 11.3	91.0 ± 8.2	94.2 ± 2.7	84.4 ± 6.0	80.2 ± 18.6	94.7 ± 4.3	84.9 ± 9.5	84.4 ± 10.6	97.6 ± 1.7	92.3 ± 5.4	84.8 ± 5.0

Table S5.13. **Effect of varying M_{env} (window size for envelope computation) on model performance.** The value used in the model was $M_{env} = 11$ samples. All other model parameters were kept at their reference values (Table 5.3, main text).

f_c (Hz)	1AD			2AD			Search			Chew			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
2.5	83.2 ± 8.0	93.5 ± 5.4	82.6 ± 11.5	91.2 ± 7.3	95.3 ± 2.5	87.2 ± 6.0	79.1 ± 21.4	94.7 ± 4.0	84.0 ± 8.5	88.9 ± 6.1	97.3 ± 1.6	91.6 ± 4.5	85.6 ± 5.4
3	83.3 ± 7.9	93.4 ± 5.0	82.1 ± 11.3	90.2 ± 8.8	95.1 ± 2.8	86.7 ± 6.6	80.1 ± 20.7	94.7 ± 3.9	84.3 ± 8.4	88.3 ± 6.0	97.4 ± 1.2	92.0 ± 3.7	85.5 ± 5.4
4	83.5 ± 8.1	93.5 ± 5.1	82.4 ± 11.7	90.2 ± 8.0	95.4 ± 2.5	87.3 ± 5.8	80.4 ± 20.2	94.6 ± 4.1	84.0 ± 8.8	89.0 ± 5.5	97.6 ± 1.3	92.5 ± 3.8	85.8 ± 5.4
5	83.7 ± 8.2	93.8 ± 4.8	83.0 ± 11.0	90.4 ± 8.2	95.1 ± 2.8	86.7 ± 6.4	80.4 ± 19.6	94.7 ± 4.1	84.5 ± 8.6	88.8 ± 5.6	97.6 ± 1.3	92.4 ± 3.8	85.8 ± 5.0
6	82.7 ± 7.7	93.8 ± 4.8	83.0 ± 11.2	90.4 ± 8.4	95.3 ± 2.8	87.2 ± 6.6	80.7 ± 19.9	94.2 ± 4.0	83.3 ± 8.2	88.2 ± 6.9	97.4 ± 1.7	91.9 ± 4.6	85.5 ± 4.7
7	83.7 ± 7.0	93.6 ± 5.2	82.9 ± 11.6	90.2 ± 8.4	95.0 ± 2.6	86.3 ± 5.8	80.3 ± 19.7	94.8 ± 3.8	84.6 ± 8.0	88.7 ± 5.6	97.5 ± 1.4	92.3 ± 4.0	85.7 ± 5.1
7.5	82.6 ± 7.7	93.8 ± 5.1	83.0 ± 11.6	89.7 ± 8.3	95.2 ± 2.8	86.8 ± 6.5	80.2 ± 19.8	93.8 ± 4.2	82.5 ± 8.4	88.2 ± 6.8	97.4 ± 1.7	92.0 ± 4.7	85.1 ± 4.9

Table S5.14. **The effect of varying f_c (low-pass cut-off frequency for smoothing envelope) on model performance.** The value used in the model was $f_c = 5$ Hz. All other model parameters were kept at their reference values (Table 5.3, main text).

th_{amp} (g)	1AD			2AD			Search			Chew			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
0.05	83.8 ± 7.3	93.7 ± 5.0	82.9 ± 11.0	89.9 ± 7.8	94.7 ± 3.3	85.8 ± 7.2	79.9 ± 18.7	94.7 ± 3.9	84.6 ± 7.9	87.1 ± 5.8	97.2 ± 1.4	91.3 ± 4.1	85.1 ± 4.7
0.075	83.3 ± 7.4	93.7 ± 4.6	82.7 ± 10.7	89.3 ± 8.5	94.6 ± 3.2	85.4 ± 6.8	79.6 ± 18.8	94.6 ± 4.1	84.3 ± 8.6	88.2 ± 5.9	97.3 ± 1.5	91.7 ± 4.4	85.1 ± 4.7
0.1	83.3 ± 7.9	93.7 ± 5.0	82.7 ± 11.2	90.5 ± 7.9	95.2 ± 2.5	87.0 ± 5.8	80.2 ± 19.9	94.6 ± 4.1	84.0 ± 8.7	88.9 ± 5.9	97.5 ± 1.6	92.2 ± 4.3	85.7 ± 5.3
0.125	85.0 ± 9.2	94.2 ± 4.6	84.0 ± 11.2	91.2 ± 8.8	95.3 ± 2.7	87.2 ± 6.2	79.7 ± 18.8	94.8 ± 4.2	84.4 ± 9.5	89.2 ± 5.9	97.4 ± 1.6	92.2 ± 4.5	86.3 ± 5.8
0.15	84.5 ± 8.9	93.6 ± 3.5	82.0 ± 8.9	88.5 ± 8.3	94.2 ± 4.0	85.1 ± 8.2	79.9 ± 17.8	94.1 ± 3.4	82.7 ± 7.1	85.3 ± 9.7	97.5 ± 1.6	92.4 ± 4.7	84.6 ± 3.2

Table S5.15. **The effect of varying th_{amp} (amplitude threshold for segmentation algorithm) on model performance.** The value used in the model was $th_{amp} = 0.1$ g. All other model parameters were kept at their reference values (Table 5.3, main text).

th_{std} (g)	1AD			2AD			Search			Chew			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
0.005	83.2 ± 7.9	93.4 ± 4.9	82.1 ± 11.0	90.4 ± 9.2	95.1 ± 2.5	86.6 ± 5.8	79.2 ± 19.6	94.5 ± 4.0	83.9 ± 8.4	88.7 ± 6.0	97.5 ± 1.5	92.2 ± 4.3	85.4 ± 5.1
0.0075	83.2 ± 7.6	93.4 ± 4.8	82.0 ± 11.0	90.2 ± 8.8	95.1 ± 2.6	86.7 ± 6.1	80.1 ± 19.5	94.6 ± 4.1	84.0 ± 8.7	88.7 ± 5.8	97.6 ± 1.5	92.5 ± 4.1	85.5 ± 5.1
0.01	83.1 ± 8.2	93.6 ± 4.9	82.5 ± 11.3	90.0 ± 9.4	94.9 ± 2.5	86.1 ± 5.5	79.5 ± 19.7	94.5 ± 4.4	83.9 ± 9.2	88.9 ± 5.6	97.5 ± 1.4	92.4 ± 3.9	85.4 ± 5.3
0.0125	83.4 ± 7.7	93.7 ± 4.8	82.7 ± 11.0	90.3 ± 8.7	95.2 ± 2.6	86.9 ± 5.9	79.9 ± 19.7	94.5 ± 4.2	83.9 ± 8.9	89.0 ± 5.9	97.5 ± 1.5	92.3 ± 4.0	85.6 ± 5.2
0.015	83.9 ± 7.8	93.5 ± 4.9	82.3 ± 11.2	90.4 ± 8.4	95.2 ± 2.7	86.9 ± 6.2	79.8 ± 20.1	94.8 ± 4.1	84.6 ± 8.8	88.8 ± 5.5	97.5 ± 1.4	92.3 ± 3.9	85.7 ± 5.2

Table S5.16. **Effect of varying th_{std} (standard-deviation threshold for segmentation algorithm) on model performance.** The value used in the model was $th_{std} = 0.01$ g. All other model parameters were kept at their reference values (Table 5.3, main text).

M_{seg} (samples @100Hz)	1AD			2AD			Search			Chew			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
5	83.2 ± 8.0	93.4 ± 4.9	82.1 ± 11.3	90.0 ± 9.2	95.0 ± 2.9	86.4 ± 6.7	79.4 ± 20.7	94.5 ± 4.1	83.6 ± 8.7	88.8 ± 6.0	97.5 ± 1.5	92.4 ± 4.2	85.4 ± 5.5
7	83.0 ± 8.4	93.6 ± 4.9	82.5 ± 11.3	90.4 ± 8.5	95.0 ± 2.6	86.4 ± 5.8	79.5 ± 20.4	94.6 ± 4.2	83.9 ± 8.9	89.1 ± 5.7	97.5 ± 1.5	92.3 ± 4.2	85.5 ± 5.4
9	82.9 ± 7.9	93.6 ± 4.8	82.4 ± 11.1	90.4 ± 9.3	95.0 ± 2.8	86.5 ± 6.3	79.7 ± 19.8	94.6 ± 4.1	84.0 ± 8.6	88.7 ± 5.4	97.4 ± 1.5	92.1 ± 4.1	85.4 ± 5.1
11	83.2 ± 7.5	93.7 ± 4.7	82.7 ± 10.9	90.6 ± 8.3	95.2 ± 2.5	86.8 ± 5.7	80.4 ± 19.5	94.7 ± 3.9	84.2 ± 8.5	89.3 ± 5.6	97.6 ± 1.4	92.6 ± 3.9	85.8 ± 5.1
13	83.3 ± 7.6	93.6 ± 4.7	82.5 ± 11.0	90.3 ± 8.6	95.2 ± 2.6	86.7 ± 5.9	80.7 ± 19.5	94.7 ± 3.9	84.1 ± 8.6	89.0 ± 5.6	97.6 ± 1.2	92.6 ± 3.6	85.8 ± 5.3
15	83.4 ± 7.4	93.7 ± 4.8	82.7 ± 10.9	90.7 ± 8.5	95.2 ± 2.5	86.7 ± 5.5	79.9 ± 19.7	94.7 ± 4.0	84.3 ± 8.6	88.9 ± 6.0	97.5 ± 1.6	92.2 ± 4.4	85.7 ± 5.2
17	83.7 ± 7.8	93.6 ± 4.8	82.6 ± 11.2	90.3 ± 9.1	95.3 ± 2.7	87.1 ± 6.1	80.3 ± 19.4	94.8 ± 3.8	84.6 ± 8.2	89.4 ± 5.3	97.5 ± 1.4	92.4 ± 4.0	85.9 ± 5.3

Table S5.17. **Effect of varying M_{seg} (window size to compute standard deviation in segmentation algorithm) on model performance.** The value used in the model was $M_{seg} = 11$ samples. All other model parameters were kept at their reference values (Table 5.3, main text).

w (s)	1AD			2AD			Search			Chew			Overall Accuracy (%)
	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	Sen. (%)	Spec. (%)	Prec. (%)	
1	79.2 ± 8.5	91.7 ± 3.4	76.5 ± 8.9	86.1 ± 9.2	93.6 ± 3.7	83.0 ± 7.2	76.5 ± 16.4	93.0 ± 3.9	79.0 ± 8.6	85.0 ± 7.5	97.3 ± 1.2	91.4 ± 4.1	81.7 ± 3.8
1.1	80.7 ± 8.6	91.7 ± 3.4	76.9 ± 9.1	86.1 ± 10.1	94.1 ± 3.3	83.9 ± 6.6	76.9 ± 16.5	93.0 ± 4.0	79.4 ± 8.6	85.6 ± 7.6	97.6 ± 1.3	92.3 ± 4.2	82.3 ± 3.6
1.2	81.3 ± 8.4	93.1 ± 3.8	80.5 ± 9.3	86.5 ± 11.0	93.8 ± 3.5	83.4 ± 6.7	79.2 ± 14.9	93.4 ± 4.2	81.1 ± 8.9	86.5 ± 7.1	97.5 ± 1.5	91.8 ± 4.8	83.4 ± 4.2
1.3	82.0 ± 7.7	93.2 ± 4.4	81.0 ± 10.5	87.1 ± 11.2	94.1 ± 3.4	84.1 ± 7.0	80.2 ± 17.9	93.9 ± 3.6	81.7 ± 8.4	86.7 ± 6.8	97.6 ± 1.3	92.3 ± 4.1	84.0 ± 4.6
1.4	83.0 ± 7.4	93.1 ± 5.2	81.7 ± 12.3	88.6 ± 9.4	94.7 ± 3.2	85.8 ± 7.2	80.0 ± 17.8	94.3 ± 4.1	83.5 ± 8.5	87.7 ± 6.6	97.6 ± 1.4	92.4 ± 4.1	84.8 ± 4.7
1.5	83.0 ± 8.2	93.5 ± 4.8	82.2 ± 11.1	90.0 ± 8.7	95.2 ± 2.7	86.9 ± 6.1	80.0 ± 19.6	94.6 ± 4.3	84.1 ± 9.0	89.0 ± 5.9	97.5 ± 1.5	92.2 ± 4.2	85.5 ± 5.3
1.6	84.8 ± 8.1	93.8 ± 5.0	83.3 ± 10.7	90.2 ± 8.7	95.2 ± 2.7	86.9 ± 6.3	79.9 ± 20.7	95.0 ± 4.4	85.4 ± 9.6	89.8 ± 5.8	97.7 ± 1.7	93.0 ± 4.5	86.2 ± 4.9
1.7	86.4 ± 8.7	94.5 ± 6.2	86.1 ± 12.2	91.7 ± 8.6	95.3 ± 2.8	87.4 ± 6.5	79.9 ± 23.1	95.1 ± 5.0	86.7 ± 10.7	90.6 ± 6.0	97.9 ± 1.6	93.8 ± 4.4	87.2 ± 5.6
1.8	86.5 ± 9.2	94.4 ± 5.8	85.8 ± 11.5	90.8 ± 8.1	95.0 ± 3.2	86.6 ± 7.2	79.5 ± 23.9	95.2 ± 4.9	87.5 ± 11.0	90.5 ± 7.4	97.8 ± 2.3	93.8 ± 5.7	86.8 ± 5.3
1.9	86.1 ± 9.9	94.6 ± 6.1	86.3 ± 11.9	91.1 ± 7.9	94.8 ± 3.0	86.1 ± 6.9	79.5 ± 23.9	95.0 ± 4.9	87.3 ± 11.5	90.5 ± 7.3	98.0 ± 2.5	94.3 ± 6.0	86.7 ± 5.3
2	86.6 ± 10.0	94.7 ± 5.6	86.2 ± 11.2	92.0 ± 8.1	94.8 ± 3.4	86.3 ± 7.3	79.8 ± 23.0	95.2 ± 4.8	87.7 ± 11.2	89.0 ± 10.4	97.8 ± 2.4	93.9 ± 5.9	86.9 ± 4.9

Table S5.18. **The effect of varying w (window size describing context around main event) on model performance.** The value used in the model was $w = 1.5$ s. All other model parameters were kept at their reference values (Table 5.3, main text).

References

- Béliveau, A., Spencer, G. T., Thomas, K. A., & Roberson, S. L. (1999). Evaluation of MEMS capacitive accelerometers. *IEEE Design & Test of Computers*, 16(4), 48-56.
- Bourke, A. K., O'Brien, J. V., & Lyons, G. M. (2007). Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait & posture*, 26(2), 194-199.
- Branco, P., Torgo, L., & Ribeiro, R. P. (2016). A survey of predictive modeling on imbalanced domains. *ACM Computing Surveys (CSUR)*, 49(2), 31.
- Brown, D. D., Kays, R., Wikelski, M., Wilson, R., & Klimley, A. P. (2013). Observing the unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry*, 1(1), 20.
- Camomilla, V., Dumas, R., & Cappozzo, A. (2017). Human movement analysis: the soft tissue artefact issue.
- Carroll, G., Slip, D., Jonsen, I., & Harcourt, R. (2014). Supervised accelerometry analysis can identify prey capture by penguins at sea. *Journal of Experimental Biology*, 217(24), 4295-4302.
- Chakravarty, P., Cozzi, G., Ozgul, A., & Aminian, K. (2019a). A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods in Ecology and Evolution*, 10(6), 802-814.
- Chakravarty, P., Maalberg, M., Cozzi, G., Ozgul, A., & Aminian, K. (2019b). Behavioural compass: animal behaviour recognition using magnetometers. *Movement Ecology*, 7(1), 28.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321-357.
- Clutton-Brock, T. H., & Manser, M. (2016). Meerkats: cooperative breeding in the Kalahari. *Cooperative breeding in vertebrates*, 294, 317.

- Dawkins, M. S. (1989). Time budgets in red junglefowl as a baseline for the assessment of welfare in domestic fowl. *Applied Animal Behaviour Science*, 24(1), 77-80.
- Esseiva, J., Caon, M., Mugellini, E., Khaled, O. A., & Aminian, K. (2018, April). Feet fidgeting detection based on accelerometers using decision tree learning and gradient boosting. In *International Conference on Bioinformatics and Biomedical Engineering* (pp. 75-84). Springer, Cham.
- Ferraris, F., Grimaldi, U., & Parvis, M. (1995). Procedure for effortless in-field calibration of three-axis rate gyros and accelerometers. *Sensors and Materials*, 7, 311-311.
- Hammond, T. T., Springthorpe, D., Walsh, R. E., & Berg-Kirkpatrick, T. (2016). Using accelerometers to remotely and automatically characterize behavior in small animals. *Journal of Experimental Biology*, 219(11), 1618-1624.
- Hansson, G. A., Asterland, P., Holmer, N. G., & Skerfving, S. (2001). Validity and reliability of triaxial accelerometers for inclinometry in posture analysis. *Medical and Biological Engineering and Computing*, 39(4), 405-413.
- He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, 21(9), 1263-1284.
- Hill, A. V. (1938). The heat of shortening and the dynamic constants of muscle. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, 126(843), 136-195.
- Iwata, T., Sakamoto, K. Q., Takahashi, A., Edwards, E. W., Staniland, I. J., Trathan, P. N., & Naito, Y. (2012). Using a mandible accelerometer to study fine-scale foraging behavior of free-ranging Antarctic fur seals. *Marine Mammal Science*, 28(2), 345-357.
- Johnson Jr, C. R., Sethares, W. A., & Klein, A. G. (2011). *Software receiver design: build your own digital communication system in five easy steps*. Cambridge University Press.

- Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, *348*(6240), aaa2478.
- Koshy, C. S., Flores, P., & Lankarani, H. M. (2013). Study of the effect of contact force model on the dynamic response of mechanical systems with dry clearance joints: computational and experimental approaches. *Nonlinear Dynamics*, *73*(1-2), 325-338.
- Martiskainen, P., Järvinen, M., Skön, J. P., Tiirikainen, J., Kolehmainen, M., & Mononen, J. (2009). Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Applied animal behaviour science*, *119*(1-2), 32-38.
- McFadden, P. D., & Smith, J. D. (1984). Vibration monitoring of rolling element bearings by the high-frequency resonance technique—a review. *Tribology international*, *17*(1), 3-10.
- Müller, R., & Schrader, L. (2003). A new method to measure behavioural activity levels in dairy cows. *Applied Animal Behaviour Science*, *83*(4), 247-258.
- Naito, Y., Bornemann, H., Takahashi, A., McIntyre, T., & Plötz, J. (2010). Fine-scale feeding behavior of Weddell seals revealed by a mandible accelerometer. *Polar Science*, *4*(2), 309-316.
- 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.
- Pesaran, H. H., & Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters*, *58*(1), 17-29.
- Phillips, C. L., Parr, J. M., & Riskin, E. A. (2003). *Signals, systems, and transforms* (p. 209). Upper Saddle River: Prentice Hall.
- Robert-Coudert, Y., & Wilson, R. P. (2005). Trends and perspectives in animal-attached remote sensing. *Frontiers in Ecology and the Environment*, *3*(8), 437-444.

- Sakamoto, K. Q., Sato, K., Ishizuka, M., Watanuki, Y., Takahashi, A., Daunt, F., & Wanless, S. (2009). Can ethograms be automatically generated using body acceleration data from free-ranging birds?. *PloS one*, 4(4), e5379.
- Suzuki, I., Naito, Y., Folkow, L. P., Miyazaki, N., & Blix, A. S. (2009). Validation of a device for accurate timing of feeding events in marine animals. *Polar Biology*, 32(4), 667-671.
- Tandon, N., & Choudhury, A. (1999). A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology international*, 32(8), 469-480.
- Usui, S., & Amidror, I. (1982). Digital low-pass differentiation for biological signal processing. *IEEE Transactions on Biomedical Engineering*, (10), 686-693.
- Viviant, M., Trites, A. W., Rosen, D. A., Monestiez, P., & Guinet, C. (2010). Prey capture attempts can be detected in Steller sea lions and other marine predators using accelerometers. *Polar biology*, 33(5), 713-719.
- Watanabe, S., Izawa, M., Kato, A., Ropert-Coudert, Y., & Naito, Y. (2005). A new technique for monitoring the detailed behaviour of terrestrial animals: a case study with the domestic cat. *Applied Animal Behaviour Science*, 94(1-2), 117-131.
- Watanabe, Y. Y., & Takahashi, A. (2013). Linking animal-borne video to accelerometers reveals prey capture variability. *Proceedings of the National Academy of Sciences*, 110(6), 2199-2204.
- Wilson, R. P., White, C. R., Quintana, F., Halsey, L. G., Liebsch, N., Martin, G. R., & Butler, P. J. (2006). Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *Journal of Animal Ecology*, 75(5), 1081-1090.
- Wilson, R. P., Neate, A., Holton, M. D., Shepard, E. L., Scantlebury, D. M., Lambertucci, S. A., ... & Marks, N. (2018a). Luck in food finding affects individual performance and population trajectories. *Current Biology*, 28(23), 3871-3877.

Wilson, R. P., Holton, M. D., di Virgilio, A., Williams, H., Shepard, E. L., Lambertucci, S., ... & Srivastava, M. (2018b). Give the machine a hand: A Boolean time-based decision-tree template for rapidly finding animal behaviours in multisensor data.

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

CHAPTER 6

Behavioural Time- Energy Budgets and Body Mass-Dependent Foraging Strategies in Meerkats⁶

⁶ This study will be submitted as a journal paper titled, “Chakravarty P., Cozzi G., Harrison N., Maag N., Manser M., Ozgul A., Aminian K. *Behavioural time-energy budgets and foraging strategies in meerkats: a study with validated accelerometer-based behaviour recognition.*”

Abstract

1. While the effect of asymmetry in status (dominant versus subordinate individuals) on social behaviour is well-documented in meerkats (*Suricata suricatta*), little is known about the influence of asymmetry in body mass on foraging behaviour. Meerkats can differ by >20% in body mass, and forage throughout the day in cohesive groups in areas with limited prey, and yet past studies have found that foraging conflict is infrequent. Little is known about the strategies meerkats employ to avoid foraging conflict while still meeting their energy requirements, and this may have been because manually observing fine-scale foraging behaviour continuously for long durations is challenging.
2. I address this issue by employing miniature accelerometers to remotely observe four fine-scale foraging behaviours – searching, one-armed digging, two-armed digging, and prey chewing – along with three other behaviours – resting, vigilance, and running. I combine behaviour recognition and a new energy expenditure estimation approach to compare foraging strategy, foraging efficiency, and cooperative activity (vigilance) for individuals of different sex, age, and body mass. For this, we deployed accelerometers on 26 wild meerkats to record triaxial acceleration data at 50 Hz/axis continuously for up to 6 days.
3. I found that lighter females (<550 grams in our dataset) almost exclusively searched with little digging, while heavier females invested more in digging than in searching. Given prey distribution in the Kalahari, this implies that lighter females targeted low-value surface prey while heavier females targeted both high-value below-ground prey and low-value surface prey. I also found that females were vigilant for significantly longer compared to males.
4. I suggest the presence of an ‘evolutionarily stable foraging strategy’ wherein lighter females ‘settle’ for lower-value prey near the surface to avoid foraging conflict with heavier females, thereby also minimising energy losses due to intensive digging. That our results on sex-differences in vigilance contribution contrasted with those of previous studies suggests that the cooperative component of vigilance may lie in its duration rather than frequency of occurrence. Our results demonstrate the potential of combining accurate and robust behaviour recognition with energy expenditure estimation to uncover subtle behavioural differences associated with complex social and demographic processes.

6.1 | Introduction

The benefits of sociality in group-living animals – such as reduced predation risk (Gosling 1986) – are tempered by costs such as competition for resources (Ward & Webster 2016, chapter 4.2.2). Outcomes of contests are determined by asymmetries between individuals (Maynard Smith & Parker 1976), such as body mass differences. In group-living species, where individuals have full information on these asymmetries, a typical ‘evolutionarily stable strategy’ (ESS) is to let the asymmetry cue settle the contest without escalation (Maynard Smith & Parker 1976). For instance, asymmetry in social status leads to dominant individuals having reduced costs of competition since subordinate individuals are more likely to avoid rather than engage in conflict with them (Packer & Pusey 1985). However, little is known on how individuals in groups adjust to asymmetry in individual traits, such as body mass, when food resources are limited and space has to be shared.

In meerkats (*Suricata suricatta*), monomorphic group-living mongooses that breed cooperatively (MacDonald & Doolan 1997), social and environmental factors interact to produce individual variation in body mass (Ozgul et al. 2014) that can exceed 20% across adult individuals (Maag et al. 2019). Intense intrasexual competition for breeding positions favours selection for aggression and body mass in female meerkats (Hodge et al. 2008). Dominant females that monopolise group breeding attempts (Clutton-Brock et al. 1999a; Young et al. 2006) employ aggressive tactics to maintain their breeding monopoly (Doolan & MacDonald 1996a; O’Riain et al. 2000), and compete for food more frequently (Flower 2011). Subordinate females have strong incentive to minimise conflict, since they subsequently have higher rates of reproductive senescence if they experience greater aggression (Sharp & Clutton-Brock 2011). The fact that, overall, meerkats compete infrequently for food (Flower 2011; Doolan & MacDonald 1996b), however, suggests that individuals may be taking known asymmetries into account to pre-emptively adjusting their foraging strategy in order to reduce the risk of conflict. This would point to the evolution of a ‘foraging ESS’ which is, however, currently undocumented because of the logistical difficulty of continuously tracking fine-scale foraging effort.

Meerkats are generalist, opportunistic carnivores that mainly feed on insect larvae and adult insects, with rarer contribution from small reptiles and scorpions (Doolan & MacDonald 1996b). The depth of insects below the surface of the sand depends on season: in the wet summer, large beetle larvae (Coleoptera) are abundant and active near the surface, whereas in the dry winter, these descend deeper into the soil, leaving only small items such as ants and small beetles on the soil surface (Doolan & MacDonald 1996b). Energetically rewarding items such as reptiles (e.g. *Nucrus tessellata*, *Ptenopus garrulous*, *Chondrodactylus angulifer*) and scorpions (of genera *Parabuthus* and *Opisthophthalmus*, Thornton & McAuliffe 2006) are almost always found relatively deep under the surface (Doolan & MacDonald 1996b). Thus, valuable prey items are rarely found without digging deep (Doolan & MacDonald 1996b), whereas low-value prey items are typically found without much effort on or just below the surface (Doolan & MacDonald 1996b). It is likely that prey-value stratification according to required foraging effort on the one hand, and reproductive benefits of avoiding conflict on the other, would select for differences in foraging strategies across asymmetric individuals. Knowing what these ‘evolutionarily stable’ foraging strategies for meerkats are and which asymmetries cause them, however, requires characterisation of allocation of time and energy to the three different foraging modes: searching and scratching, one-armed digging, and two-armed digging.

Miniature accelerometers affixed to the body have found widespread utility in remotely and non-invasively recognising behaviour (Martiskainen et al. 2009; Nathan et al. 2012; Chakravarty et al. 2019) and quantifying energy expenditure (*EE*) (Montoye et al. 1983 in humans; Wilson et al. 2006; Gleiss, Wilson & Shepard 2011; Qasem et al. 2012) at fine temporal resolution across >120 animal species (Brown et al. 2013). Two recent models were shown to recognise common coarse- and fine-scale meerkat behaviour with high accuracy and robustness. These behavioural categories were: (i) four ‘coarse-scale’ behaviours – resting, vigilance, foraging (a compound category comprising of searching, digging, and prey ingestion), and running (Chakravarty et al. 2019), and (ii) four ‘fine-scale’ behaviours – searching, one-armed digging, two-armed digging, and prey chewing (Chakravarty et al. 2020, *in review*). There are three main approaches to quantify *EE* from accelerometer data. In the first method, common in animal studies (Wilson et al. 2006; Gleiss, Wilson &

Shepard 2011), nongravitational acceleration (termed ‘dynamic body acceleration’ in these studies), obtained after subtracting the static component (due to the effect of gravity) from measured acceleration, is used as a proxy for rate of *EE*. The two others, common in human studies, consist of activity counts (sum of rectified acceleration over a fixed epoch) (Chen & Bassett Jr. 2005), and converting the type of activity derived from acceleration profile into metabolic equivalent of task (MET) (Bonomi et al. 2009) based on a published compendium (Ainsworth et al. 2000). The advantage of the last method is the possibility to have a tangible and realistic value of *EE*; however, it requires knowing the duration of each daily task.

The main objective of this study was to evaluate whether asymmetries in individual traits are linked to differences in foraging strategies in free-ranging, wild meerkats. Inspired from an existing approach in human *EE* estimation, and fusing the accelerometer-based coarse- and fine-scale behaviour recognition (CoFiBRec) algorithms into a single model, time-energy budgets for meerkats were constructed for the most common behaviours they engage in: resting, vigilance, searching, one-armed digging, two-armed digging, prey chewing, and running. Foraging efficiency was quantified as duration of chewing per unit of energy invested in foraging effort. Different individual traits such as body mass, age, and sex are considered to investigate whether asymmetries between individuals give rise to differences in foraging strategy (proportions, durations, and energy expended in searching and digging), foraging efficiency, and cooperative activity (vigilance duration) between and within sexes.

6.2 | Material and Methods

6.2.1 | Data collection and preparation

Fieldwork was conducted at the Kalahari Meerkat Project (Clutton-Brock & Manser 2016) from April 2018 to August 2019. Radio collars (Figure 6.1) containing recording units developed by CDD Ltd (Athens, Greece) were deployed on 26 adult individuals (aged 12 months or more) with the goal of recording 7 days/individual. Unit size was 37 mm x 22 mm x 22 mm, and collars weighed <25 grams; collars of

this size and weight have been shown not to affect meerkat behaviour (Golabek, Jordan & Clutton-Brock 2008). Units contained a triaxial accelerometer recording at 50 Hz/axis with a range of ± 8 g (± 78.48 m/s²). Battery life was 4-8 days, and the accelerometer's internal clock was updated every four hours through the collar GPS. Individuals for which recording duration was <1 day were removed from the dataset since I was interested in full-day behaviour. For the other individuals, only those days of recording were retained for which the device recorded data for $\geq 90\%$ of both the morning (sunrise to solar noon) and afternoon (solar noon to sunset) hours. Accelerometer calibration was done according to Ferraris, Grimaldi & Parvis 1995, and the accelerometer axes arranged to conform with the axis arrangement specified in Chakravarty et al. 2019 and Chakravarty et al. 2020 (*in review*): surge (positive direction: tail-to-head), sway (positive direction: leftward), and heave (positive direction: ventral-to-dorsal). Body mass was obtained by training individuals to stand on an electronic platform balance (Clutton-Brock et al. 1999b). Individuals were weighed 'Weight gain' was measured by subtracting morning weight (when individuals emerged from their sleeping burrow) from evening weight (measured last thing before entering the burrow in the evening).

6.2.2 | Coarse- and fine-scale behaviour recognition model (CoFiBRec)

The coarse- (Chapter 3) and fine-scale (Chapter 5) behaviour recognition (CoFiBRec) algorithms were fused into a single model (Figure 6.2) that takes as input triaxial acceleration data and outputs seven behaviours that meerkats engage in for >96% of the time (Chakravarty et al. 2019): resting, vigilance, searching, one-armed digging, two-armed digging, prey chewing, and running. Full details on model parameters are provided in Appendix S6.1.

6.2.3 | Deriving estimates of *EE*

Here, an estimate of *EE* is proposed based on MET-based coding of each behaviour obtained from Section 6.2.2. This approach, used in humans, provided a better correlation with actual *EE* than activity counts (Bonomi et al. 2009).



Figure 6.1. **Meerkats equipped with collars.** Three individuals exhibiting vigilance behaviour.

Nevertheless, this new approach is compared to the main existing method for *EE* estimation in animal studies using gravity compensation to obtain dynamic body acceleration (Qasem et al. 2012).

6.2.3.1 MET-based coding: literature-derived rates of *EE* (LiDREE)

In the Compendium of (human) Physical Activities (Ainsworth et al. 2000), standard MET values for different activities were compiled from previously published and unpublished data and literature, as multiples of resting metabolic rate corresponding to 1 MET (1 kcal/kg.h) when lying quietly. Resting metabolic rate (RMR) of meerkats has previously been measured to be 241 kJ/day (0.167 kJ/min) (Scantlebury, Clutton-Brock & Speakman 2004). Using a similar strategy as

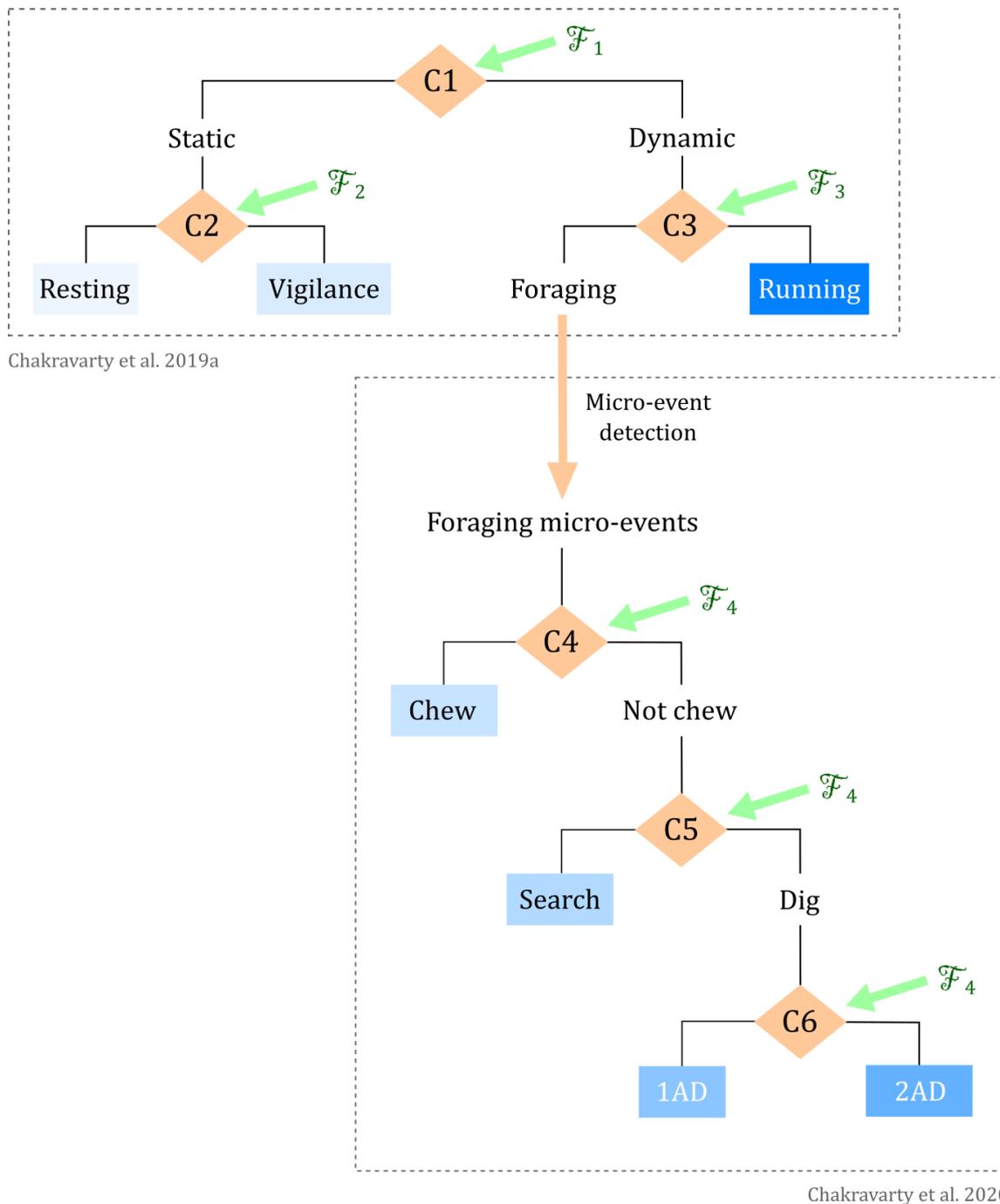


Figure 6.2. **The coarse- and fine-scale behaviour recognition (CoFiBRec) model.** Seven behaviours are classified using two recently developed models. The first model (Chakravarty et al. 2019) splits behaviours into two static behaviours (resting, vigilance) and two dynamic (foraging, running) behaviours. The second model (Chakravarty et al. 2020, *in review*) resolves the compound category ‘foraging’ into four finer behaviours: searching, one-armed digging (1AD), two-armed digging (2AD), and chewing. The model uses linear decision boundaries at each node (C1-C6), and a total of 11 features grouped into four sets (\mathcal{F}_1 - \mathcal{F}_4).

Ainsworth et al. 2000, I defined MET for each behaviour as multiples of known meerkat RMR. For this, I leveraged data and models from past studies of animal *EE*, and assigned the following METs to the behaviours of interest: resting, 1; vigilance, 1.2; chewing, 2; searching, 3; one-armed digging, 4; two-armed digging, 5; and running: 16.

EE during resting was taken to be equal to RMR, resulting in a MET of 1. The MET value of 1.2 for vigilance was taken from the value for the ‘alert’ state in thirteen-lined ground squirrels (*Ictidomys tridecemlineatus*) (Table 4 in Karasov 1992) since, like the squirrels, meerkats have similar upright posture, and are stationary and alert during vigilance. The value of 2 for chewing was chosen to be intermediate between 1.2 for relatively slow chewers such as humans (Levine, Baukol & Pavlidis 1999) and 2.9 for relatively rapid chewers such as short-tailed shrews (*Blarina*) (Martinsen 1969). The value of 4 for 1AD was taken from studies on diggers such as badgers (*Taxidea taxus*) (Lampe 1976) and pocket gophers (*Thomomys talpoides*) (Andersen & MacMahon 1981). I assigned a value of 5 for 2AD following the visual observation that in 2AD, compared to 1AD, a larger quantity of sand is flung for a greater distance (~2 body lengths compared to ~half body length for 1AD), and higher peak acceleration is observed compared to 1AD (Chakravarty et al. 2020, *in review*).

To estimate MET for searching and running, I derived a relationship between moving speed $V_{km/hr}$ and rate of *EE* $E_{kJ/min}$ (equation 6.1) for meerkats from the multi-species data and model presented in Taylor, Schmidt-Nielsen & Raab 1970 (details in Appendix S6.2). Considering average speed to be 1-2 km/hr during searching, equation 6.1 yields rates of expenditure of 0.44-0.64 kJ/min. Considering average running speed to be 12 km/hr (which is 37.5% of the maximum reported running speed of meerkats of 32 km/hr in <https://www.natgeokids.com/au/discover/animals/general-animals/meerkat-facts/>), equation 6.1 yields a rate of *EE* of 2.64 kJ/min. Dividing by the meerkat RMR of 0.167 kJ/min (Scantlebury, Clutton-Brock & Speakman 2004), I obtained MET values of 2.6-3.8 and 15.8 for searching and running, which I approximated to 3 and 16, respectively.

$$E_{kJ/min} = 0.2V_{km/hr} + 0.24 \quad (6.1)$$

6.2.3.2 | Vectorial dynamic body acceleration (VeDBA)

VeDBA is commonly used as a proxy for energy expended by logger-bearing animals (Qasem et al. 2012). It is computed by first obtaining dynamic body acceleration (denoted by X_{dyn} , Y_{dyn} , Z_{dyn} for dynamic acceleration recorded along the three axes of the accelerometer) by subtracting static acceleration from accelerometer data, and then computing the vectorial norm (equation 6.2) (Qasem et al. 2012). Static acceleration is obtained by smoothing acceleration recorded along each axis using a running mean over two seconds (Shepard et al. 2008).

$$VeDBA = \sqrt{X_{dyn}^2 + Y_{dyn}^2 + Z_{dyn}^2} \quad (6.2)$$

VeDBA assumes that the gravity component of acceleration is compensated by removing static acceleration, which is not always true, particularly in the presence of rotational movements where change in inclination and kinematic acceleration can occur at the same rate (Van Hees et al. 2013). Nevertheless the method is used widely in animal *EE* studies, though where it was found recently that VeDBA correlates with *EE* only when computed separately for each activity (Jeanniard-du-Dot et al. 2017). Therefore, I computed VeDBA for each behaviour separately (more details in Appendix S6.3) and compared it to LiDREE.

6.2.4 | Descriptive metrics and statistical analysis

Three classes of metrics were computed to characterise daytime behaviour: (i) behaviour durations, (ii) total *EE*, and (iii) foraging efficiency.

Durations were computed for each of the seven fine-scale behaviours plus those of compound behaviours such as foraging (searching, 1AD, 2AD, and chewing), foraging effort (searching, 1AD, and 2AD), and digging (1AD and 2AD). Total *EE* was computed using equation 6.3:

$$EE = \sum_{i=1}^7 c_i D_i \quad (6.3)$$

where D_i denotes the duration of behaviour i (in the list of behaviours in Section 2.2) having units of minutes, and c_i the coefficient of rate of EE for behaviour i (having units of kJ/min in the case of LiDREE, where each MET value has been multiplied by RMR, and m/s^2 in the case of VeDBA). The total EE computed with LiDREE coefficients is henceforth denoted by EE_{LiD} , and that using VeDBA coefficients by EE_{VeDBA} . Taking duration of chewing as a proxy for foraging success (though noting that chewing duration may not necessarily have a linear relationship with energy gain or mass gain), foraging efficiency FE was defined as the duration of chewing per energy expended in foraging effort, and computed according to equation 6.4:

$$FE = \frac{D_{chew}}{c_{Search}D_{search} + c_{1AD}D_{1AD} + c_{2AD}D_{2AD}} \quad (6.4)$$

Each metric was computed for three time slots: (i) sunrise to solar noon (henceforth referred to as ‘morning’), (ii) solar noon to sunset (‘afternoon’), and (iii) sunrise to sunset (‘full day’ or ‘daytime hours’).

Correlations between morning and afternoon metrics were computed to examine whether morning behaviour influenced afternoon behaviour. Correlations between full-day metrics, and age and body mass were performed to examine within-sex differences. `corrcoef` (MATLAB R2019b) was used to compute the Pearson correlation coefficient r_p and p-value P for the null hypothesis that the two variables were uncorrelated. The Wilcoxon rank sum test (WRST), computed using `ranksum` (MATLAB R2019b), was used to examine between-sex differences by comparing metrics computed from data recorded on females versus males. The null hypothesis was that the two sets of metrics were sampled from continuous distributions with equal medians.

6.3 | Results

6.3.1 | Collected data

For nine of the 26 individuals on which accelerometers were deployed, hardware malfunction caused the device to stop recording before the first day of recording was over – these individuals were removed from the dataset. The device identity for one individual (for which 1 day of data was collected) was lost – this individual was also excluded from the dataset. Of the remaining 16 individuals, three did not have days of recording where the unit recorded data for $\geq 90\%$ of both the morning and afternoon hours. The final dataset (Table 6.1) consists of 47 full days of data from 13 individuals (5 females and 5 males in winter, and 3 females in summer), corresponding to >90 million triaxial acceleration samples (counting only the daylight hours; night-time recording is not included in this figure). Animal body mass (mean across days of recording, Table 6.1) ranged from 513 grams for #8 to 763 grams for #6, while age ranged from 12 months (#4, #8) to 48 months (#10). There had been rainfall in the 14 days preceding each recording done in the summer (#5, #6, #7), while no rainfall was recorded in the 14 days prior to any of the winter recordings. Individuals #5 and #6 were the dominant females in their respective groups; #5 was lactating, while #6 was pregnant.

6.3.2 | Time budgets

Visualisation of triaxial acceleration signals and CoFiBRec-classified behaviour (Figure 6.3) yielded results that agreed with prior experience with groundtruthed acceleration data recorded on meerkats (Chakravarty et al. 2019; Chakravarty et al. 2020, *in review*). Duration distributions for coarse-scale behaviours obtained with CoFiBRec (Table S6.6) agreed with those observed in a previous groundtruthing study with video-annotated behaviours (Table S6.5). There is currently no reference for the comparison of duration distributions of fine-scale behaviours predicted by CoFiBRec since the groundtruthing study conducted previously (Chakravarty et al. 2020, *in review*) used a balanced dataset to minimise model bias due to imbalanced data. Nevertheless, the higher proportions of searching and 1AD, and lower proportions of 2AD and chewing (Figure 6.4) agree with

S.No.	Individual name	Sex	Age (months)	Weight (grams)	# Days of recording	Month and year
1.	VVHF109 ₂₅	F	25	581 ± 14	6	Aug-Sep '18
2.	VVHF110	F	25	537 ± 6	5	Aug-Sep '18
3.	VLF230	F	22	627 ± 9	6	Sep '18
4.	VHMF010	F	12	577 ± 10	6	Sep '18
5.	VVHF109 ₃₀	F	30	590	1	Feb '19
6.	VRUF001	F	13	763	1	Feb '19
7.	VLF241	F	13	602	1	Mar '19
8.	VLF246	F	12	513 ± 2	4	Jul '19
9.	VJXM101	M	30	na	6	Jul '19
10.	VLM221	M	48	733 ± 10	3	Jul '19
11.	VZUM033	M	16	660 ± 4	5	Aug '19
12.	VZUM034	M	16	600 ± 1	2	Aug '19
13.	VHMM008	M	23	717	1	Aug '19

Table 6.1. **Summary of collected data.** Accelerometers were deployed on a total of 26 individuals, of which 13 had to be removed because of hardware malfunction. The remaining 13 (8 females, 5 males) individuals yielded a total of 47 full days of recording – corresponding to > 90 million triaxial acceleration records – from both seasons: the cold, dry winter (May - September), and the hot, wet summer (October - April). Mean and standard deviation across days of average body mass per day are reported for individuals for which data are available for multiple days. Body mass was not available (na) for individual #9.

field observations.

Meerkats were active for most of the daytime hours: the mean proportion of dynamic behaviour (running and foraging) was 67.1% in winter and 51.8% in summer (Figure 6.4 a,c) (standard means across individuals; individual-wise time budgets are reported in Table S6.6). Most dynamic behaviour (98.9% in winter, 97.4% in summer) consisted of foraging. While the mean proportion of digging was similar for both seasons (22.6% in winter, 26.6% in summer), the mean proportion of searching in winter (36.3%) was almost twice that in summer (19.9%).

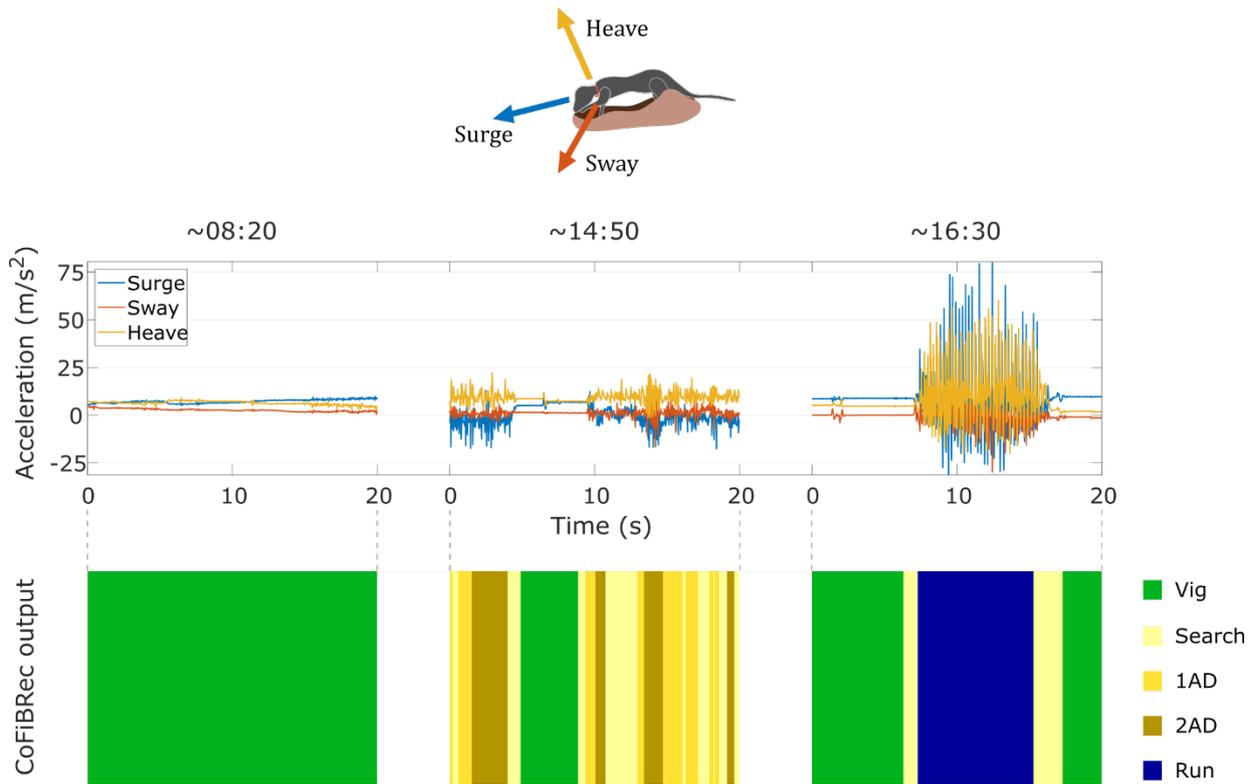


Figure 6.3. **Triaxial acceleration and recognised behaviour.** Twenty-second snapshots of triaxial acceleration data (upper panel) from individual #3 at different times of day are shown here to illustrate behaviour recognition (lower panel) using CoFiBRec. The individual presumably exhibited sunning in the morning (~08:20), where body posture is identical to that during bipedal/sitting vigilance. Busy foraging was observed in the afternoon (~14:50), with rapid transitions between different foraging modes. A period of high-intensity running between two vigilance bouts was observed in the late afternoon (~16:30). 1AD: one-armed digging; 2AD: two-armed digging.

Though the mean proportion of time spent being vigilant was similar for both seasons (26% in winter, 22.3% in summer), there was a marked increase (by 73%) in vigilance proportion on winter mornings (15.6%) compared to those during summer (9.0%). A nearly four-fold increase in the mean proportion of resting was observed in summer (25.8%) compared to winter (6.9%). Cumulative plots of resting duration versus time of day (Figure S6.2; cumulative plots of duration of each of the other six behaviours versus time of day are provided in Figures S6.3-S6.8) revealed that meerkats rested throughout the midday hours in summer whereas there was virtually no resting during this period in winter. The proportion of each behaviour in the afternoon was positively correlated with its proportion in the morning: resting $r_p = 0.63$, $P = 1.7e-6$; vigilance $r_p = 0.38$, $P = 0.008$; searching $r_p = 0.89$, $P = 5.3e-17$; one-armed digging $r_p = 0.70$, $P = 3.7e-8$;

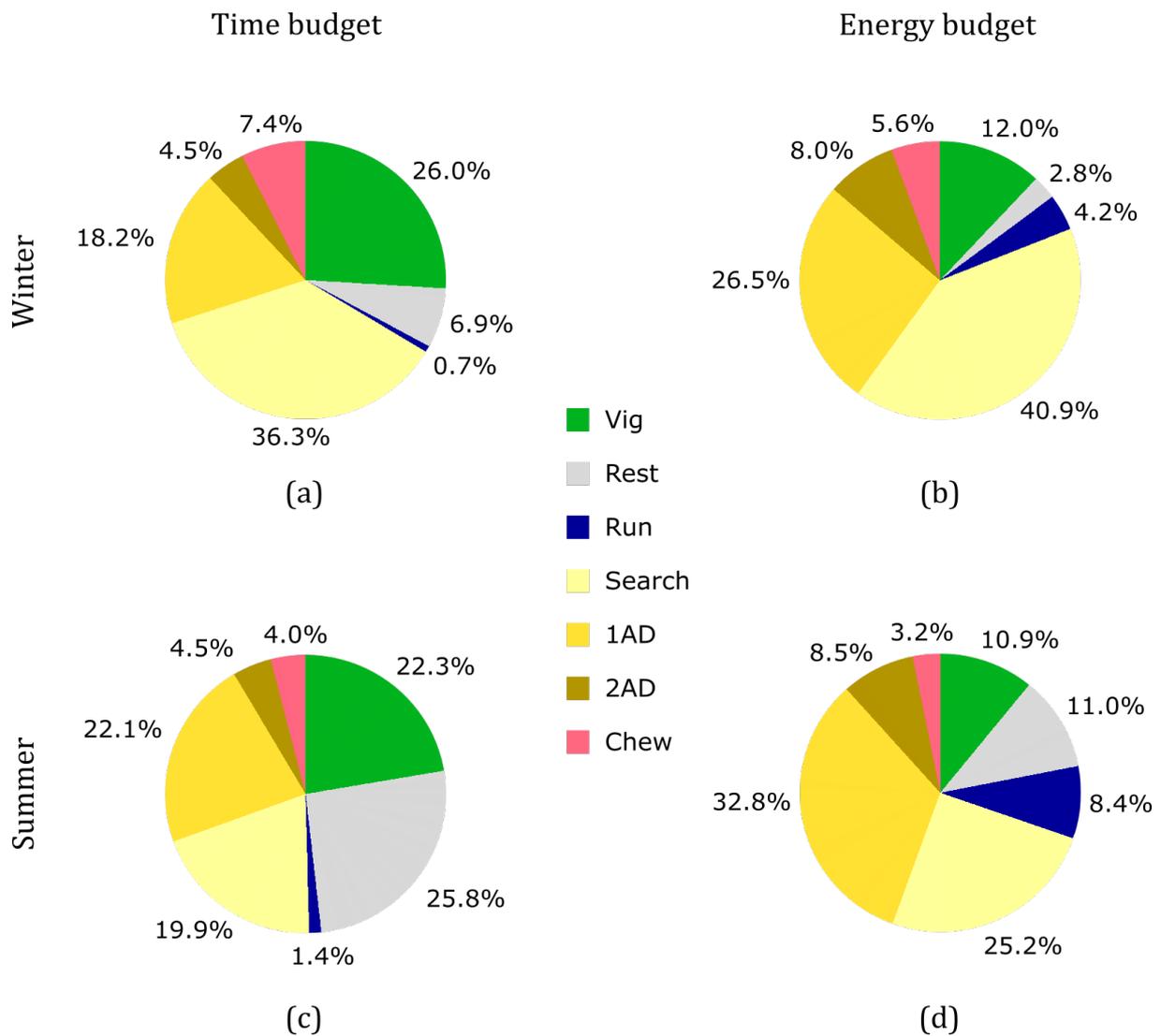


Figure 6.4. **Time and energy budgets by season.** From behavioural sequences obtained using CoFiBRec, proportions of duration (presented here as percentage) of each behaviour were computed for the daytime hours (sunrise to sunset). The proportions shown were averaged over data from 10 individuals (5 males, 5 females) totalling 44 days of recording in the winter, and 3 individuals (3 females) totalling 3 days of recording in the summer. Meerkats spent most of their time and energy on dynamic activity, of which the major constituent was foraging. About a quarter of the time was spent being vigilant. Resting duration increased dramatically in the summer in response to the scorching heat during the midday hours. *EE* was calculated using LiDREE intensity coding. 1AD: one-armed digging; 2AD: two-armed digging.

two-armed digging $r_p = 0.77$, $P = 1.8e-10$; chewing $r_p = 0.90$, $P = 1.5e-17$; and running $r_p = 0.53$, $P = 1.2e-4$.

6.3.3 | *EE* computation, and energy budgets

EE_{LiD} was computed for the entire day by assuming MET value of 1 for the hours before sunrise and after sunset, when meerkats retire to underground burrows to sleep. I obtained 24-hour *EE* values of 413 ± 31 kJ (standard mean and error, $N = 13$. Individual-wise values reported in Table S6.8). Considering only subordinate females, EE_{LiD} values were 404 ± 33 kJ ($N = 6$), which were not significantly different from values of 391 ± 93 kJ (WRST: $P = 0.50$) and 468 ± 89 kJ (WRST: $P = 0.36$) reported from measurements with doubly labelled water for allo-lactating and helper subordinate females, respectively (Scantlebury et al. 2002). VeDBA values (in m/s^2) for the seven behaviours were computed to be: resting, 0.54; vigilance, 1.47; chewing, 3.14; searching, 4.12; one-armed digging, 5.98; two-armed digging, 4.91; and running, 10.30 (standard deviation for each value and further details in Appendix S6.3). There was a very strong positive correlation between EE_{LiD} and EE_{VeDBA} of $r_p = 0.98$, $P = 3.4e-31$.

Due to the strong positive correlation between EE_{LiD} and EE_{VeDBA} , and realistic EE_{LiD} values, I present the remaining energy-expenditure results using LiDREE intensity coding only. Meerkats allocated more than three-fourths of their daytime (sunrise to sunset) energy budget to dynamic activities: 85.2% in winter, and 78.1% in summer (Figure 6.4 b,d) (standard means across individuals; individual-wise energy budgets are reported in Table S6.7).

6.3.4 | Behavioural differences

For this analysis, metric values computed for each day were considered as separate data points since both behaviour patterns and body mass can vary from day to day for the same individual.

6.3.4.1 | Within-sex differences

Females

I first controlled for season by removing the three individuals (#5, #6, #7, Table 6.1) that were recorded in summer. Data were available for a total of 27 full days of recording from 5 females. A strong negative correlation between body mass

and foraging efficiency ($r_p = -0.79$, $P = 7.2e-7$) was found. This occurred because of two reasons. First, search duration was significantly negatively correlated with body mass ($r_p = -0.74$, $P = 9.2e-6$) (Figure 6.5 a) and digging duration was significantly positively correlated with body mass ($r_p = 0.77$, $P = 2.1e-6$) (Figure 6.5 b). Both 1AD ($r_p = 0.76$, $P = 3.7e-6$) and 2AD ($r_p = 0.80$, $P = 5.3e-7$) had strong, significant positive correlations with body mass. Since MET value for searching was smaller than those for both 1AD and 2AD (Section 2.3), this had the effect of reducing foraging effort for lighter females relative to their heavier counterparts (denominator, equation 6.4). Second, chewing duration was significantly negatively correlated with body mass ($r_p = -0.73$, $P = 1.5e-5$) (Figure 6.5 c). A combination of both these factors led to the decrease in foraging efficiency with increasing body mass (Figure 6.5 d). As a consequence, two distinct, body mass-dependent foraging modes emerged (Figure 6.5 a,b): lighter (<550 grams) females invested effort almost exclusively in searching (312 ± 19 minutes across days) and little in digging (41 ± 10 mins), while heavier females (>550 grams) invested more in digging (200 ± 46 mins) than searching (182 ± 39 mins). Coarse-scale behaviour recognition (Figure 6.2, upper panel) alone would not have been able to resolve the observed differences in foraging strategy: the duration of the compound category ‘foraging’ (including searching, 1AD, 2AD, and chewing) had no significant correlation with body mass ($r_p = -0.05$, $P = 0.79$).

Age did not have significant correlation with foraging efficiency ($r_p = -0.27$, $P = 0.18$), chewing duration ($r_p = -0.36$, $P = 0.07$), search duration ($r_p = 0.01$, $P = 0.96$), or digging duration ($r_p = -0.05$, $P = 0.82$). While there was no significant correlation between age and EE_{LiD} ($r_p = -0.10$, $P = 0.61$) (Figure 6.6 a), body mass had a positive, significant correlation with EE_{LiD} ($r_p = 0.41$, $P = 0.03$) (Figure 6.6 c). Vigilance duration had a significant positive correlation with age ($r_p = 0.5$, $P = 0.008$), but no significant correlation with body mass ($r_p = -0.20$, $P = 0.32$). Weight gain, available for all 5 females for a total of 20 of 27 days, was 16 ± 12 grams across days, with weight being lost on only two days (2 grams by #2 and 5 grams #3). Weight gain was not significantly correlated with either foraging efficiency ($r_p = -0.25$, $P = 0.33$) or chewing duration ($r_p = -0.19$, $P = 0.43$).

Males

Data were available for a total of 17 full days of recording from 5 males recorded in winter; body mass data were not available for one male (#9, Table 6.1). In males, chewing duration was significantly negatively correlated with age ($r_p = -0.65$, $P = 0.005$), which led to a significant negative correlation of foraging efficiency with age ($r_p = -0.63$, $P = 0.007$). No correlation was found between weight gain (available for 7 of 17 days from 3 males) and foraging efficiency ($r_p = 0.05$, $P = 0.91$) or chewing duration ($r_p = 0.05$, $P = 0.91$). There were no significant correlations of durations of searching, digging or vigilance, or *EE* with either age or body mass.

6.3.4.2 | Between-sex differences

Compared to males, females contributed significantly more to vigilance in terms of total duration during the day (WRST: $P = 0.025$; 180 ± 53 mins for females and 144 ± 51 mins for males). Females also spent significantly more energy compared to males (WRST: $P = 0.01$; EE_{LiD} was 292 ± 37 kJ for females and 271 ± 19 kJ for males).

6.3.4.3 | Dominants versus subordinates

Dominant females #5 (lactating) and #6 (pregnant) (recorded during the summer), spent 342 kJ and 357 kJ of energy during the daytime hours, respectively, which was more than any of the winter females (292 ± 37 kJ), winter males (271 ± 19 kJ), or the sole subordinate summer female (#7, Table 6.1) (231 kJ) (means and standard deviations across days of recording). The dominant females also performed running, the most energetically costly behaviour (Section 2.3.1), for longer durations (12.3 mins for #5; 13.7 mins for #6) compared to any of the winter females (5.5 ± 2.6 mins), winter males (3.5 ± 1.5 mins), or the sole subordinate summer female (4.5 mins) (Figure S6.8). Additionally, the duration of 1AD for #6 (pregnant) (250 mins) was higher compared to all other meerkats (116.6 ± 58 mins).

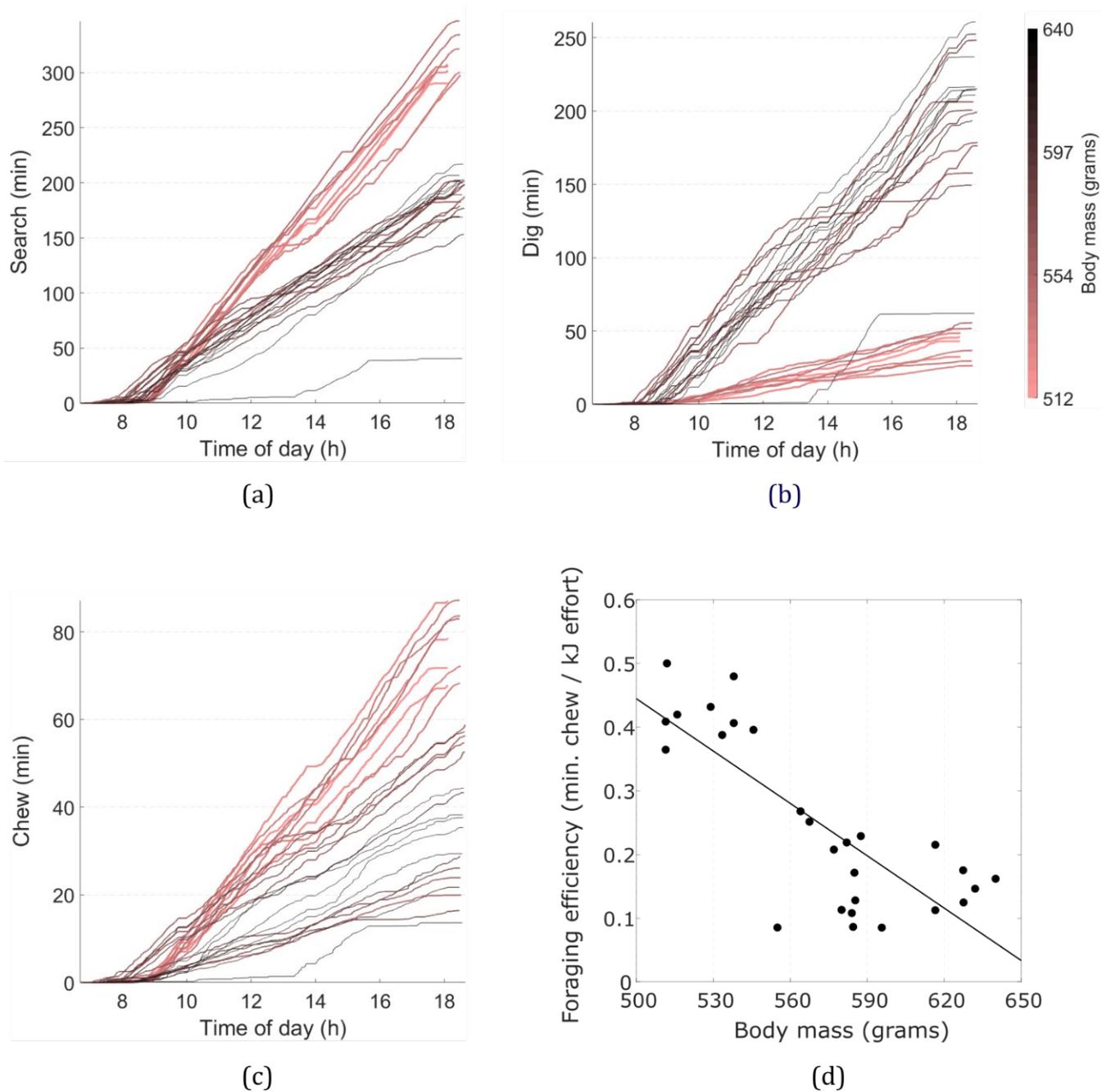


Figure 6.5. Body mass-dependent foraging strategies in females. Two distinct, body mass-dependent foraging modes were found in subordinate females recorded in winter: lighter females invested almost exclusively in searching (a) and little in digging (b), while heavier females invested more in digging than searching. In addition, lighter females chewed for longer (c). This led to a significant decrease in foraging efficiency (d), defined as minutes of chewing per kJ or foraging effort (searching and digging), with body mass ($r_p = -0.79$, $P = 7.2e-7$). No ‘light’ males (<550 grams) were recorded, so one cannot conclude whether this is a strategy specific to females or not.

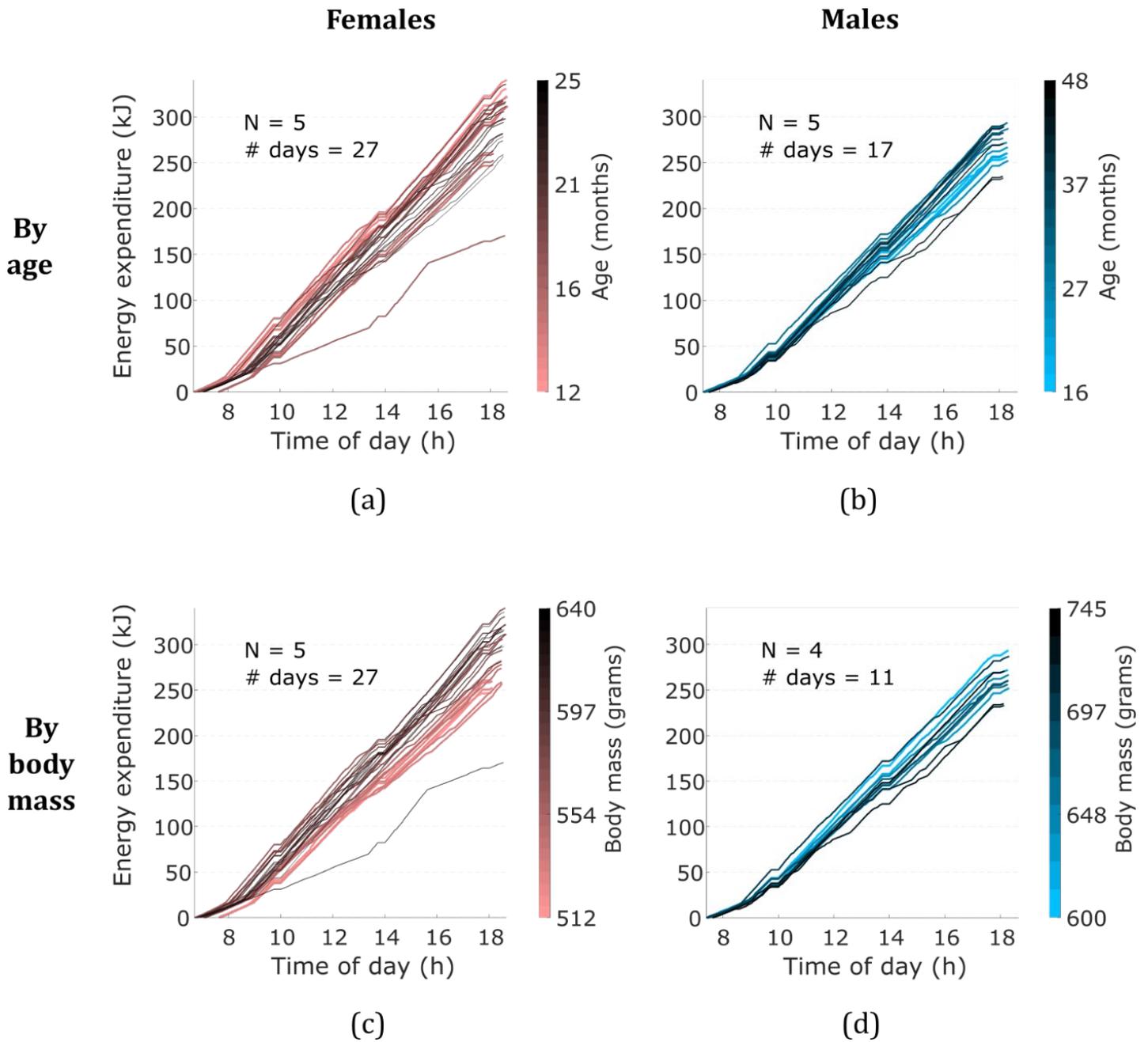


Figure 6.6. **The evolution of energy expenditure with time of day.** Females were found to spend significantly more energy compared to males (WRST: $P = 0.01$). There was no significant correlation between age and energy expenditure for either males or females. However, female energy expenditure increased with body mass ($r_p = 0.41$, $P = 0.03$). *EE* was calculated using MET-based coding (LiDREE).

6.4 | Discussion

I constructed behavioural time-energy budgets for wild meerkats from triaxial acceleration data by employing validated, coarse- and fine-scale behaviour recognition (CoFiBRec), and *EE* estimation techniques. The obtained time budgets were consistent with findings from previous studies based on manual observation. For the first time, *EE* estimation using behaviour-wise MET, usually applied in human studies, was applied in an animal study, and yielded *EE* values in kJ that were in good agreement with field measurements made in a previous study measured using doubly labelled water. Further, there was strong agreement between *EE* estimates obtained using both MET-based coding and vectorial dynamic body acceleration. Fine-scale information on different modes of foraging effort – low-intensity searching versus intensive digging – revealed that asymmetries in body mass in females led to contrasting foraging strategies. Lighter females invested virtually exclusively in searching, whereas heavier females invested more in digging than in searching. Coarse-scale behaviour recognition would not have been able to resolve these differences in foraging strategy. Older females contributed more to vigilance than younger females, and females contributed to vigilance (in terms of duration) more than males. Finally, both dominant females recorded (one pregnant, one lactating) ran more and expended more energy during the daytime hours compared to all other individuals.

CoFiBRec yielded realistic time budgets that were consistent with observations of meerkat behaviour in past studies. First, dynamic behaviour was recognised for most part of the day, and a majority of dynamic behaviour was classified as foraging, which is consistent with day-long manual observations that meerkats are active throughout the day, and that most activity is devoted to foraging (Doolan & MacDonald 1996b). Second, long bouts of resting were detected during the midday hours in summer but not in winter; this was consistent with previous observations that meerkats retreat into burrows for a ‘siesta’ period in summer to avoid the extremely high midday temperatures, and that siestas cease as temperatures drop in the winter (Doolan & MacDonald 1996b). Third, in each winter recording, but none of the summer recordings, CoFiBRec detected a long period of vigilance right after sunrise – this may have been due to sunning behaviour that meerkats perform on cold winter mornings after emerging from their sleeping burrow, where

they are static, upright, and facing the Sun (Demartsev et al. 2018). This shows one limitation of accelerometer-based behaviour classification in that, even though the behaviour is correctly classified based on the biomechanics involved (posture in this case), the actual interpretation depends on knowledge of *a priori* context to which the accelerometer is 'blind'. The nearly two-fold increase in the mean searching budget from summer (19.9%) to winter (36.3%) may be a response to decreased prey availability in winter owing to low rainfall (English, Bateman & Clutton-Brock 2012), forcing meerkats to search for longer as compared to the summer, when the arrival of the rainy season makes prey abundant again (Doolan & MacDonald 1996b). Finally, I found that, in both seasons, behaviour proportions in the morning and afternoon did not differ significantly.

Four observations indicate that lighter females targeted low-value surface prey while heavier females targeted both high-value below-ground prey and low-value surface prey: (i) lighter females nearly exclusively searched, while heavier females dug more than they searched, (ii) the duration of chewing was higher for lighter females as compared to heavier females, (iii) virtually all females (18 of 20 days of recording) gained weight over the course of the day, but weight gain was not correlated with chewing duration, and (iv) heavier females spent more energy due to greater investment in digging but nevertheless gained weight at the end of the day. The nearly exclusively search-based strategy employed by lighter females could only have given them access to low-value prey items near the surface of the soil, such as ants and small beetles (Doolan & MacDonald 1996b). On the other hand, the dig-and-search-based strategy employed by heavier females would have given them access to both higher-value prey such as reptiles and scorpions located deeper below the surface of the soil (Doolan & MacDonald 1996b) as well as small surface prey. Heavier females' increased *EE* due to significantly more digging must have been offset by greater energy intake from larger prey. The higher chewing duration for lighter females suggests that they made up for the low value of prey items through volume, which would also explain why they would have to search significantly more than heavier females, since prey items are dispersed over a large area (Doolan & MacDonald 1996b).

Differences in targeted prey while cohabiting the same patch may indicate the presence of an 'evolutionarily stable foraging strategy', wherein lighter females

‘settle’ for lower-value prey items near the soil surface that neither require extensive digging to find (lighter females will presumably have lower energy reserves, too) nor would represent costly losses in the case of successful scrounging attempts made by heavier females. Heavier females, with presumably greater energy requirements, would accrue minimal benefit by engaging in conflict with lighter females for small prey items. Instead, they leverage their presumably higher energy reserves to invest in digging for higher-value prey while also keeping a lookout for low-value prey through searching. This strategy is thus different from the foraging niche specialisation that has previously been observed in social carnivores, where individuals living with access to the same resources occupy distinctive foraging niches (Robertson et al. 2014). Since all males recorded in our study were ‘heavy’ (>550 grams), one cannot conclude whether this body mass-strategy is seen exclusively in females, or both males and females.

The duration of neither total foraging nor vigilance significantly correlated with body mass, suggesting that even though female meerkats face high selection for body mass (Hodge et al. 2008), they balance total foraging budget with contribution to cooperative behaviour (vigilance). Instead, within the time available for foraging, they modulate their relative investments in digging and searching. Asymmetries in age, known to affect the probability of acquiring breeding positions in females (Hodge et al. 2008), led to differences in contribution to cooperative activity (older subordinate adult females were vigilant for longer compared to younger females) but not in foraging strategy.

A secondary outcome of this study was insight into cooperative behaviour. I found that female meerkats displayed vigilance for significantly longer durations compared to males. This is consistent with the expectation that since female meerkats are more likely to breed in their natal group than males, their philopatry would manifest as increased cooperative behaviour (Clutton-Brock et al. 2002). In the cited study, vigilance (or raised guarding), also a cooperative behaviour, had been found to be performed more frequently (i.e. a greater number of times) by males than by females. The increased frequency of vigilance in males in Clutton-Brock et al. 2002 was explained by the large direct fitness benefits males likely to disperse would gain through information about the presence of females in neighbouring groups. The increased duration of vigilance in females compared to males found in our study

suggests that the cooperative aspect of vigilance may lie in its duration rather than frequency.

Even though data from several days are available in this study, the sample size in terms of number of individuals is low (e.g. five females in the foraging strategy analysis), and represents a limitation of this study. Further, by including data from two years (2018, 2019), some random effects may have been introduced. The effects of age and body mass were assumed to be additive in nature (i.e. acting separately) in the female foraging strategy study, but one cannot exclude the possibility of a multiplicative effect of age and body mass from our results. Finally, hardware malfunction was a major limiting factor in this study. The nominal battery life was 7 days, and deployments were carried out on 26 individuals, so in theory one should have obtained 182 days of data; however, we obtained only 47 days of data, which is roughly a quarter of the total data that could have been collected. Nevertheless, the relationships obtained (e.g. foraging strategy versus body mass in females in the winter) were strong, allowing some insight into foraging and vigilance behaviour in meerkats despite the limited data.

In conclusion, the results in this study demonstrate that accelerometer-based analysis of fine-resolution behaviour can capture behavioural differences arising from environmental (season), individual (body mass, age, sex), social (dominance status), and physiological (pregnancy, lactation) factors.

Appendices

Appendix S6.1 | Feature sets and decision boundaries in CoFiBRec

The hierarchical behaviour recognition scheme (Figure 6.2, Chapter 6) consists of a total of 6 nodes (classifiers C1–C6; one classifier per node) employing 11 features grouped into four feature sets (\mathcal{F}_1 – \mathcal{F}_4). Classifiers C1 to C3 belong to the first method for coarse-scale behaviour recognition (Chakravarty et al. 2019a). Here, acceleration data are first split into two-second windows (with an overlap of 50% between successive windows), and three features are computed for each window. C1 classifies the window as belonging to activity that is either static or dynamic. C2 separates static activity into resting and vigilance. C3 separates dynamic activity into foraging and running. This model was shown to have an overall accuracy of >95% and high robustness to inter-individual variation.

Nodes C4 to C6 belong to the second method, where the compound category ‘foraging’ is resolved into four fine-scale foraging behaviours (Chakravarty et al. 2020, *in review*). Here, surge acceleration from each continuous portion of the triaxial acceleration signal classified as ‘foraging’ by the first method is first enveloped and segmented to locate behavioural microevents (i.e. impulsive movements made during foraging that produce brief shock signals in accelerometer data). Then, eight features (\mathcal{F}_4) are computed for each detected microevent (Appendix S6.1). C4 separates chewing microevents from the other categories. C5 separates digging microevents (1AD and 2AD) from searching. Finally, C6 separates 2AD from 1AD. This model was shown to have an overall accuracy of >85% and high robustness to inter-individual variation and changes in model parameters.

Note that coarse-scale behaviour recognition (CoBRec) outputs one behaviour label for each second of acceleration data, whereas fine-scale behaviour recognition requires one behaviour label for each sample of triaxial acceleration. To fuse the two models, CoBRec’s stream of window-wise behaviour labels B_w was converted to sample-wise behaviour labels B_s such that B_s had the same number of entries as the number of triaxial acceleration samples in the day-long signal. Consider $B_{w,i}$, the window-wise behaviour label at time t_i seconds. $B_{w,i-1}$ corresponds to time $t_i - 1$ seconds, and $B_{w,i+1}$ corresponds to time $t_i + 1$ seconds. To all samples in B_s that correspond to a time value

between $t_i - 0.5$ seconds and $t_i + 0.5$ seconds, I assigned the behaviour label $B_{w,i}$. This converted CoBRec's output from second-wise to sample-wise behaviour labels, and fused the two models together. In summary, CoFiBrec (Figure 2 in main manuscript) takes as input day-long triaxial acceleration data, and outputs a sequence of behavioural codes B_s containing 7 unique codes – one for each behaviour of interest. B_s has the same number of entries as the number of triaxial acceleration samples, and thereby the same time resolution of 0.02 seconds (since accelerometer sampling frequency is 50 Hz/axis).

Parameter values of the coarse- and fine-scale behaviour recognition model (CoFiBRec) used in this study (Section 2.2) are provided here. CoFiBRec uses six class boundaries (Table S6.1) computed from a total of eleven features (Table S6.2) grouped into four sets. The following conventions for triaxial acceleration data were followed: surge (positive direction: tail-to-head), sway (positive direction: leftward), and heave (positive direction: ventral-to-dorsal) (Figure 3 in main manuscript). Acceleration sampling frequency was 50 Hz/axis.

Class separation	Class boundary	Feature set and model coefficients		
Static vs Dynamic	C1 If $\vec{A}_1 \cdot \vec{X}_1 \geq 0$, then Dynamic, otherwise Static.	\mathcal{F}_1	\vec{X}_1	\vec{A}_1
			meanSurge	-0.04998
			stdNorm	1
			bias	-0.10292
Resting vs Vigilance	C2 If Static AND $\vec{A}_2 \cdot \vec{X}_2 \geq 0$, then Vigilance, otherwise Resting.	\mathcal{F}_2	\vec{X}_2	\vec{A}_2
			meanSurge	1
			bias	-0.23140
Foraging vs Running	C3 If Dynamic AND $\vec{A}_3 \cdot \vec{X}_3 \geq 0$, then Running, otherwise Foraging.	\mathcal{F}_3	\vec{X}_3	\vec{A}_3
			stdNorm	1
			fftPeakPowerAvg	0.02976
			bias	-1.19868

Chew vs Not chew	C4	If Foraging AND $\vec{A}_4 \cdot \vec{X}_4 \geq 0$, then Not chew, otherwise Chew.	\mathcal{F}_4	\vec{X}_4	\vec{A}_4
				cxMinMovMeanSurge	-0.76383
				cxStdSway	4.08750
				cxFFTpeakpowerHeave	0.21407
				cxFFTdomfreqNorm	-0.11742
				cxFFTpowerfracHeave	5.01153
				cxCorrSurgeHeave	-4.62578
				cxNumAscDescHeave	-0.27854
				spMinSurge	-2.18355
	bias	9.12378			
Search vs Dig	C5	If Foraging AND Not chew AND $\vec{A}_5 \cdot \vec{X}_4 \geq 0$, then Dig, otherwise Search.	\mathcal{F}_4	\vec{X}_4	\vec{A}_5
				cxMinMovMeanSurge	-4.28971
				cxStdSway	-0.24563
				cxFFTpeakpowerHeave	0.15869
				cxFFTdomfreqNorm	0.11295
				cxFFTpowerfracHeave	-1.73690
				cxCorrSurgeHeave	-1.42750
				cxNumAscDescHeave	0.17014
				spMinSurge	-2.08527
	bias	-10.26811			
1AD vs 2AD	C6	If Foraging AND Not chew AND Dig AND $\vec{A}_6 \cdot \vec{X}_4 \geq 0$, then 2AD, otherwise 1AD.	\mathcal{F}_4	\vec{X}_4	\vec{A}_6
				cxMinMovMeanSurge	0.80317
				cxStdSway	-7.82587
				cxFFTpeakpowerHeave	0.10451
				cxFFTdomfreqNorm	-0.71621
				cxFFTpowerfracHeave	2.38598
				cxCorrSurgeHeave	-1.54460
				cxNumAscDescHeave	-0.00845
				spMinSurge	-0.64160
	bias	3.20601			

Table S6.1. **CoFiBRec parameters.** Class boundaries C1-C6, and feature sets \mathcal{F}_1 - \mathcal{F}_4 and their respective coefficients used in our coarse- and fine-scale behaviour recognition model (CoFiBRec) for meerkats.

S.No.	Feature name	Computation	Reference
1.	meanSurge	Mean of surge acceleration.	
2.	stdNorm	Standard deviation of vectorial norm of triaxial acceleration.	Chakravarty et al. 2019
3.	fftPeakPowerAvg	Maximum of axis-averaged (term-wise sum across axes divided by number of axes) squared coefficients of Fourier transforms of acceleration from each of the three axes.	
4.	cxMinMovMeanSurge	Minimum of moving mean of surge acceleration in context of microevent.	
5.	cxStdSway	Standard deviation of sway acceleration in context of microevent.	Chakravarty et al. 2020
6.	cxFFTpeakpowerHeave	Maximum squared coefficient of Fourier transform of heave acceleration in context of microevent.	
7.	cxFFTdomfreqNorm	Frequency at which maximum squared coefficient of Fourier transform of norm of contextual triaxial acceleration is obtained.	
8.	cxFFTpowerfracHeave	Power fraction in range 2.5 to 5 Hz for heave acceleration in context of microevent.	
9.	cxCorrSurgeHeave	Coefficient of correlation between surge and heave acceleration in context of microevent.	
10.	cxNumAscDescHeave	Number of ascents and descents (signal paths from a negative peak to the following positive peak, and vice versa) in heave acceleration in context of microevent.	
11.	spMinSurge	Minimum of surge acceleration in span of microevent.	

Table S6.2. **Feature computation.** Features 1-3 were computed in a moving window of length 2 seconds with 50% overlap between successive windows. Features 4-11 were computed in a window of size 1.5 seconds centred at the location of each detected microevent.

Appendix S6.2 | Derivation of expression for energetic cost of running versus speed for meerkats

I used Taylor, Schmidt-Nielsen & Raab 1970's data on running versus mass in a range of mammals (Table S6.3) to estimate energy expenditure in meerkats during searching and running (Section 2.3.1).

Mammal	Mass (grams)	Slope, m	Intercept, c
White mouse (<i>Mus musculus</i>)	21	2.83	3.34
Kangaroo rat (<i>Dipodomys merriami</i>)	41	2.01	1.75
Kangaroo rat (<i>Dipodomys spectabilis</i>)	100	1.13	1.81
Ground squirrel (<i>Citellus tereticaudus</i>)	236	0.66	1.27
White rat (<i>Rattus norvegicus</i>)	384	1.09	1.48
Dog (<i>Canis familiaris</i>)	2600	0.34	0.75
Dog (<i>Canis familiaris</i>)	18000	0.17	0.65

Table S6.3. **Energetic cost of running for mammals of different mass.** Data are referenced from Taylor, Schmidt-Nielsen & Raab 1970 for the expression: $M = mV + c$, where M is the oxygen consumption in ml O₂ · g⁻¹·hr⁻¹, V is running velocity in km.hr⁻¹, m is the slope of the straight line, and c the line's intercept.

Observing that a straight-line fit had been obtained by Taylor, Schmidt-Nielsen & Raab 1970 between energy expenditure M and the logarithm of body mass (Figure 4 in their paper), I reasoned that there would also be a linear relationship between the logarithm of body mass with both slope m and intercept c . These lines were obtained using MATLAB's `fitlm` function (version R2019b) as:

$$m = -0.8071 \times \log_{10}(\text{mass}_{\text{grams}}) + 3.2007 \quad (\text{S6.1})$$

with an R^2 of 0.77, and

$$c = -0.73831 \times \log_{10}(x) + 3.4309 \quad (\text{S6.2})$$

with an R^2 of 0.72.

The average body mass of meerkats in our dataset was 603 ± 67 grams (mean and standard deviation over # of days of recording). Approximating average body mass to 600 grams, and calculating m and c according to equations S1 and S2, respectively, I obtain the following equation for energy expenditure during meerkat running:

$$M_{meerkat} = 0.95845 \times V_{\text{km/hr}} + 1.1496 \quad (\text{S6.3})$$

Finally, the energy expenditure units (of $M_{meerkat}$) of $\text{ml O}_2 \cdot \text{g}^{-1} \cdot \text{hr}^{-1}$ can be converted to $\text{kJ} \cdot \text{hr}^{-1}$ by approximating 1 ml O_2 as 5 calories

(<https://www.ucl.ac.uk/~ucbcdab/enbalance/measuring.htm>), or 21 J, and the average mass of a meerkat to be 600 grams. Multiplying equation S3 by the multiplier of

$$\frac{21 \text{ J}}{\text{ml O}_2} \cdot 600 \text{ g} \cdot \frac{1 \text{ kJ}}{1000 \text{ J}} * \frac{1 \text{ hr}}{60 \text{ min}} = 0.21$$

results in equation 1 of the main manuscript.

Appendix S6.3 | Computation of VeDBA for the seven behaviours of interest, and comparison with MET-based coding

I first computed ‘VeDBA per triaxial acceleration sample’ as a measure of rate of energy expenditure for each behaviour from video-annotated bouts of resting, vigilance and running collected previously in Chakravarty et al. 2019, and searching, 1AD, 2AD, and chewing in Chakravarty et al. 2020 (*in review*). For a given behaviour, consider bout b_i (a bout is defined here as a continuous sequence of the behaviour under consideration uninterrupted by any other behaviour) of duration two seconds or more containing N_{b_i} triaxial acceleration samples. I first computed VeDBA according to equation 2 (in main manuscript, as defined in Qasem et al. 2012) for each of the N_{b_i} acceleration samples, and then computed their mean to obtain $VeDBA_{b_i}$. Then, I computed the mean of $VeDBA_{b_i}$ values across all available bouts of this behaviour, and denoted this as the VeDBA value corresponding to this behaviour. This was done for each of the seven behaviours (Table S6.4).

Behaviour	# bouts	Duration (s)	VeDBA (m/s ²)
Resting	70	56.2 ± 63.0	0.54 ± 0.98
Vigilance	1733	20.2 ± 52.6	1.47 ± 0.98
Chewing	78	7.5 ± 8.7	3.14 ± 0.96
Searching	93	3.9 ± 1.8	4.12 ± 1.28
1AD	94	4.5 ± 2.4	5.98 ± 1.18
2AD	64	5.2 ± 4.4	4.91 ± 0.95
Run	266	4.5 ± 2.4	10.30 ± 2.94

Table S6.4. **Behaviour-wise VeDBA.** VeDBA (in m/s²) was calculated from video-validated bouts of each of the seven behaviours of interest from a total of 10 meerkats (data was collected as reported in Chakravarty et al. 2019 & Chakravarty et al. 2020, *submitted, in review*). The standard mean and error of VeDBA across bouts is reported here for each behaviour. 1AD: one-armed digging; 2AD: two-armed digging.

Energy-expenditure estimates for the daytime hours (sunrise to sunset) based on VeDBA (EE_{VeDBA}) and MET-based coding (this study) (EE_{LiD}) were strongly linearly correlated (Section 3.3 in main manuscript). A straight line (Figure S6.1) fit using least-squares (`fitlm` function in MATLAB R2019b) (equation S4) yielded an R^2 value of 0.95:

$$EE_{LiD} = 0.117EE_{VeDBA} + 25 \quad (S6.4)$$

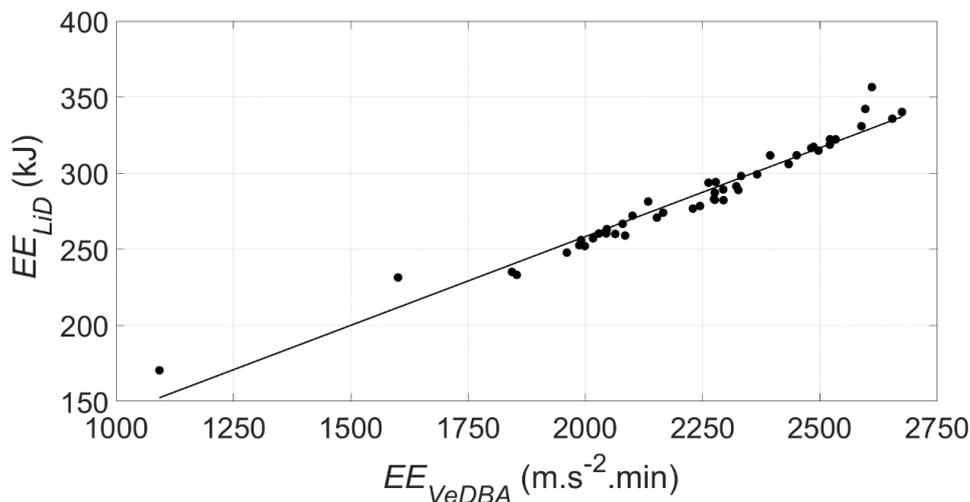


Figure S6.1. **Comparing VeDBA and MET-based coding.** Energy expenditure estimated based on VeDBA (Qasem et al. 2012) (EE_{VeDBA}) and MET-based coding (this study) (EE_{LiD}) over the daytime hours (sunrise to sunset) showed very strong agreement, with an R^2 of 0.95 for least-squares straight-line fit.

Appendix S6.4 | Comparison of CoFiBRec’s behaviour-proportion distributions with those recorded in a previous study

Good agreement was found between duration distributions recorded in a groundtruthing (i.e. with video annotation) study (Chakravarty et al. 2019) performed previously (Table S6.5) and those found using CoFiBRec in the current study (Table S6.6). Note that the coarse-scale behaviour ‘foraging’ (Table S6.5) is composed of four finer-scale modes: searching, one-armed digging, two-armed digging, and chewing (Table S6.6).

Individual #	Resting	Vigilance	Foraging	Running
1	25.4	55.1	18.7	0.8
2	1.3	41.6	56.8	0.3
3	0.0	18.7	80.8	0.5
4	0.0	63.6	34.4	2.0
5	0.0	23.1	74.9	2.1
6	10.8	23.7	64.1	1.4
7	1.8	56.4	41.3	0.5
8	0.0	1.4	98.2	0.4
9	0.0	70.9	28.4	0.7
10	0.0	29.8	68.9	1.3
11	8.3	22.5	68.2	1.0
Overall	4.3 ± 7.9	37.0 ± 21.9	57.7 ± 24.4	1.0 ± 0.6

Table S6.5. Behaviour proportions derived from data in Chakravarty et al. 2019. Proportion of time spent (expressed as percentage of sum of durations of all four coarse-scale behaviours) in each behaviour for each individual recorded in a previous study (Chakravarty et al. 2019). This table is derived from data presented in Chakravarty et al. 2019. In the previous study, individuals were continuously recorded and simultaneously filmed over a period of three hours during the morning; groundtruth data consisted of annotated videos from which behaviour durations were found.

Appendix S6.5 | Individual-wise time and energy budgets

Individual #	Resting	Vigilance	Searching	1AD	2AD	Chewing	Running
1	5.2 ± 1.7	35.3 ± 3.9	27.9 ± 2.3	22.7 ± 3.6	4.7 ± 0.9	3.3 ± 0.7	0.8 ± 0.2
2	3.7 ± 3.2	31.3 ± 6.2	47.1 ± 3.3	5.2 ± 1.8	0.6 ± 0.2	11.6 ± 1.2	0.5 ± 0.3
3	17.7 ± 14.7	20.5 ± 7.3	25.8 ± 9.9	24.7 ± 8.5	5.8 ± 2.3	5.0 ± 1.6	0.5 ± 0.2
4	6.0 ± 2.8	25.0 ± 3.3	27.9 ± 2.7	26.0 ± 4.6	5.8 ± 0.8	8.1 ± 0.8	1.3 ± 0.2
5	9.4	35.0	20.5	25.0	4.5	4.1	1.6
6	28.7	11.5	12.2	33.9	7.0	4.8	1.9
7	39.6	20.5	26.9	7.3	2.0	3.1	0.7
8	2.8 ± 1.0	23.7 ± 1.6	51.9 ± 1.3	6.3 ± 1.1	1.0 ± 0.1	13.1 ± 1.4	1.3 ± 0.3
9	5.1 ± 2.4	17.4 ± 3.5	40.0 ± 1.9	23.4 ± 1.0	5.7 ± 0.7	8.0 ± 1.0	0.5 ± 0.1
10	4.0 ± 3.1	35.4 ± 13.9	36.0 ± 6.1	16.2 ± 3.2	3.6 ± 0.9	4.4 ± 0.9	0.5 ± 0.2
11	7.0 ± 2.0	26.6 ± 3.3	42.4 ± 1.7	11.0 ± 1.0	2.6 ± 0.3	9.7 ± 0.9	0.8 ± 0.2
12	10.4 ± 0.3	24.4 ± 6.0	27.1 ± 1.7	24.4 ± 1.2	7.7 ± 2.5	5.3 ± 0.5	0.8 ± 0.1
13	7.6	20.1	37.5	21.8	7.0	5.8	0.2

Table S6.6. **Individual-wise time budgets.** Proportion of time spent (expressed as percentage of sum of durations of all seven behaviours) in each behaviour for each individual. 1AD: one-armed digging, 2AD: two-armed digging.

Individual #	Resting	Vigilance	Searching	1AD	2AD	Chewing	Running
1	2.0 ± 0.7	16.1 ± 2.5	31.5 ± 1.8	34.1 ± 4.3	8.9 ± 1.4	2.5 ± 0.4	4.9 ± 1.6
2	1.6 ± 1.4	15.9 ± 3.8	59.4 ± 1.5	8.8 ± 2.7	1.3 ± 0.3	9.8 ± 0.7	3.2 ± 1.9
3	8.5 ± 10.7	10.6 ± 7.9	27.9 ± 8.1	36.0 ± 7.7	10.4 ± 3.1	3.7 ± 0.8	2.9 ± 0.7
4	2.1 ± 1.0	10.4 ± 1.5	29.0 ± 3.5	35.8 ± 5.0	9.9 ± 1.0	5.6 ± 0.7	7.3 ± 1.1
5	3.5	15.6	22.8	37.1	8.3	3.1	9.7
6	9.9	4.8	12.7	47.0	12.1	3.3	10.3
7	19.7	12.3	40.2	14.6	5.0	3.1	5.2
8	1.1 ± 0.4	10.8 ± 1.0	59.1 ± 0.6	9.5 ± 1.5	1.9 ± 0.3	10.0 ± 1.2	7.6 ± 1.5
9	1.8 ± 0.8	7.2 ± 1.5	41.2 ± 2.1	32.1 ± 1.0	9.8 ± 1.1	5.5 ± 0.7	2.5 ± 0.6
10	1.5 ± 1.1	17.2 ± 7.8	42.4 ± 3.3	25.5 ± 2.6	7.1 ± 1.1	3.4 ± 0.4	2.9 ± 0.7
11	2.7 ± 0.7	12.5 ± 1.8	49.8 ± 1.9	17.1 ± 1.4	5.2 ± 0.6	7.6 ± 0.6	5.2 ± 1.0
12	3.7 ± 0.1	10.5 ± 3.1	29.1 ± 0.4	34.8 ± 0.1	13.6 ± 3.7	3.8 ± 0.1	4.5 ± 0.9
13	2.7	8.6	40.0	30.9	12.4	4.2	1.3

Table S6.7. **Individual-wise energy budgets.** Proportion of energy expended (expressed as percentage of the sum of energy expended across all seven behaviours) on each behaviour for each individual. Energy expenditure was estimated using MET-based coding. 1AD: one-armed digging, 2AD: two-armed digging.

Individual #	Individual name	Energy expenditure, EE_{LiD} (kJ)
1	VVHF109 ₂₅	428.9 ± 14.3
2	VVHF110	396.5 ± 11.7
3	VL230	414.6 ± 60.0
4	VHMF010	444.7 ± 11.6
5	VVHF109 ₃₀	448.2 ± 0.0
6	VRUF001	466.4 ± 0.0
7	VL241	349.3 ± 0.0
8	VL246	391.6 ± 7.0
9	VJXM101	420.2 ± 5.3
10	VLM221	381.6 ± 20.9
11	VZUM033	390.8 ± 5.5
12	VZUM034	415.0 ± 15.1
13	VHMM008	418.2 ± 0.0

Table S6.8. **Individual-wise estimates of 24-hour energy expenditure.** Assuming an intensity code of 1 for the hours before sunrise and after sunset, when meerkats are known to sleep in underground burrows, 24-hour energy expenditure (estimated using MET-based coding). Mean ± standard deviation across days is indicated for individuals for which more than one day of data was available.

Appendix S6.6 | Behaviour duration profiles over the course of the day by sex and season

This section presents the time-evolution of profiles of duration versus time of day for each of the seven behaviours of interest: resting (Figure S6.2), vigilance (Figure S6.3), searching (Figure S6.4), one-armed digging (Figure S6.5), two-armed digging (Figure S6.6), chewing (Figure S6.7), and running (Figure S6.8). Duration profiles are presented separately for males and females, and colour-coded according to season (winter: blue; summer: mustard).

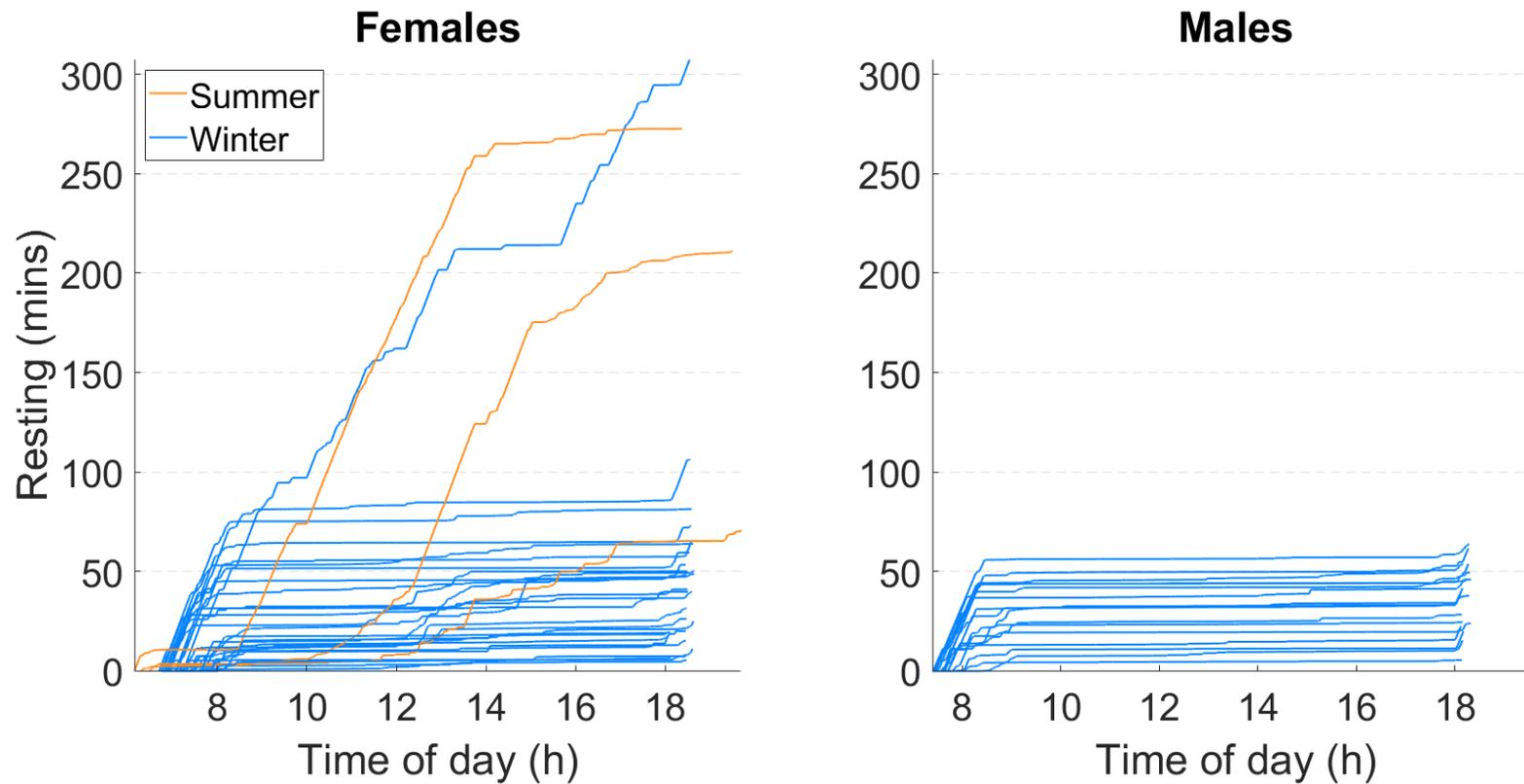


Figure S6.2. **Resting duration evolution with time of day.** Cumulative duration of resting versus time of day, by sex and season. Winter temperatures allowed meerkats to avoid resting during the midday hours, a luxury that was not available in the scorching summer. An exception was one female in the winter (#6, on 06-Sep-2018) who was reported by a field assistant to have stayed below ground due to bad weather for the most part of the day, thus leading to high resting duration. The other exception was #5, a lactating dominant female, who preferred to continue foraging despite the heat.

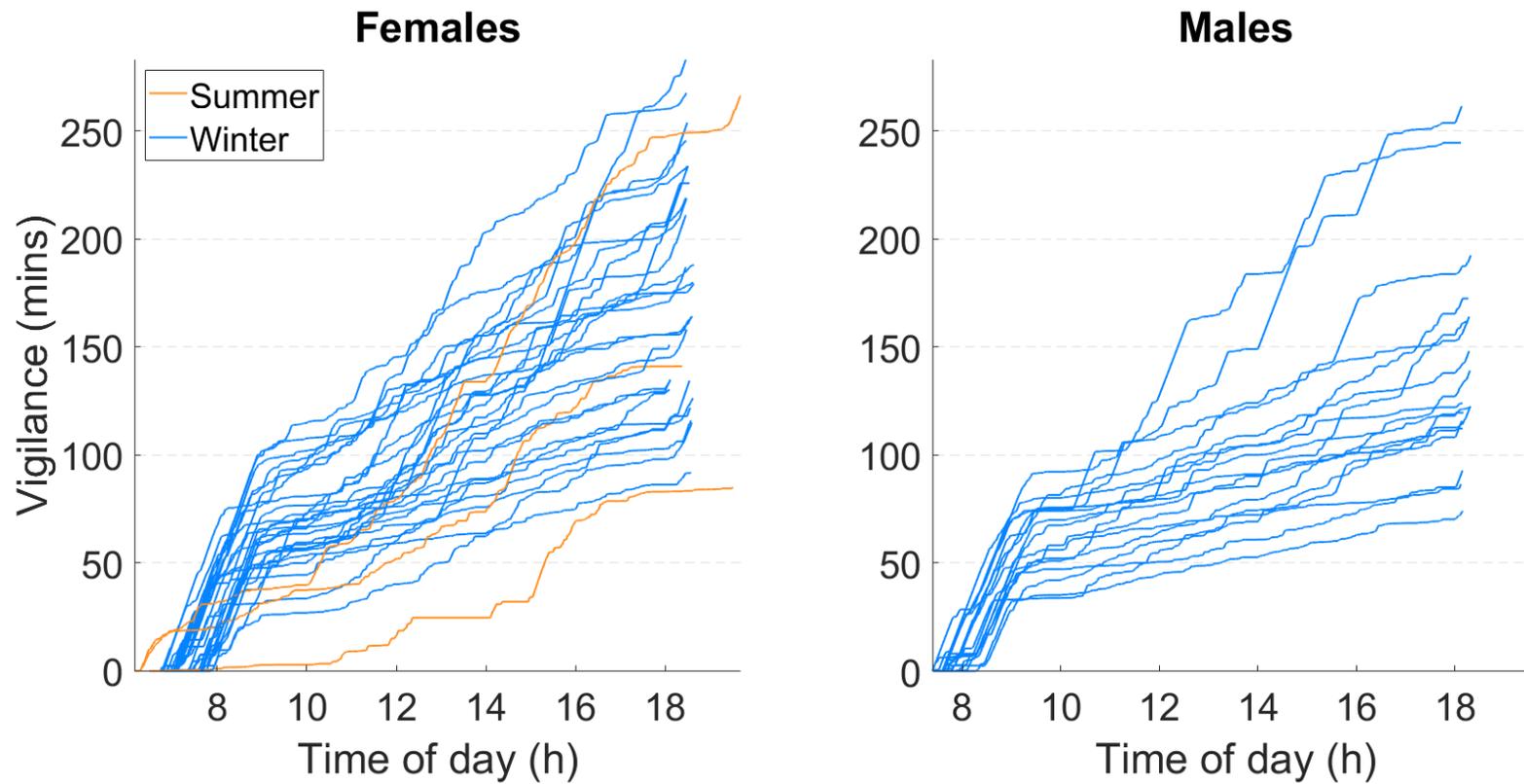


Figure S6.3. **Vigilance duration evolution with time of day.** Cumulative duration of vigilance versus time of day, by sex and season. Winter days started off with sunning behaviour, where meerkat posture is upright with belly facing the Sun, identical to bipedal vigilance – this is evidenced by a sharp, linear increase in vigilance duration directly after sunrise. Note that this behaviour is not displayed in the summer because of the high temperatures.

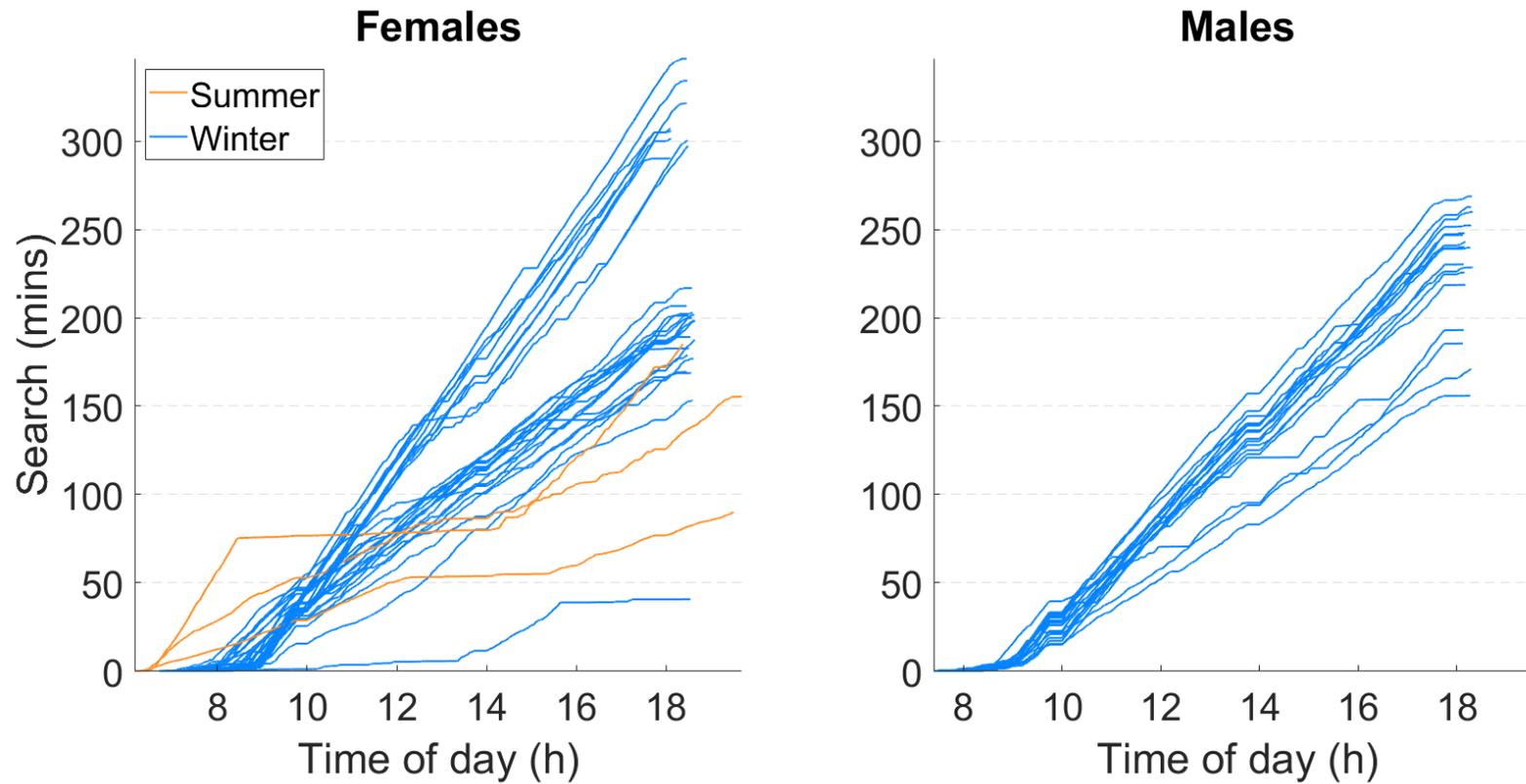


Figure S6.4. **Searching duration evolution with time of day.** Cumulative duration of searching versus time of day, by sex and season. Meerkats started searching right after sunrise in the summer, while the need for sunning behaviour on cold winter mornings delayed the start of the day's searching. Two distinct foraging strategies were observed: one involving nearly exclusive searching for lighter females, and the other involving more digging and less searching for heavier females (Figure 6.5).

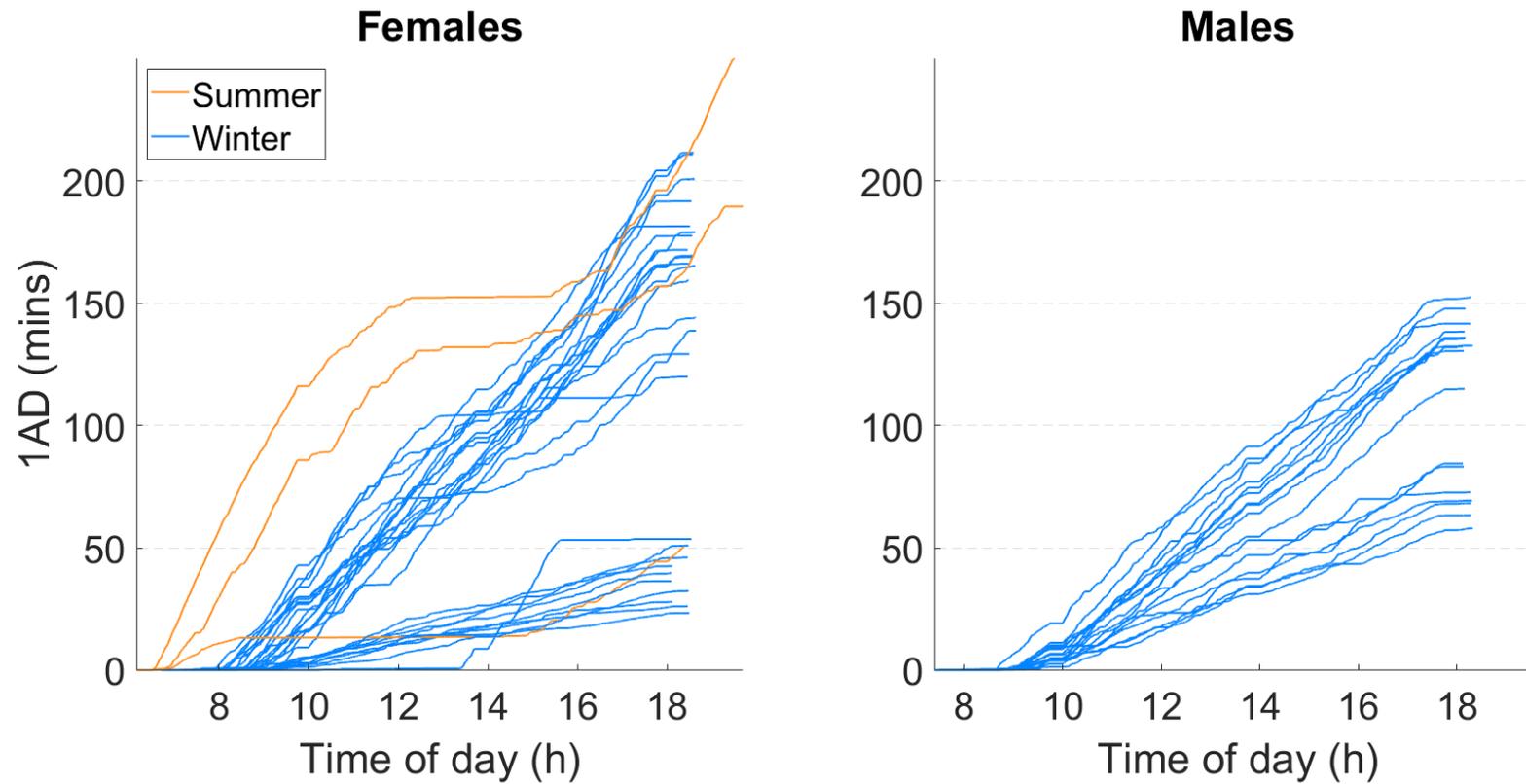


Figure S6.5. **One-armed digging duration evolution with time of day.** Cumulative duration of one-armed digging versus time of day, by sex and season. Meerkats started digging right after sunrise in the summer, while the need for sunning behaviour on cold winter mornings delayed the start of the day's digging. Two distinct foraging strategies were observed: one favouring digging in heavier females, and the other favouring searching in lighter females (Figure 6.5).

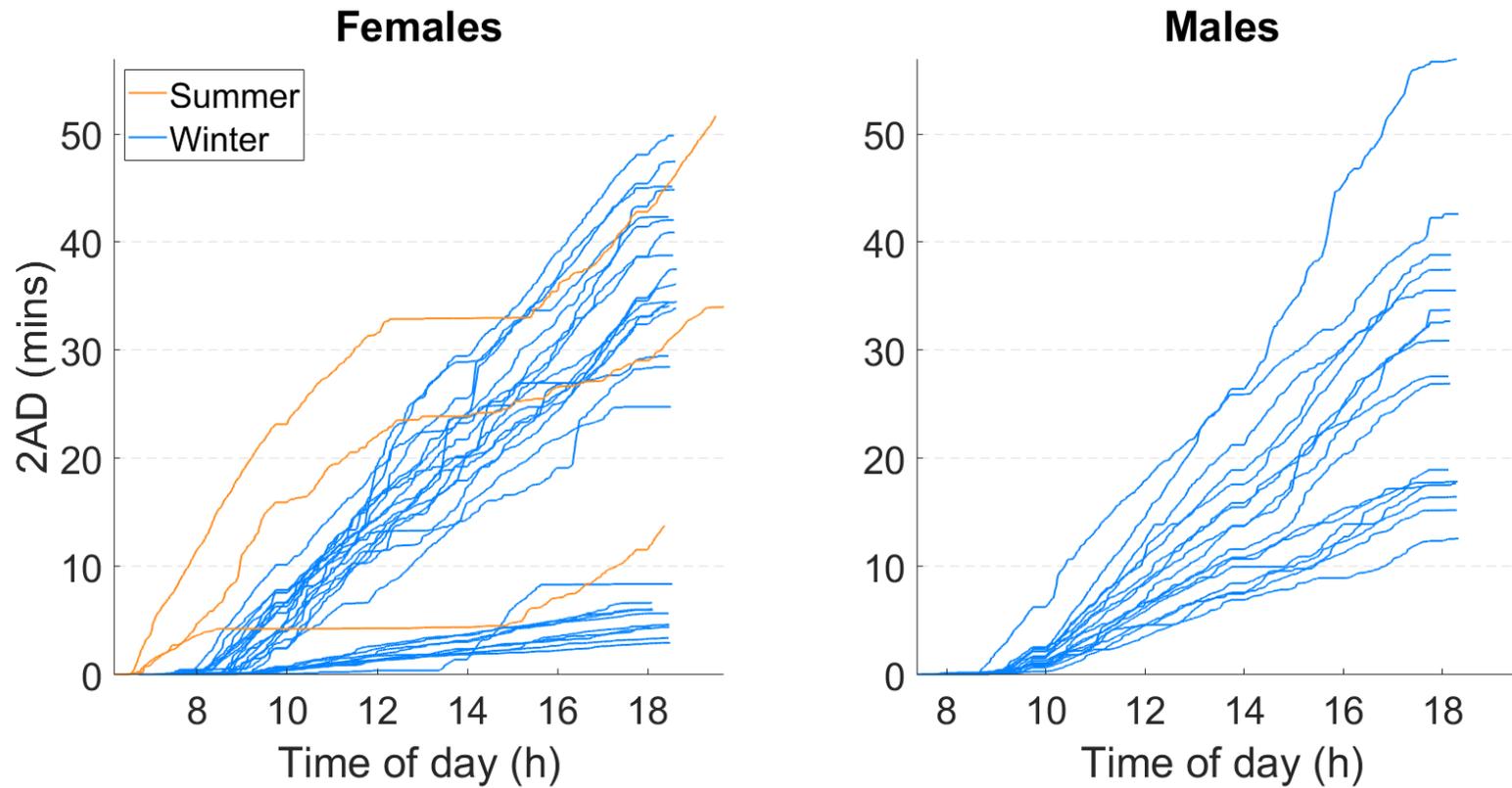


Figure S6.6. **Two-armed digging duration evolution with time of day.** Cumulative duration of two-armed digging versus time of day, by sex and season. The patterns observed here are similar to those observed for one-armed digging, except that the overall duration of two-armed digging is much lower compared to that of one-armed digging, presumably because of the higher energetic cost of two-armed digging (Section 6.2.3.1).

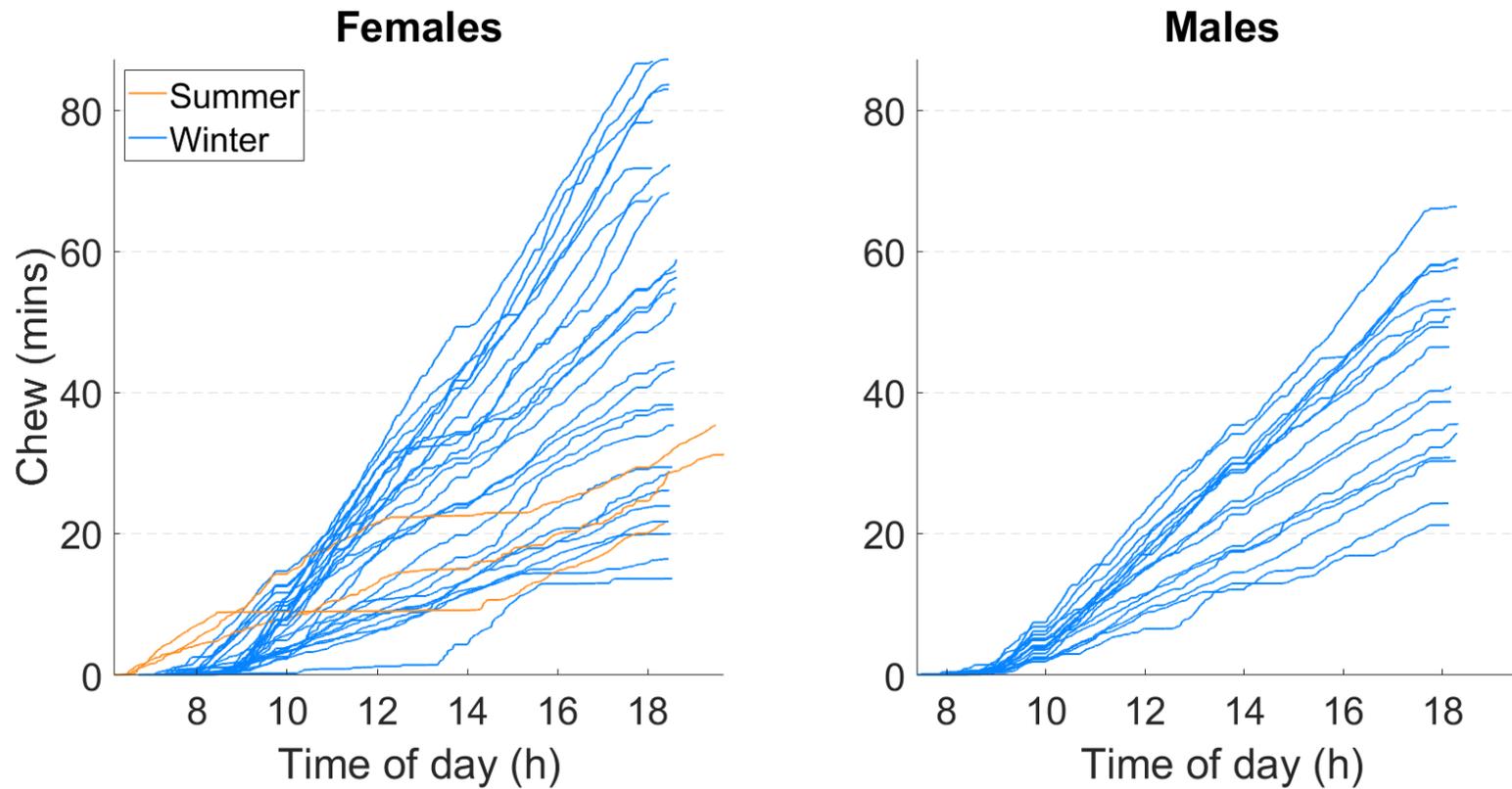


Figure S6.7. **Chewing duration evolution with time of day.** Cumulative duration of chewing versus time of day, by sex and season. Both males and females met with foraging success throughout the day in both seasons. Note, however, that chewing duration does not necessarily reflect energy intake, since different prey items will inevitably differ in their nutritive content.

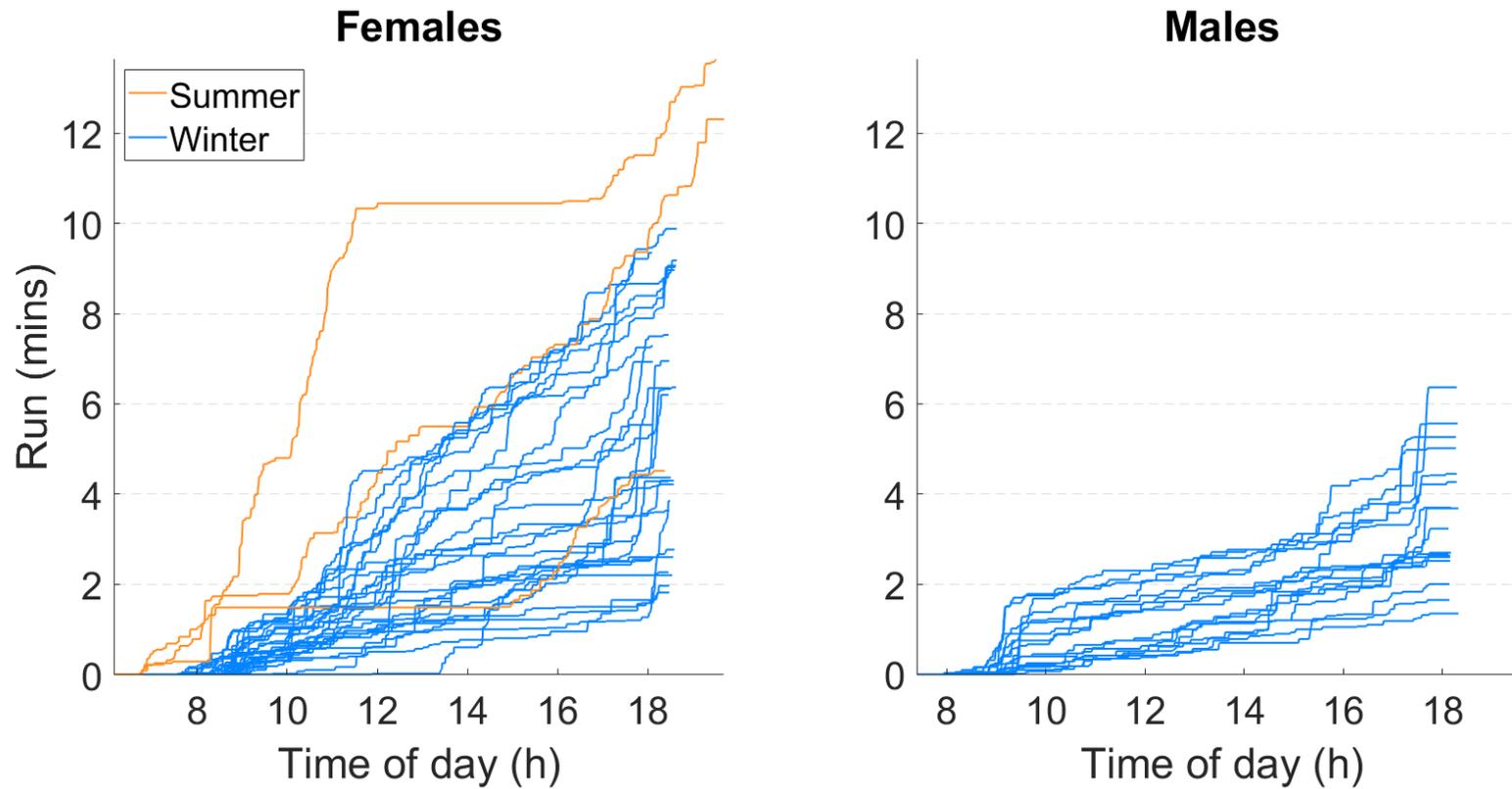


Figure S6.8. **Running duration evolution with time of day.** Cumulative duration of running versus time of day, by sex and season. Meerkats exhibited short running bouts (evidenced by short, near-vertical rises in the curves) throughout the day. Both dominants (#5, #6) ran more than all other meerkats. Owing to the high energy expenditure involved in an intensive activity such as running (Section 6.2.3.1), however, the total duration of running was much lower than any of the other behaviours.

References

- Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, A. M., Strath, S. J., ... & Jacobs, D. R. (2000). Compendium of physical activities: an update of activity codes and MET intensities. *Medicine & science in sports & exercise*, 32(9), S498-S516.
- Andersen, D. C., & MacMahon, J. A. (1981). Population dynamics and bioenergetics of a fossorial herbivore, *Thomomys talpoides* (Rodentia: Geomyidae), in a spruce-fir sere. *Ecological Monographs*, 51(2), 179-202.
- Bonomi, A. G., Plasqui, G., Goris, A. H., & Westerterp, K. R. (2009). Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. *Journal of applied physiology*, 107(3), 655-661.
- Brown, D. D., Kays, R., Wikelski, M., Wilson, R., & Klimley, A. P. (2013). Observing the unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry*, 1(1), 20.
- Chakravarty, P., Cozzi, G., Ozgul, A., & Aminian, K. (2019). A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods in Ecology and Evolution*, 10(6), 802-814.
- Chakravarty, P., Cozzi, G., Dejnabadi H., Léziart P-A., Manser M., Ozgul, A., & Aminian, K. (2020, *in review*). Seek and learn: automated identification of microevents in animal behaviour using envelopes of acceleration data and machine learning.
- Chen, K. Y., & Bassett Jr., D. R. (2005). The technology of accelerometry-based activity monitors: current and future. *Medicine & Science in Sports & Exercise*, 37(11), S490-S500.
- Clutton-Brock, T. H., Maccoll, A., Chadwick, P., Gaynor, D., Kinsky, R., & Skinner, J. D. (1999a). Reproduction and survival of suricates (*Suricata suricatta*) in the southern Kalahari. *African Journal of Ecology*, 37(1), 69-80.
- Clutton-Brock, T. H., Gaynor, D., McIlrath, G. M., Maccoll, A. D. C., Kinsky, R., Chadwick, P., ... & Brotherton, P. N. M. (1999b). Predation, group size and mortality in a cooperative mongoose, *Suricata suricatta*. *Journal of Animal Ecology*, 68(4), 672-683.

- Clutton-Brock, T. H., Russell, A. F., Sharpe, L. L., Young, A. J., Balmforth, Z., & McIlrath, G. M. (2002). Evolution and development of sex differences in cooperative behavior in meerkats. *Science*, *297*(5579), 253-256.
- Clutton-Brock, T. H., Russell, A. F., & Sharpe, L. L. (2003). Meerkat helpers do not specialize in particular activities. *Animal Behaviour*, *66*(3), 531-540.
- Clutton-Brock, T. H., Russell, A. F., & Sharpe, L. L. (2004). Behavioural tactics of breeders in cooperative meerkats. *Animal Behaviour*, *68*(5), 1029-1040.
- Demartsev, V., Strandburg-Peshkin, A., Ruffner, M., & Manser, M. (2018). Vocal turn-taking in meerkat group calling sessions. *Current Biology*, *28*(22), 3661-3666.
- Doolan, S. P., & MacDonald, D. W. (1996a). Dispersal and extra-territorial prospecting by slender-tailed meerkats (*Suricata suricatta*) in the south-western Kalahari. *Journal of Zoology*, *240*(1), 59-73.
- Doolan, S. P., & MacDonald, D. W. (1996b). Diet and foraging behaviour of group-living meerkats, *Suricata suricatta*, in the southern Kalahari. *Journal of Zoology*, *239*(4), 697-716.
- English, S., Bateman, A. W., & Clutton-Brock, T. H. (2012). Lifetime growth in wild meerkats: incorporating life history and environmental factors into a standard growth model. *Oecologia*, *169*(1), 143-153.
- Ferraris, F., Grimaldi, U., & Parvis, M. (1995). Procedure for effortless in-field calibration of three-axial rate gyro and accelerometers. *Sensors and Materials*, *7*(5), 311-330.
- Flower, T. P. (2011). *Competition for food in meerkats (Suricata suricatta)* (Doctoral dissertation, University of Pretoria).
- Gleiss, A. C., Wilson, R. P., & Shepard, E. L. (2011). Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure. *Methods in Ecology and Evolution*, *2*(1), 23-33.
- Golabek, K. A., Jordan, N. R., & Clutton-Brock, T. H. (2008). Radiocollars do not affect the survival or foraging behaviour of wild meerkats. *Journal of Zoology*, *274*(3), 248-253.
- Gosling, L. M. (1986). The evolution of mating strategies in male antelopes. *Ecological aspects of social evolution*, 244-281.
- Halsey, L. G., & White, C. R. (2010). Measuring energetics and behaviour using accelerometry in cane toads *Bufo marinus*. *PloS one*, *5*(4).

- Hodge, S. J., Manica, A., Flower, T. P., & Clutton-Brock, T. H. (2008). Determinants of reproductive success in dominant female meerkats. *Journal of Animal Ecology*, 77(1), 92-102.
- Jeanniard-du-Dot, T., Guinet, C., Arnould, J. P., Speakman, J. R., & Trites, A. W. (2017). Accelerometers can measure total and activity-specific energy expenditures in free-ranging marine mammals only if linked to time-activity budgets. *Functional Ecology*, 31(2), 377-386.
- Karasov, W. H. (1992). Daily energy expenditure and the cost of activity in mammals. *American Zoologist*, 32(2), 238-248.
- Lampe, R. P. (1976). *Aspects of the predatory strategy of the North American badger, Taxidea taxus*, PhD thesis, University of Minnesota, Minneapolis.
- Levine, J., Baukol, P., & Pavlidis, I. (1999). The energy expended in chewing gum. *New England Journal of Medicine*, 341(27), 2100-2100.
- Maag, N., Cozzi, G., Bateman, A., Heistermann, M., Ganswindt, A., Manser, M., ... & Ozgul, A. (2019). Cost of dispersal in a social mammal: body mass loss and increased stress. *Proceedings of the Royal Society B*, 286(1896), 20190033.
- MacDonald, D. W., & Doolan, S. P. (1997). Band structure and failures of reproductive suppression in a cooperatively breeding carnivore, the slender-tailed meerkat (*Suricata suricatta*). *Behaviour*, 134(11-12), 827-848.
- Martinsen, D. L. (1969). Energetics and activity patterns of short-tailed shrews (*Blarina*) on restricted diets. *Ecology*, 50(3), 505-510.
- Martiskainen, P., Järvinen, M., Skön, J. P., Tiirikainen, J., Kolehmainen, M., & Mononen, J. (2009). Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Applied animal behaviour science*, 119(1-2), 32-38.
- Molvar, E. M., & Bowyer, R. T. (1994). Costs and benefits of group living in a recently social ungulate: the Alaskan moose. *Journal of Mammalogy*, 75(3), 621-630.
- Montoye, H. J., Washburn, R., Servais, S., Ertl, A., Webster, J. G., & Nagle, F. J. (1983). Estimation of energy expenditure by a portable accelerometer. *Medicine and Science in Sports and Exercise*, 15(5), 403-407.

- 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.
- O’Riain, M. J., Bennett, N. C., Brotherton, P. N. M., McIlrath, G., & Clutton-Brock, T. H. (2000). Reproductive suppression and inbreeding avoidance in wild populations of co-operatively breeding meerkats (*Suricata suricatta*). *Behavioral Ecology and Sociobiology*, 48(6), 471-477.
- Ozgul, A., Bateman, A. W., English, S., Coulson, T., & Clutton-Brock, T. H. (2014). Linking body mass and group dynamics in an obligate cooperative breeder. *Journal of Animal Ecology*, 83(6), 1357-1366.
- Packer, C., & Pusey, A. (1985). Asymmetric contests in social mammals: respect, manipulation and age-specific aspects. *Evolution: essays in honour of John Maynard Smith*, 173-186.
- Qasem, L., Cardew, A., Wilson, A., Griffiths, I., Halsey, L. G., Shepard, E. L., ... & Wilson, R. (2012). Tri-axial dynamic acceleration as a proxy for animal energy expenditure; should we be summing values or calculating the vector?. *PloS one*, 7(2).
- Robertson, A., McDonald, R. A., Delahay, R. J., Kelly, S. D., & Bearhop, S. (2014). Individual foraging specialisation in a social mammal: the European badger (*Meles meles*). *Oecologia*, 176(2), 409-421.
- Scantlebury, M., Russell, A. F., McIlrath, G. M., Speakman, J. R., & Clutton-Brock, T. H. (2002). The energetics of lactation in cooperatively breeding meerkats *Suricata suricatta*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1505), 2147-2153.
- Scantlebury, M., Clutton-Brock, T. H., & Speakman, J. R. (2004, December). Energetics of cooperative breeding in meerkats *Suricata suricatta*. In *International Congress Series* (Vol. 1275, pp. 367-374). Elsevier.
- Sharp, S. P., & Clutton-Brock, T. H. (2011). Competition, breeding success and ageing rates in female meerkats. *Journal of evolutionary biology*, 24(8), 1756-1762.

- Shepard, E. L., Wilson, R. P., Liebsch, N., Quintana, F., Laich, A. G., & Lucke, K. (2008). Flexible paddle sheds new light on speed: a novel method for the remote measurement of swim speed in aquatic animals. *Endangered Species Research*, 4(1-2), 157-164.
- Smith, J. M., & Parker, G. A. (1976). The logic of asymmetric contests. *Animal behaviour*, 24(1), 159-175.
- Taylor, C. R., Schmidt-Nielsen, K., & Raab, J. L. (1970). Scaling of energetic cost of running to body size in mammals. *American Journal of Physiology-Legacy Content*, 219(4), 1104-1107.
- Thornton, A., & McAuliffe, K. (2006). Teaching in wild meerkats. *Science*, 313(5784), 227-229.
- Van Hees, V. T., Gorzelniak, L., Leon, E. C. D., Eder, M., Pias, M., Taherian, S., ... & Brage, S. (2013). Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. *PloS one*, 8(4).
- Ward, A., & Webster, M. (2016). Sociality: the behaviour of group-living animals.
- Wilson, R. P., White, C. R., Quintana, F., Halsey, L. G., Liebsch, N., Martin, G. R., & Butler, P. J. (2006). Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *Journal of Animal Ecology*, 75(5), 1081-1090.
- Young, A. J., Carlson, A. A., Monfort, S. L., Russell, A. F., Bennett, N. C., & Clutton-Brock, T. (2006). Stress and the suppression of subordinate reproduction in cooperatively breeding meerkats. *Proceedings of the National Academy of Sciences*, 103(32), 12005-12010.

CHAPTER 7

Conclusion

7.1 | Main contributions, general discussion, and limitations

The main aim of this thesis was to provide general, robust methods to quantify behaviour and energy expenditure (EE) in animals from inertial sensor data. Though the field of biologging has been witness to rapid hardware miniaturisation and breadth of tagging across taxa, methodology to remotely observe behaviour has nevertheless favoured ‘off-the-shelf’ machine learning which, despite its ready availability, is blind to the specifics of the nature of animal behaviour and inertial sensor data. This ‘black box’ approach has led to complex models that do not recognise common behaviours with high accuracy, the inability to deal with nonstationary signals recorded during fine-scale behaviour, and validation approaches that do not address the main goal of performing well on data from new individuals. This thesis addresses these issues by proposing robust and accurate biomechanically driven learning approaches, requisite signal processing techniques for inertial sensor signals, and appropriate cross-validation techniques for quantification of model performance. Secondly, existing accelerometer-based techniques for remote, non-invasive quantification of EE in animals have mainly been validated for locomotor behaviours, and make assumptions about the nature of recorded acceleration signals that may not be valid during complex or short-duration behaviours. This thesis proposes a new method for estimating EE from recognised behaviour that does not make such assumptions, is sensitive to impulsive, short-duration behaviours as well as static behaviours, and affords the generality of leveraging EE measurements and empirical relationships from past studies on the same or multiple surrogate species. These developments enable both behaviour and energy to be accurately quantified using a single, low-power sensor. Thirdly, I come full circle by applying this ‘engineering’ development to the ‘ecological’ question of how differences in individual traits, such as body mass, affect foraging strategy in free-ranging, wild, group-living meerkats.

7.1.1 | Biomechanically driven learning

The main contribution here is to have increased the robustness, accuracy, and generality of behaviour recognition schemes by making the classification process ‘biomechanically aware’. This was achieved through three new developments:

I. ***Grouping behavioural classes by biomechanical similarity.***

This offered a way to encode visually perceived characteristics of behaviour in terms of biomechanical similarities and differences. This was used to structure a hierarchical classification scheme that sequentially separated compound behavioural categories into finer behaviours, e.g. grouping vigilance and resting as ‘static’, foraging and running as ‘dynamic’, one- and two-armed digging as ‘digging’.

II. ***One good feature per biomechanical descriptor.***

This involved structuring the feature selection process so that behaviours are represented in ‘biomechanical space’ rather than ‘blind feature space’. Here, each feature corresponds to the most discriminative measure of a single biomechanical quantity (e.g. posture, intensity, periodicity, coordination, etc.), and the dimension of the ‘biomechanical space’ is equal to the number of biomechanical quantities used to represent behaviours. This reduced model complexity (in terms of number of features employed) while simultaneously solving the ‘blindness’ problem of black-box feature selection techniques, where several features representative of the same biomechanical quantity (e.g. intensity) may appear near the top of a black-box technique’s ranked list, leaving out key features that may be necessary to separate specific behaviours accurately. For instance, *cxFFTpowerfracHeave*, Chapter 5, an important feature that encodes the distinctive periodic motion in the 2.5 – 5 Hz frequency band during two-armed digging and not during other fine-scale foraging modes, was not selected by traditional feature selection techniques).

III. ***The concept of microevents as building blocks of complex behaviour.***

This offered a way to tackle the nonstationarity of acceleration signals recorded during complex behavioural classes. Signals in such cases are difficult to interpret even visually (in contrast to other more evident differences such as resting versus running). Formulating complex behaviour as being composed of brisk, individual impulse-

like movements (microevents) – e.g. a digging swipe, or a head jerk during chewing – allowed separation of shock-like sections of the signal from smoother sections of the signal that weren't characteristic of microevents. This led to the selection of features with greater discriminatory power, leading to models with higher accuracies. This is a general and powerful approach that could be used in many diverse circumstances.

Biomechanically aware classification developed for accelerometers was shown to have advantages beyond the immediate scope of behaviour recognition as well:

- (i) ***The same model can accommodate data from multiple sensors***
The study with magnetometers (Chapter 4) demonstrated that the same classification scheme developed for accelerometers (Chapter 3) could be used to separate coarse-scale behaviours. In fact, the very possibility of making a one-to-one biomechanical comparison between the accelerometer- and magnetometer-based models led to the novel finding that even for behaviours that accelerometers can recognise with high accuracy, i.e. dynamic behaviour (foraging, running), the magnetometer shows higher robustness to inter-individual variation. This finding is topical in light of the recent interest in making the most of multisensor tags (Williams et al. 2019).

- (ii) ***Model interpretability can provide insights on underlying assumptions and limits***
The sole reference for purely data-driven machine learning is groundtruth data. The biomechanically driven hierarchical model, however, contains an additional biomechanical 'reference' implicit in the model structure. This additional reference actually provided new insight in the following two cases on steps further upstream that would normally be 'inaccessible' to black-box machine learning:

- *Unlabelled resting posture in meerkats revealed through misclassification*

The underlying hypothesis during model development was that vigilance and resting would have two different postures (Chapter 3). Resting misclassification and the interpretability of the model, however, revealed that resting was actually carried out in two postures – belly-flat (as assumed) and curled-up (new) – revealing insights into limitations of the groundtruthing process itself.

- *A top-of-carapace sensor in leopard tortoises can't distinguish head movements during feeding*

With data collected using an accelerometer affixed to a scute at the top of the carapace (Figure 7.1) in 10 free-living, wild leopard tortoises (*Stigmochelys pardalis*) inhabiting the southern Kalahari desert, we⁷ applied a similar coarse-scale behaviour recognition scheme as the one developed for meerkats to separate the most common tortoise behaviours: resting, eating, and walking. It was found that the model succeeded in separating static (resting, eating) and dynamic (walking) behaviours with high (90.5%) accuracy, but that the sensitivity and precision of eating recognition were both <75%. Eating and resting were found to have similar levels of intensity, which confounded their separation. Upon watching the videos again to understand this somewhat surprising result, we realised that while eating, the carapace (to which the sensor is affixed) doesn't move when the leopard tortoise head cranes its neck to bite off grass and chew it. This led to the insight that to detect eating, future studies must reflect on other sensor locations to capture the head movement that a scute-fixed sensor might miss.

⁷ This study was a Master thesis project done by Laure Bruyère at EPFL under my direct supervision and the overall direction of Prof. Kamiar Aminian. The project was proposed by Prof. Marta Manser at the University of Zurich (UZH), and all data were collected by Seija-Mari Filli as part of a side project in her own MSc thesis at UZH.



Figure 7.1. **Leopard tortoise tagged with an accelerometer.** In an additional study done during this PhD (see footnote 7), we used the coarse-scale behaviour recognition scheme developed for meerkats to separate the most common leopard tortoise (*Stigmochelys pardalis*) behaviours: resting, eating, and walking. The model distinguished static (resting, eating) and dynamic (walking) behaviours with high (90.5%) accuracy, and revealed that to detect eating, future studies must reflect on other sensor locations to capture the head movement that a scute-fixed sensor will miss.

(iii) ***Accurate coarse-scale behaviour recognition models can be built with data from a single individual***

In an analysis aimed at assessing the generalisation ability of the accelerometer-based coarse-scale behaviour recognition model, we⁸ varied the number of meerkat recordings used to train the model and found that results with models trained on data from only one (any one) individual (and tested on the remaining four individuals) were similar

⁸ This study was part of a semester project (involving 1 day of work per week for 14 weeks) done by Mahmoud Zgolli at EPFL; the project was proposed and directly supervised by me, with the overall direction of Prof. Kamiar Aminian.

to those obtained during leave-one-individual-out (LOIO) cross-validation (CV), where models were trained on four individuals and tested on one individual. This indicates that using biomechanically aware models might even reduce the field effort necessary to collect enough training data. Further, this is the first study (though unpublished yet), as far as I know, to carry out leave-several-individuals-out CV for wild animal behaviour recognition.

(iv) ***The concept of microevents is applicable during static behaviour, too.***

The concept of microevents was not only useful in elucidating fine-scale dynamic behaviour in meerkats (Chapter 5), but also events during static behaviour, such as head turns during vigilance⁹. A different algorithm using accelerometer data, based on the physics of circular motion to model the turning of meerkat's torso during a head-turn during bipedal vigilance behaviour, detected head-turns with an overall accuracy of 97%. Head turns can be useful to characterise vigilance behaviour (e.g. varying levels of alertness to predator or rival-group presence) and, in tandem with other sensors, such as microphones, can be used to remotely obtain a better understanding of the context of calls performed during vigilance (Rauber & Manser 2017).

7.1.2 | Cross-validation (CV)

The main contributions here are: (i) the finding that not all CV techniques are equal, (ii) that LOIO CV is the most appropriate for animal behaviour recognition studies, and (iii) a new SMOTE-based data resampling approach for addressing inter-individual imbalance in data.

Results with three CV techniques were compared in Chapter 3: equally distributed 10-fold (EQDIST), Stratified 10-fold (STRAT), and LOIO. Model

⁹ This study was part of two semester projects done by Hamza Chaoui and Marion Groperrin at EPFL; the projects were proposed and directly supervised by me, with the overall direction of Prof. Kamiar Aminian.

performance during EQDIST answers the question: how well can the model separate behaviours of interest using the chosen features? There are two simplifying assumptions in EQDIST: first, we force each behaviour to have the same frequency of occurrence, and second, by pooling all data together, we do not account for inter-individual differences. Model performance during STRAT answers the question: how well can the model separate behaviours when the naturally occurring frequency of occurrence and duration of each behaviour (which are often unequal in continuous recordings of behaviour, e.g. a lot of medium-intensity foraging, and little high-intensity running in meerkats) are taken into account? There is still one simplifying assumption in STRAT: data from all individuals are pooled together, so it does not take inter-individual differences into account. Model performance during LOIO answers the question: how well can the model recognise behaviours recorded on a new individual unseen by the model during training? Both simplifying assumptions used in EQDIST are relaxed in LOIO, and thus LOIO is the most general of the three. Since the ultimate goal of most animal behaviour recognition algorithms is to be able to perform well on data from new individuals, LOIO CV, despite being strict and despite data being difficult to obtain on multiple individuals due to logistic difficulties, is the most appropriate test of robustness of animal behaviour recognition models. This is the first study, as far as I know, to perform LOIO CV on data collected on wild individuals. The only other study to have performed LOIO CV in animal behaviour recognition is Hammond et al. 2016, but training data were obtained from captive animals in their case.

While working on identifying fine-scale foraging modes in meerkats (Chapter 5), I encountered the hitherto unaddressed data-science problem of inter-individual imbalance in data, i.e. when one or more behavioural classes are not approximately equally represented across different individuals. I solved it by extending the popular SMOTE approach; the dataset balanced with the new technique will have equal representation of each class from each individual. This method should be used when the frequency of occurrence of different behavioural classes in the dataset does not reflect naturally occurring frequency distributions of these classes. This was the case in Chapter 5, where all chewing microevents were annotated across all individuals, and then digging and searching microevents were annotated across individuals until their total representation in the dataset (in terms of duration) became roughly equal to

that of chewing. This was done to prevent the classical problem of inter-class imbalance in data that the data-science community has traditionally encountered (Branco, Torgo & Ribeiro 2016). The proposed data resampling technique will be especially useful for animal behaviour recognition studies, where, in many cases, animals can only be observed opportunistically, and the resulting dataset may not reflect naturally occurring frequency distributions of the different behaviours of interest. Note, however, that LOIO CV performed on a balanced dataset like the one described here (Chapter 5) would be a ‘softer’ version of the LOIO CV performed on naturally imbalanced datasets (Chapter 3), and, consequently, possibly less robust.

7.1.3 | MET-based energy expenditure (EE) estimation

The main contribution here is a new method to estimate EE that is based on metabolic equivalent of task (MET) and behaviour recognition. In this method, durations of recognised behaviour can be multiplied by behaviour-wise MET coefficients (given as multiples of resting metabolic rate, RMR) to obtain total EE. If RMR is available in kilo-Joules (kJ), then MET-based EE can yield values in kJ as well.

MET-based EE correlated very strongly with EE determined using vectorial dynamic body acceleration (VeDBA), one of the prevalent methods to derive EE from acceleration data that has been shown to correlate with rate of oxygen consumption in several animal species. This points toward the possibility that MET-based EE may also be applicable across a wide range of species. Three additional factors make the MET method a powerful technique to estimate EE:

- (i) *The possibility to obtain MET coefficients from past studies on the same or surrogate species, or from established empirical relationships.*

The fact that MET coefficients for different activities for the meerkat, such as vigilance, searching, digging, and chewing, were all obtained from EE measurements on other species in past studies, and that the resulting total EE was in good agreement with measurements made using the doubly labelled water technique in a past study on a similar population of meerkats, offers a novel and powerful extension to the

idea of using surrogate animal species when actual measurements on the study species are not possible. When MET coefficients were not available even for surrogates, such as for running and searching in meerkats, an empirical relationship from a past study (Taylor, Schmidt-Nielsen & Raab 1970) that relates moving speed to EE, was adapted (Chapter 6) to estimate the respective MET coefficients. Further, MET-based EE can be computed in kJ even when RMR is not available for some species by using established empirical relationships that relate body size to RMR (Kleiber 1932). For instance, RMR for meerkats measured using doubly labelled water was 241 kJ/day (Scantlebury et al. 2004), whereas that estimated from Kleiber's relationship was 206 kJ/day (for a 600 g meerkat), which shows fair agreement.

- (ii) *Introducing a behaviour recognition step between acceleration data and EE estimation makes EE 'immune' to non-body acceleration and imprecise gravity compensation.*

By passing through the behaviour recognition step, the MET method overcomes problems that may arise when EE is based directly on acceleration data – that of imprecise gravity compensation, and that of separating body acceleration from undesirable accelerations produced due to sensor-body impacts and soft tissue artefacts. For instance, this step improved EE estimates compared to activity counts in human studies (Bonomi et al. 2009). These problems are effectively 'delegated' to behaviour recognition where (a) no assumptions are made about the precision of gravity estimation, and (b) sensor impacts and artefacts are actually leveraged to identify complex behaviour involving microevents.

- (iii) *MET-based EE estimation only requires durations of behaviours, not the accelerometer signal itself.*

This feature lends generality to the MET-based technique due to its non-dependence on acceleration data in particular. This means that if behaviour can be known using other sensors, e.g. magnetometer

(Chapter 4), depth in the case of marine animals, or even through visual observation, MET coefficients can be used to obtain EE estimates.

Using behaviour-wise MET coefficients allows the computation of meaningful metrics to analyse behaviour that would not be possible to obtain with behaviour durations alone, such as the definition of foraging efficiency in terms of duration of chewing per kJ of foraging effort (searching and digging) for meerkats.

There are two major limitations of using MET-based coefficients to estimate EE. First, this method assigns one ‘lumped’ coefficient per behaviour, and thus variations of EE expenditure within the same activity cannot be taken into account (e.g. an average speed of 12 km/hr was assumed for all running bouts, whereas running could be faster or slower than this; also, one-armed digging could be performed less or more vigorously). Secondly, in this study, the same value of RMR of 241 kJ/day (Scantlebury et al. 2004) was used for all individuals; RMR is known to be proportional to mass^{0.75} (Kleiber 1932), and body mass can indeed vary not only across individuals (by 12.5% in the dataset in Chapter 6, computed as standard deviation divided by mean) but also for the same individual overnight (~5% of body mass in meerkat pups, Russell et al. 2002). Thus, the MET-based method is better suited for studies involving time-energy budget estimation and investigation of qualitative differences in behavioural patterns, rather than performance quantification or investigation of fine-scale physiological differences between individuals.

7.1.4 | Meerkat foraging strategy

The main contribution here is the new finding that body-mass differences in female meerkats lead to significantly different foraging strategies. The resulting differences suggest that lighter females may be targeting low-value prey and foregoing opportunities to capture high-value prey items. This may be to avoid losing high-value prey to scrounging attempts made by heavier females. Low-value prey lies near the surface, and can be consumed quickly when found. On the other hand, high-value prey is located deep beneath the surface and requires significant investment of digging effort to catch, and, consequently, significant energy losses

when prey is stolen by heavier females. This finding is significant not only because these fine-scale differences in foraging would have been difficult to observe through direct observation, since meerkats switch between different fine-scale foraging modes very quickly, but also because it demonstrates that the accelerometer-based behaviour recognition model developed in this thesis can pick up subtle differences in behaviour.

7.2 | Broader discussion and critique

The methods presented in this thesis are primarily aimed for uptake in the biologging community. Putting the methods in the broader context of their utility and limitations for uptake in other biologging studies would therefore be in order.

Apart from the new biomechanically driven approach to learning, three additional reasons may have been responsible for the high behaviour recognition accuracies obtained in this thesis: (i) precise electronic synchronisation (precision of ± 0.04 s) between acceleration signals and video (Chapter 4) during the groundtruthing stage, (ii) use of a high-enough sampling frequency (100 Hz), and (iii) accurate video annotation based on a well-defined ethogram. None of these three factors were gotten right in the very first attempt at data collection, where a different tagging device (earlier version of the tag used in Chapter 6) was deployed on eight wild meerkats using GPS-based synchronisation¹⁰ with a relatively low precision of ± 1 s (along with manufacturer firmware issues in several cases, see footnote 10, which further affected synchronisation accuracy). The accelerometer sampling frequency was 12.5 Hz, and the ethogram lumped all kinds of gait into a single category ‘moving’. The following problems were encountered in this first attempt which led to a successful second attempt: (i) manufacturer firmware issues persisted, and it became clear that the desired level of synchronisation would not be achieved with this device, (ii) the accelerometer sampling frequency was suspected to be too low since preliminary analyses indicated running frequency to be ≥ 4 Hz, and

¹⁰ GPS-based synchronisation involved filming the screen of a hand-held GPS (which had a precision of ± 1 second) to mark the beginning and end of videos, and matching them to the tag’s GPS timestamps; the tag was programmed to acquire GPS fixes once per second. Several issues in the manufacturer’s firmware were encountered in the writing of GPS timestamps to memory, which added to inaccuracies in synchronisation.

impact-generated high-frequency accelerations were expected to further exacerbate the problem of aliasing, and (iii) the category of ‘moving’, expected to yield rhythmic acceleration patterns at a characteristic frequency, was instead found to encompass too many gait styles and speeds, which led to large variation in recorded signal patterns. In light of these factors, the decision was made to shift hardware development in-house, redo groundtruthing data collection with precise synchronisation, increase the accelerometer’s sampling frequency to a ‘comfortably’ high value of 100 Hz, and increase the behaviour resolution and confidence in video annotation by separating slower forms of locomotion that were difficult to tell apart visually from ‘fully developed running’ (where it was easy to tell that the animal was galloping). Thus, improving the system used to collect groundtruth data required expert engineering knowledge that may not be easily accessible to field biologists, but was essential for obtaining high accuracy. Further, one outcome of this experience was the realisation that high behaviour recognition accuracies may only be possible to obtain when behaviours in the ethogram are defined with some knowledge of what the accelerometer can ‘see’. This was highlighted not just with the example of the behaviour label ‘moving’ given here, but also in the multiple resting postures observed for meerkats that confounded vigilance versus resting recognition (Chapter 3), and inability of a top-of-carapace sensor to distinguish leopard tortoise feeding (Section 7.1.1).

Developing good features requires some signal processing expertise that may not be readily available to non-specialists such as ecologists/biologists. Even though the features for coarse-scale behaviour recognition have been clearly laid out (Chapter 3 and 4), and a large list of 200 features have been made available for fine-scale behaviour recognition (Chapter 5), new features may be required to separate fine-scale behaviour in a different study. This is because the range and diversity of animal movement and behaviour, and spur-of-the-moment adjustments to objects or characteristics in the environment are simply too vast to create a ‘perfect’ or ‘complete’ set of features. However, new, good quality features, essential to accurately infer behaviour regardless of the type of model used, could be challenging to engineer because of the domain-specific (i.e. signal processing) knowledge required. Thus, even though the classification scheme itself was rendered transparent in this thesis, the discussion here suggests that signal processing (a central

component of feature engineering) might itself be a ‘hidden black box’ for non-specialists – ‘hidden’ since it lies ‘behind’ the traditional black box comprised of machine learning and feature selection. One solution to this is to have greater and deeper interdisciplinary collaboration between engineers and biologists, as has indeed been recognised by the biologging community recently (Williams et al. 2019). Another more long-term (and perhaps more radical) solution would be to introduce signal processing courses early on in an ecologist’s university curriculum.

It was found that the more specific the behaviour to be classified, the stronger the assumptions about sensor orientation with respect to the animal’s body needed to be. The coarse-scale behaviour recognition model’s features (Chapters 3 & 4) were based only on the surge axis, which was more robust than the other two axes to orientation changes due to sensor rotation around the animal’s neck, while the fine-scale behaviour recognition model’s features were based on data from all three axes. Nevertheless, the robustness of the model across individuals and model parameters indicated that increased sensitivity to sensor orientation changes did not pose a major problem.

Some other general aspects worth mentioning here are:

- The effect of change in sensor location on model accuracy could not be assessed in this study (since meerkats only bore collars), i.e. the model developed for a collar sensor may perform poorly for a backpack sensor for the same species. This is because signal magnitudes will, in general, be different in different locations.
- The algorithms for static behaviour detection do not yet account for noisy acceleration that may be recorded due to the medium of propagation’s movement, e.g. waves for seabirds sitting on the surface of the water, where determining when the bird is static or slowly paddling might be difficult.

7.3 | Future perspectives

Improving static behaviour recognition through transition detection

Static behaviours (vigilance and resting) were difficult to tell apart in both the accelerometer (Chapter 3) and magnetometer (Chapter 4) studies, and this may be because of dependence on only postural information. One approach that could be tried in future studies is to add a ‘transition-based correction’ layer to the coarse-scale behaviour recognition model, inspired from human studies using ‘sit-to-stand’ and ‘stand-to-sit’ transitions to improve detections of sitting and standing (Salarian et al. 2007). Here, instead of classifying static activities window-wise, one would characterise the nature of transition from the end of the preceding dynamic behaviour bout and the end of the current static bout. For instance, there would always be a negative rotation about pitch (if pitch axis points leftwards with respect to the animal) to go from dynamic behaviour (where the torso is horizontal) or resting (where the torso is horizontal or curled up) to vigilance, where the head is pointing upward. On the other hand, there would be a positive rotation about pitch in going from dynamic behaviour or vigilance to resting. This could be characterised either using a magnetometer or a gyroscope. We¹¹ performed a preliminary study on this using gyroscopes, with limited success since it was discovered that there could be transitions between different modes of vigilance (quadrupedal to bipedal) that could generate confounding positive pitch rotations. Nevertheless, this approach could be a valuable tool in other scenarios with lesser types of static behaviour.

Identifying bouts of behaviour

One valuable extension of the current work would be to detect bouts of behaviour from temporal sequences of classified behaviour. Two examples from the meerkat study would be:

- (i) *Combining successive chewing microevents to define a ‘prey ingestion event’.*

Chewing microevents occurring in the acceleration signal within a

¹¹ This study was part of a semester project done by Hugues Vinzant at EPFL; the project was proposed and directly supervised by me, with the overall direction of Prof. Kamiar Aminian.

short space of time could be grouped together to define a single prey ingestion event. From the detection of prey ingestion events, one could go further and combine information on, say, the number of constituent chewing microevents, duration of prey ingestion event, and intensity of acceleration, to qualitatively estimate the size of ingested prey items, e.g. in terms of small, medium or large, as defined in Brotherton et al. 2001. This would help fine-scale quantification of dietary composition, thereby extending the findings from previous studies based on direct observation (Doolan & MacDonald 1996). Further, variation in ingested prey size with individual traits such as body mass could help confirm the inference made in this work (Chapter 6) about lighter females primarily targeting small prey and heavier females targeting both large and small prey.

The challenge here would be to develop rules to determine when chewing microevents are sufficiently closely spaced in time for them to belong to the same ingestion event, and to account for possibly misclassified microevents occurring within the same ingestion event.

(ii) *Combining detection of searching, and one- and two-armed digging to define 'foraging attempts'.*

The composition of foraging attempts could reveal individual foraging styles at very fine scale. This would enable one to distinguish between, say, individuals with a 'precision-based' strategy, where searching is continued until a digging spot more likely to yield valuable prey is found, and those with a 'brute-force' strategy, where digging occurs more often and less selectively until prey is found. This would help investigate, for instance, whether foraging style influences acquisition or maintenance of dominance within the group. The challenge here would be to distinguish between intra-bout and inter-bout non-foraging intervals in order to determine when a bout stops, which is known to be a non-trivial problem (Rook & Huckle 1997).

Dynamical characterisation of temporal sequences of behaviour

The main application of behaviour recognition investigated in the current thesis has been based on ‘static’ metrics (metrics insensitive to the temporal order of appearance of different behavioural states, e.g. total duration of a behaviour over the full day). An interesting extension of the current work would be to analyse behavioural sequences using ‘dynamic’ metrics (metrics sensitive to temporal order of behavioural states, e.g. sample entropy, Richman & Moorman 2000, and Lempel-Ziv complexity, Paraschiv-Ionescu et al. 2012). This would be especially interesting given that human studies have shown that static metrics do not always reveal differences in patients and healthy controls, but dynamic metrics do (Paraschiv-Ionescu et al. 2012).

Unravelling demographic mechanisms through behaviour recognition

Female dispersal in meerkats leads to body mass loss and increased stress levels (Maag et al. 2019), but the behavioural mechanisms through which these effects operate are unclear. This is because dispersing meerkats are difficult to observe since they have no defined territory and can traverse several kilometres a day. Remote observation using the accelerometer-based behaviour recognition developed in this work could, with a tag with suitable battery life (~3 months) and GPS, be used to investigate the behavioural changes (e.g. foraging success, alertness levels) and space use during dispersal.

References

- Bonomi, A. G., Plasqui, G., Goris, A. H., & Westerterp, K. R. (2009). Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. *Journal of applied physiology*, *107*(3), 655-661.
- Branco, P., Torgo, L., & Ribeiro, R. P. (2016). A survey of predictive modeling on imbalanced domains. *ACM Computing Surveys (CSUR)*, *49*(2), 1-50.
- Brotherton, P. N. M., Clutton-Brock, T. H., O'riain, M. J., Gaynor, D., Sharpe, L., Kansky, R., & McIlrath, G. M. (2001). Offspring food allocation by parents and helpers in a cooperative mammal. *Behavioral Ecology*, *12*(5), 590-599.
- Doolan, S. P., & MacDonald, D. W. (1996). Diet and foraging behaviour of group-living meerkats, *Suricata suricatta*, in the southern Kalahari. *Journal of Zoology*, *239*(4), 697-716.
- Kleiber, M. (1932). Body size and metabolism. *Hilgardia*, *6*(11), 315-353.
- Maag, N., Cozzi, G., Bateman, A., Heistermann, M., Ganswindt, A., Manser, M., ... & Ozgul, A. (2019). Cost of dispersal in a social mammal: body mass loss and increased stress. *Proceedings of the Royal Society B*, *286*(1896), 20190033.
- Paraschiv-Ionescu, A., Perruchoud, C., Buchser, E., & Aminian, K. (2012). Barcoding human physical activity to assess chronic pain conditions. *PLoS one*, *7*(2), e32239.
- Rauber, R., & Manser, M. B. (2017). Discrete call types referring to predation risk enhance the efficiency of the meerkat sentinel system. *Scientific reports*, *7*, 44436.
- Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, *278*(6), H2039-H2049.
- Rook, A. J., & Huckle, C. A. (1997). Activity bout criteria for grazing dairy cows. *Applied Animal Behaviour Science*, *54*(2-3), 89-96.

- Russell, A. F., Clutton-Brock, T. H., Brotherton, P. N., Sharpe, L. L., McIlrath, G., Dalerum, F. D., ... & Barnard, J. A. (2002). Factors affecting pup growth and survival in co-operatively breeding meerkats *Suricata suricatta*. *Journal of Animal Ecology*, *71*(4), 700-709.
- Salarian, A., Russmann, H., Vingerhoets, F. J., Burkhard, P. R., & Aminian, K. (2007). Ambulatory monitoring of physical activities in patients with Parkinson's disease. *IEEE Transactions on Biomedical Engineering*, *54*(12), 2296-2299.
- Taylor, C. R., Schmidt-Nielsen, K., & Raab, J. L. (1970). Scaling of energetic cost of running to body size in mammals. *American Journal of Physiology-Legacy Content*, *219*(4), 1104-1107.
- Williams, H. J., Taylor, L. A., Benhamou, S., Bijleveld, A. I., Clay, T. A., de Grissac, S., ... & Griffiths, R. C. (2019). Optimizing the use of biologgers for movement ecology research. *Journal of Animal Ecology*.

Curriculum Vitae

Pritish Chakravarty

Nationality: Indian

Date of birth: 10 April 1990

Email: pritch3@gmail.com

Phone: +41 78 883 53 80

Address: Chemin de la Prairie 60,
CH-1007 Lausanne, Switzerland

EXPERIENCE AND EDUCATION

- 11/2015 – 06/2020 Ecole polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland
PhD in Bioengineering, School of Engineering
Supervisor: *Prof. Kamiar Aminian*
Collaborating supervisor: *Prof. Arpat Ozgul (University of Zurich)*
- 02/2015 – 07/2015 Stanford University, CA, United States
Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland
MSc Thesis
Title: *‘Three-dimensional Meshes Calculated from Biomedical Imaging to Characterise Tissue Morphology in Knee Osteoarthritis’*
Supervisors: *Dr. Julien Favre (CHUV)*
Prof. Thomas Andriacchi (Stanford, Department of Mechanical Engineering)
- 09/2013 – 07/2015 Ecole polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland
MSc in Bioengineering, School of Engineering
- 11/2012 – 06/2013 Indian Institute of Technology Bombay (IITB), Mumbai, India
Research Assistant, Department of Mechanical Engineering
Title: *‘Hydrodynamic flow focusing of blood for cytometry applications’*
Supervisor: *Prof. Amit Agrawal*
- 07/2012 – 10/2012 Ernst & Young, Mumbai, India
Associate Consultant
- 07/2008 – 07/2012 Indian Institute of Technology Bombay (IITB), Mumbai, India
BTech (Hons) in Chemical Engineering

PROFESSIONAL ACTIVITIES

Journal Referee: *Methods in Ecology and Evolution*; *IEEE Internet of Things Journal*;
Journal of Zoology.

AWARDS

Recent

- 2020 Forschungskredit postdoctoral fellowship (CHF 44,154).
- 2019 Awarded certificate for MEE 2019 paper being recognized as one of the top downloaded MEE papers in 2018-2019.
- 2019 Shortlisted for Robert May Prize 2019 (among the top eight papers by Early Career Researchers in *Methods in Ecology and Evolution*).
- 2017 Rory Wilson prize for 'Wittiest Talk' at the 6th International Bio-logging Science Symposium (BLS6), Konstanz, Germany

Past

- 2008 Studentship Award by Council of Scientific & Industrial Research (CSIR), India.
- 2007 First in city-level Mathematics Olympiad in Chandigarh, India.
- 2006 National Talent Search Examination (NTSE) scholarship.
- 2006, 08 Certificates of Merit in English and Mathematics by the Central Board of Secondary Education (CBSE) for being among the top 0.1% in the national examinations.

PUBLICATIONS

Journal

Published

1. **Chakravarty, P.**, Maalberg, M., Cozzi, G., Ozgul, A., & Aminian, K. (2019). Behavioural compass: animal behaviour recognition using magnetometers. *Movement Ecology*, 7(1), 28. Extensive press coverage: [EPFL News](#), [British Ecological Society](#), [ScienceDaily](#), [SwissInfo](#), [Le Matin](#), [Le Soir](#), [24 Heures](#), [20 Minutes](#), [Tribune de Genève](#).
2. **Chakravarty, P.**, Cozzi, G., Ozgul, A., & Aminian, K. (2019). A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods in Ecology and Evolution*, 10(6), 802-814.
3. Tripathi, S., **Chakravarty, P.**, & Agrawal, A. (2014). On non-monotonic variation of hydrodynamically focused width in a rectangular microchannel. *Current Science*, 1260-1274.

Under review

4. **Chakravarty P.**, Cozzi G., Dejnabadi H., Léziart P.A., Manser M., Ozgul A., Aminian K. (2020, 'minor revision'). Seek and learn: Automated identification of microevents in animal

behaviour using envelopes of acceleration data and machine learning. *Methods in Ecology and Evolution*.

In preparation/to be submitted

5. **Chakravarty P.**, Cozzi G., Harrison N., Maag N., Manser M., Ozgul A., Aminian K. Behavioural time-energy budgets and foraging strategies in meerkats: a study with validated accelerometer-based behaviour recognition.

Conference

1. **Chakravarty P.**, Cozzi G., Ozgul A., Aminian K. *An accelerometer-based behaviour recognition model as a tool to investigate meerkat dispersal*. The 6th International Bio-Logging Science Symposium, Konstanz, Germany, 25-29 September, 2017.
2. Jolles-Haeberli, B., Meyer, V., Cavinato, A., **Chakravarty, P.**, Omoumi, P., & Favre, J. (2017, May). Cartilage thickness and bone density are related in healthy tibias. In *Orthopaedic Proceedings* (Vol. 99, No. SUPP_9, pp. 66-66). The British Editorial Society of Bone & Joint Surgery.
3. Jolles-Haeberli, B., Meyer, V., **Chakravarty, P.**, Cavinato, A., Omoumi, P., & Favre, J. (2015, June). Subchondral bone mineral density is heterogeneous on the femoral condyles and varies with osteoarthritis. In *Swiss Medical Weekly* (Vol. 145, pp. 56S-56S).

TEACHING

Teaching assistant:

Sensors in medical instrumentation (Spring 2016, 2017, 2018)

Biomedical optics (Fall 2014)

Linear algebra (Fall 2014)

Students supervised (co-advised with Prof. K. Aminian):

- | | |
|------|---|
| 2019 | Laure Bruyère
<i>Master thesis: Behaviour recognition in wild leopard tortoises in the Kalahari using scute-fixed accelerometers</i> |
| 2019 | Hugues Vinzant
<i>Semester project: Can sensor fusion improve animal behaviour classification when the accelerometer reaches its limits?</i> |
| 2018 | Marion Groperrin
<i>Semester project: Meerkat head turns: how post-classification biomechanical event detection using accelerometers can augment behaviour recognition performance</i> |
| 2018 | Pierre-Alexandre Léziart
<i>Semester project: Characterising and detecting successful foraging events in wild meerkats</i> |

- 2018 Mahmoud Zgolli
Semester project: Frequently frugal: What is the minimum accelerometer sampling frequency required for accurate meerkat behaviour recognition?
- 2018 Hamza Chaoui
Semester project: Detecting head turns during vigilance behaviour in wild meerkats using inertial sensors
- 2017 Maiki Maalberg
Master thesis: Meerkat behaviour recognition using triaxial magnetometers
- 2016 Gauvain Rudler
Semester project: Quantifying the table tennis service using inertial sensors attached to the racquet

Institute Student Mentor (IITB, India):

Guided 15 freshmen through their first year of college – helping them adjust to a new environment away from home, and assisting them with their academic, extra-curricular, social, and emotional issues.

TECHNICAL SKILLS

Scientific computing	<i>MATLAB, Weka, some Python</i> Past: <i>COMSOL, SolidWorks, ImageJ</i>
Graphical design	<i>Inkscape</i>

LANGUAGES

Bengali: Native
Hindi: Native
English: Near-native
French: Fluent
Punjabi: Intermediate

INTERESTS, HOBBIES

Vipassanā meditation, ancient Indian spiritual knowledge.

Past:

Table tennis
 (played for the Lausanne Club de Tennis de Table in the Ligue Nationale C; represented Switzerland (through EPFL) at the 2016 European Universities Games; captained the

university team while in India).

Guitar/songwriting/singing

(played at the Fête de la Musique Lausanne 2016 and Fête de la Musique Monthey 2016; several informal performances in numerous bands in Switzerland and India).

Squash

(won the regional tournament *Championnat Vaudois* in 2015; reached the top 32 in the 36th Bombay Gymkhana Squash Open in 2011).

Awarded the Ambrose Otieno Award and Institute Sports Citation for exceptional performance in multiple sports over four years at IITB, India.