Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Wearable technologies for behavioural assessment in captive cheetah (*Acinonyx jubatus)*

A thesis presented

in partial fulfilment of the requirements

for the degree of

Master of Science (Zoology)

Massey University, Palmerston North,

New Zealand

Yilin Wang

2026

Abstract

Monitoring animal behaviour is important for zoo animal welfare assessment, but direct observation is limited by observer time, animal visibility and the difficulty of recording behaviour continuously. Triaxial accelerometers can support long-term behaviour monitoring. Previous cheetah studies have used accelerometer data for behavioural classification and feeding event detection, but less attention has been given to linking behaviour with time budgets and enclosure use. This thesis developed and applied a collar-based triaxial accelerometer approach for behavioural assessment in two adult female cheetahs (*Acinonyx jubatus*) at Auckland Zoo. Triaxial acceleration data were used to develop a series of random forest models across different levels of behaviour resolution, with video-labelled behaviours used as the reference data for model training and evaluation. After the most optimal behavioural structure was selected, it was applied to video-based time-budget analysis, collar-video comparison and enclosure-use records. Model performance improved when rare or visually similar behaviours were removed or combined into broader behavioural classes. The model selected had the most useful structure because it retained six behaviourally meaningful categories, active, drink, eat, groom/allogroom, lying asleep and resting alert, while maintaining strong classification performance. In the video-based time budget, both cheetahs spent most retained behavioural time in resting alert and lying asleep with public viewing side (Zone 3) and three-tier vertical stand (Zone 6) recorded most often. Triaxial accelerometer monitoring can be useful for assessing broad behavioural patterns in captive cheetahs and when combined with enclosure-use analysis, can also help identify habitat features relevant to husbandry decisions, enclosure design and long-term welfare monitoring.

Acknowledgements

Firstly, I would like to express my deepest gratitude to my supervisors, Dr Chris Andrews, Professor David Thomas, and Professor Murray Potter, for their guidance, patience and valuable feedback throughout this thesis. Their knowledge, advice and encouragement, as well as the time and effort they devoted to this project, were very important in helping me develop the research, improve my analysis and complete this thesis.

I am also very grateful to Auckland Zoo for allowing this research to be carried out. I would especially like to thank Lauren Booth, Team Leader Carnivores, and the experienced carnivore keepers, including Kirsten, Ellie, Sue, Tory, Emma and Nick, for their support during the study. Their assistance with collar attachment, removal and reattachment when required, as well as their knowledge of Qia and Quartz and their practical help, made this project possible.

Finally, I would like to thank my family for their support throughout my Master's study. I am especially grateful to my father, Guosheng, my mother, Hongya, and my husband, Mingliang Chi, for their financial and personal support throughout my study. Their support allowed me to focus on my research and complete this thesis.

Table of Contents

Abstract i

Acknowledgements ii

Chapter 1 – Literature review 11

1.1 Introduction 11

1.2 The Five Domains Model 12

1.2.1 Welfare Assessing in Zoos 14

1.3 Felid welfare in captivity 17

1.3.1 Evaluating welfare in captive felids 17

1.3.2 Factors affecting felid welfare 20

1.4 Behavioural monitoring in captive felids 23

1.5 The use of triaxial accelerometers in animal behaviour research 25

1.6 Study site and species 27

1.7 Thesis aim and objectives 29

Chapter 2 – Development of machine learning models for classifying captive cheetah behaviour from collar-based triaxial accelerometer data 31

2.1 Introduction 31

2.2 Materials and Methods 32

2.2.1 Study location and Animals 32

2.2.2 Experimental design 33

2.2.3 Scoring behaviour from video footage 35

2.2.4 Processing the acceleration data 37

2.2.5 Machine learning using random forests 37

2.3 Result 42

2.3.1 Model 1 (24 behavioural classes) 44

2.3.2 Model 2 (16 behavioural classes) 44

2.3.3 Model 3 (11 behavioural classes) 46

2.3.4 Model 4 (eight behavioural classes) 47

2.3.5 Model 5 (six behavioural classes) 48

2.3.6 Model 6 (three behavioural classes) 50

2.4 Discussion 50

2.5 Conclusion 55

Chapter 3 – Enclosure use and time budget analysis of two captive cheetahs *(Acinonyx jubatus*) 57

3.1 Introduction 57

3.2 Materials and Methods 58

3.2.1 Study location and animals 58

3.2.2 Habitat management 58

3.2.3 Video recording and behavioural scoring 60

3.2.4 Processing the video data 62

3.2.5 Collar-video behavioural validation 62

3.3 Results 64

3.3.1 Video observed behavioural time budget 64

3.3.2 Collar versus video behavioural validation 65

3.3.3 Enclosure zone uses and behavioural distribution by zone 67

3.4 Discussion 68

3.5 Conclusion 74

Chapter 4 - General Discussion 76

4.1 Conclusions 78

5.0 References 80

Appendix 88

List of Figures

Figure 1. 1. Conceptual transition from the Five Freedoms to the Five Domains model, moving from a focus on eliminating negative states toward a broader framework that also encompasses positive aspects of animal welfare, (adapted from FAWC, 1993; Mellor et al., 2020). 13

Figure 1. 2. Animals experience alternating positive and negative affective states over time, with overall welfare determined by the balance and frequency of positive experiences (ZAA, 2025). 17

Figure 1. 3. Visible tail markings used to distinguish Qia and Quartz during video scoring. Qia had a larger white marking near the distal end of the tail, whereas Quartz had a smaller white marking in the same area. 28

Figure 1. 4. Overview of the cheetah enclosure at Auckland Zoo, showing the indoor habitat, public viewing area and main enclosure features. The red line indicates the main outdoor enclosure boundary. 28

Figure 2. 1. Diagram of the indoor and outdoor habitats, showing the locations of the four surveillance cameras used to monitor the animals. 33

Figure 2. 2. (A) ActiGraph® wGT3X-BT unit. (B) ActiGraph® wGT3X-BT device set in the Collar. (C) Cheetah with collar. 34

Figure 2. 3. Synchronisation of accelerometer (30 Hz) and video data. The highlighted sections indicate the period when the collar was static on the ground, which was used as a reference point for aligning the accelerometer data with the video footage. Based on this reference event, the accelerometer data were adjusted by +39 s for Quartz (A) and +55 s for Qia (B) to synchronise the collar-derived data with video records. Note: Timestamps are not shown on the x-axis because the raw data were sampled at 30 Hz, resulting in a very fine time scale 38

Figure 2. 4. Number of training observations for each behavioural category included across the six RF model development stages. After excluding the other and out of sight categories, the full dataset consisted of 223,909 observations with 70% (156,736 observations) used for training and 30% (67,173 observations) for test. The last four behaviours listed in Model 1 were removed from subsequent models due to limited sample size and/or poor model performance. Abbreviation: L, left; R, right. 40

Figure 2. 5. Comparison of model performance across the six random forest models. (A) Distribution of F1 scores across behavioural classes for each model. (B) CV% between observed and predicted behavioural prevalence for each model. CV% represents the variation between observed and detected prevalence. 43

Figure 3. 1. Layout of the cheetah enclosure showing the eight analytical zones, four surveillance cameras, and main habitat features used for behavioural and spatial analysis, including the indoor habitat, keeper's office, water bowl, pergola, public viewing area, and three-tier vertical stand. 59

Figure 3. 2. Video observed behavioural time budget for Qia (white) and Quartz (grey) based on Model 5 behavioural categories. Percentages were calculated from the retained behavioural dataset after non-model behaviours were excluded. 64

Figure 3. 3. Hourly regression relationships between video-observed and collar-classified behavioural percentages for six behaviours (A) Active, (B) Lying asleep, (C) Resting alert, (D) Groom/allogroom, (E) Drink, and (F) Eat. Each data point represents one hourly comparison between video-observed and collar-classified percentages. Regression lines, R^2 values and p values are shown for each behaviour. 66

List of Tables

Table 2. 1. A standardised ethogram for the Felidae including definitions for all base behaviours derived from Stanton et al. (2015). 36

Table 2. 2. Calculations for the parameters used to evaluate model performance. 42

Table 2. 3. Performance characteristics of random forest Models 1 and 2, including sensitivity, specificity, balanced accuracy, F1 score, observed prevalence, detection prevalence and coefficient of variation (CV%) and for each behaviour. CV% represents the variation between observed and detected prevalence. Values in the Average row are presented as mean ± SD across behavioural classes. Behaviours highlighted in red have a CV% > 20, while cells highlighted in orange indicate performance values < 0.70. 45

Table 2. 4. Performance characteristics of random forest Models 3 to 6. The table shows sensitivity, specificity, balanced accuracy, F1 score for each behaviour, observed prevalence, detection prevalence and coefficient of variation (CV%). CV% represents the variation between observed and detected prevalence. Values in the Average row are presented as mean ± SD across behavioural classes. Behaviours highlighted in red have a CV% > 20, while orange shading indicates performance values < 0.70. 47

Table 2. 5. Confusion matrices showing the percentage of predicted and observed behaviours for Models 3 and 4. Correct classifications are highlighted in yellow, while red cells indicate misclassifications greater than 15%. 49

Table 2. 6. Confusion matrices showing the percentage of predicted and observed behaviours for Models 5 and 6. Correct classifications are highlighted in yellow, while red cells indicate misclassifications greater than 15%. 50

Table 3.1. Description of the eight zones used for enclosure use analysis. 60

Table 3.2. Behavioural categories and corresponding behaviours used for video scoring. Base behavioural definitions were adapted from Stanton et al. (2015). Out of sight and keeper-led walking were recorded as additional observed categories and were not part of the six Model 5 behavioural categories. 61

Table 3.3. Calculations used for behavioural time-budget; enclosure use and collar-video validation analyses. 63

Table 3.4. Daily comparison between video observed behaviours and collar datasets percentages for Qia and Quartz. Values are presented as mean ± SEM; Difference and CV% are reported as percentage points. 65

Table 3.5. Total enclosure zone use. Values are presented as minutes (% of total observation time), calculated from 7140 min of observation for each cheetah. Brief zone descriptions are provided for reference. 67

Table 3.6. Behavioural distribution across enclosure zones for Qia and Quartz based on the full video-observed dataset. Values are presented as minutes (% of total observation time), calculated from 7140 mins of observation for each cheetah. Cells with no recorded observations are shown as 0. Highlighted cells indicate the main behavioural records in that zone. 70

Appendix

Appendix 1. Description of identifier variables used for model building. 88

Appendix 2. Confusion matrix showing the percentage of predicted and observed behaviours for Model 1. Correct classifications are highlighted in yellow, while red cells indicate misclassifications greater than 15%. 89

Appendix 3. Confusion matrix showing the percentage of predicted and observed behaviours for Model 2. Correct classifications are highlighted in yellow, while red cells indicate misclassifications greater than 15%. 90

Chapter 1 – Literature review

1.1 Introduction

Animal welfare assessment is an important part of modern zoo management, as animals housed in zoos depend on people to provide appropriate physical, social and behavioural conditions. Contemporary welfare science has moved beyond the narrow focus of avoiding poor welfare and increasingly considers both negative and positive welfare states. The Five Domains Model provides a useful framework for this approach because it links nutrition, physical environment, health and behavioural interactions to infer an animal's overall mental state (Mellor & Beausoleil, 2015; Mellor et al., 2020). In zoo settings, this framework is relevant because welfare assessment must consider not only whether animals are healthy, but also whether their environments allow them to express a broad range of species-appropriate behaviours (Miller et al., 2020).

Scientific evidence is increasingly central to zoo animal management, particularly where husbandry practices have historically been shaped by tradition, practical experience and assumptions about species requirements. Zoo animal welfare research has emphasised the need for scientific evidence to support management decisions and improve welfare assessment in practice (Melfi, 2009; Salas et al., 2024). Behaviour is one of the most widely used welfare indicators because it can provide non-invasive information about how animals respond to their environment. Behavioural data may also help identify welfare concerns earlier than some physiological or pathological measures, since behaviour can reflect an animal's initial attempts to cope with environmental challenges (Mohapatra et al., 2014).

Behavioural assessment in captive felids is complicated by the fact that resting and inactivity may occupy substantial parts of the daily activity pattern (Seyrling et al., 2022). Rather than treating these periods as an absence of useful information, activity budgets can be used to describe how behaviour is organised over time and how this organisation may relate to husbandry routines or enclosure conditions. Previous studies of captive felids have used activity budgets to describe activity patterns, resting behaviour, feeding related responses and abnormal repetitive behaviours (Mohapatra et al., 2014; Weller & Bennett, 2001). For felids

housed in zoos, these behavioural summaries become more informative when combined with spatial records, because enclosure structure, shelter, visual barriers and the quality of available space may influence where behaviours occur (Hunter et al., 2014).

Although direct behavioural observations remain valuable, they are constrained by observation time, animal visibility and the difficulty of recording behaviour continuously over long periods (Grünewälder et al., 2012). These constraints are especially relevant for felids, which are crepuscular and tend to spend long periods inactive or hidden from view (Weller & Bennett, 2001). Wearable technologies (e.g., external sensors) can extend behavioural monitoring by recording body movement and posture beyond the period of direct observation (Brown et al., 2013). When paired with video labelled observations, accelerometer data can be used by trained machine learning (ML) models that classify behaviours from movement patterns (Valletta et al., 2017).

The present thesis examines how wearable technologies and machine learning can support zoo animal welfare assessment in captive cheetahs. By linking collar-derived movement data with observed behaviour, this research explores how behavioural monitoring can be extended beyond direct observation and used to provide more continuous records of zoo animal behaviour. The following literature review provides the background for this research by considering zoo animal welfare assessment, welfare in captive felids, behavioural monitoring methods, and the use of triaxial accelerometers and machine learning in animal behaviour research.

1.2 The Five Domains Model

Broom (1986) defined animal welfare as “the state of an animal as it attempts to cope with its environment”. This definition emphasises that welfare is related to how an animal responds to the conditions in which it lives. Therefore, the development of objective indicators is important for assessing welfare across animal management systems, including zoological institutions.

The five freedoms framework is one of the earliest attempts to create a comprehensive checklist for animal welfare (Webster, 2008). It has achieved global recognition and remains

embedded in legislation, policy, and animal management practice (McCulloch, 2013). However, this framework has been criticised for being misinterpreted as completely attainable or even as animal "rights", and for concentrating primarily on the elimination of negative states (Mellor, 2016). This has led to the development of new models for welfare assessment (Figure 1.1).

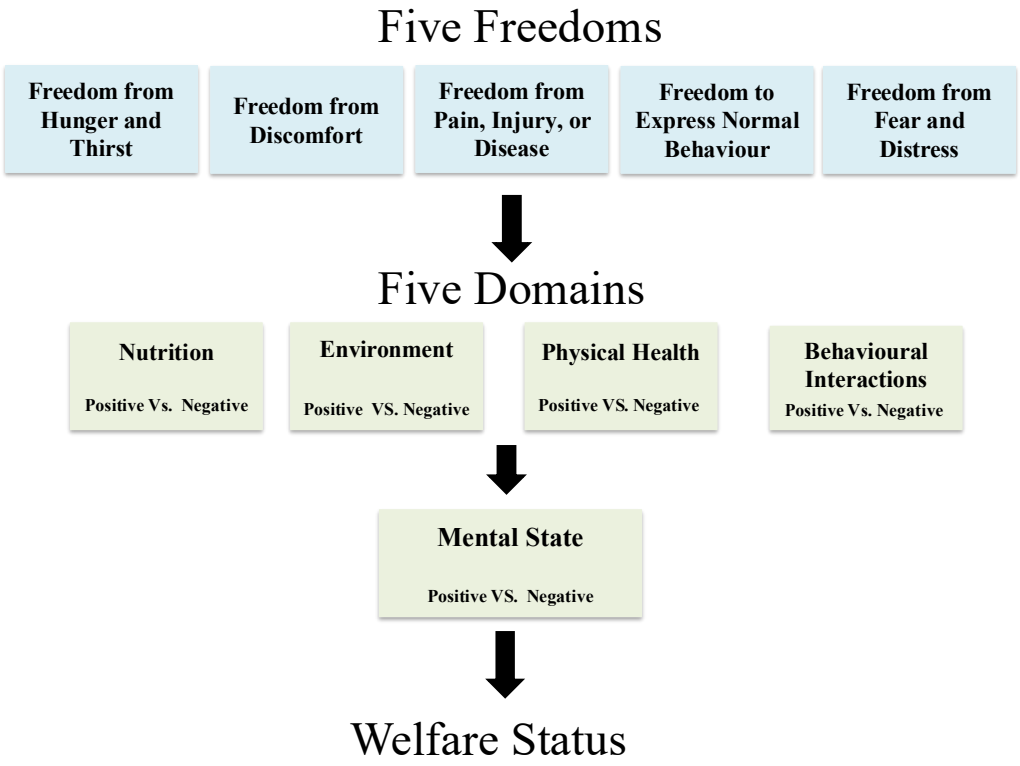


Figure 1. 1. Conceptual transition from the Five Freedoms to the Five Domains model, moving from a focus on eliminating negative states toward a broader framework that also encompasses positive aspects of animal welfare, (adapted from FAWC, 1993; Mellor et al., 2020).

The five domains model was introduced by Mellor and Reid (1994) to provide a more systematic evaluation of the welfare impacts of scientific procedures on animals. It builds on the foundational principles of the five freedoms model. The new model initially focused on identifying and grading the negative effects of research, teaching, and testing activities on sentient animals (Mellor & Reid, 1994). Later revisions have expanded the model to include positive affective states, such as comfort, pleasure, and positive social interactions (Mellor & Beausoleil, 2015). Further refinement emphasised human and animal interactions as important factors influencing animal welfare, especially in zoos and companion animal care settings (Mellor et al., 2020).

Presently, the five domains model comprises five interrelated domains: nutrition, environment, physical health, behaviour, and mental state. The first three domains focus on internal disturbances arising from nutritional deficits, environmental inadequacies, and health related conditions (Mellor et al., 2009). These physical, functional states, together with the animal's external circumstances, can be assessed through observable indicators in Domain 4 (Mellor et al., 2020). Mental state (Domain 5) reflects the animal's overall welfare or subjective experience in affective terms. Rather than being assessed independently, this domain serves as a reminder to derive affective interpretations from the states, conditions, and behaviours identified in Domains 1 through 4 (Mellor et al., 2020). In 2020, Domain 4 was renamed from "Behaviour" to "Behavioural Interactions" to emphasise that behaviour should be interpreted in relation to the animal's opportunities and interactions, rather than treated only as a list of observed behaviours (Mellor et al., 2020). Behaviour can therefore provide information on how animals respond to the conditions in which they are kept.

1.2.1 Welfare Assessing in Zoos

When animals are under human care, there is an obligation to provide comprehensive support for their well-being and to maximise their quality of life (Broom, 2002). According to the five domains models, limitations in nutrition, environment, or behaviour can influence both physical health and mental state. Therefore, animal management strategies should not only meet baseline requirements, but also actively promote rewarding experiences and minimise adverse ones across the individual's lifespan (Green & Mellor, 2011; Mellor, 2012).

1.2.1.1 Nutrition

Nutrition is one of the basic requirements of welfare. An animal's nutritional requirements should be assessed with consideration of its age, activity, body size, condition, reproductive state, physiological status, and overall health (Rees, 2011). In zoos, each species should have a feeding plan that provides adequate nutrition, promotes natural feeding behaviour, is reliably consumed by the animals, and remains practical to implement (EZNC, 2010). For captive carnivores, nutrition is not only related to food intake, but also to food presentation and opportunities to express natural feeding behaviour (Bond & Lindburg, 1990). Feeding enrichment can encourage captive felids to work for food and show feeding-related behaviours that are closer to natural feeding patterns (Bashaw et al., 2003). Therefore, feeding regimes can

influence both physical health and behavioural welfare.

1.2.1.2 Environment

Environment is also a key component of welfare assessment. When wild animals are confined, they face multiple restrictions. If the environment fails to meet the species-specific needs, their welfare may be compromised. Traditionally, zoo animal welfare was often assessed by the amount of space available (Estevez & Christman, 2006). In current welfare assessment, it is recognised that both the quantity and the quality of space play crucial roles in determining animal welfare. Analysing how animals use their space is important for identifying their biological needs, informing exhibit design, and enhancing welfare in zoos (Estevez & Christman, 2006). Retreat spaces are also important because they give animals places to rest while out of sight from visitors or other individuals. This can help reduce stress, social conflict, and the risk of injury (Rochlitz, 1999). In fact, research in felids has demonstrated that enclosure complexity/enrichment can be more important (within reason) than enclosure size (Moreira et al., 2007).

Behaviour and space use in captive environments can provide useful information about animals' needs, preferences, and psychological states (Mench, 1997). If an artificial environment does not provide suitable features, animals may experience discomfort and stress. These negative experiences can lead to physiological and behavioural changes and may reduce welfare (Würbel, 2001). In contrast, larger and more enriched enclosures can improve the wellbeing of many captive species (Mallapur et al., 2002). Space use is also influenced by enclosure design and social conditions. Physical structures, such as walls and barriers, can affect how animals move and where they spend time within an enclosure (Jensen et al., 2003). Social conditions may also affect movement patterns, as animals may need to move around other individuals during daily activities (Estevez & Christman, 2006). Other environmental features, such as shade, shelter and solar exposure, may also help explain where animals choose to rest or spend time.

1.2.1.3 Physical health

Physical health is another important domain in zoo welfare assessment. Zoological institutions are responsible for providing appropriate husbandry, preventive healthcare, and veterinary treatment for the animals in their care (Backues et al., 2011). Preventive healthcare,

such as routine veterinary health checks, can help identify health problems before they become obvious and support animal welfare (Kershaw et al., 2020). Physical health is also closely linked to behaviour, as pain, illness, or poor physical condition may change an animal's activity level, movement pattern, and interaction with the environment.

1.2.1.4 Behavioural Interactions

Behavioural interactions are another important part of the Five Domains Model. This domain considers how animals interact with their physical environment, other animals and humans, and whether they have opportunities to express species-appropriate behaviours (Mellor et al., 2020). Behaviour is also one of the most directly observable sources of welfare information in zoo animals. Changes in activity, resting, feeding, social behaviour, avoidance behaviour or abnormal repetitive behaviour can show how animals respond to the conditions in which they are kept (Manteca et al., 2016).

1.2.1.5 Mental State

Affective experiences in animals range from negative to positive, and overall welfare reflects the balance between them (Figure 1.2). Welfare is not a fixed state. A short negative experience does not necessarily indicate poor overall welfare if positive experiences are frequent and sustained, while repeated or prolonged negative experiences can shift welfare in a poorer direction (ZAA, 2025). Some positive affective experiences arise as adjuncts to behaviours originally motivated by negative internal states. For example, behaviours driven by hunger or thirst can result in pleasurable sensations during feeding or drinking, while behaviours triggered by thermal discomfort may produce rewarding experiences of warmth or cooling (Mellor & Beausoleil, 2015). Promoting good welfare in wild animals involves management that supports positive mental states and limits negative experiences, thereby minimising suffering across the animal's lifetime (Green & Mellor, 2011). Welfare is enhanced when animals experience emotions such as playfulness, contentment, happiness and a sense of security, which are often linked to good physical health together with appropriate and beneficial interactions with their environment and conspecifics (Mellor, 2012).

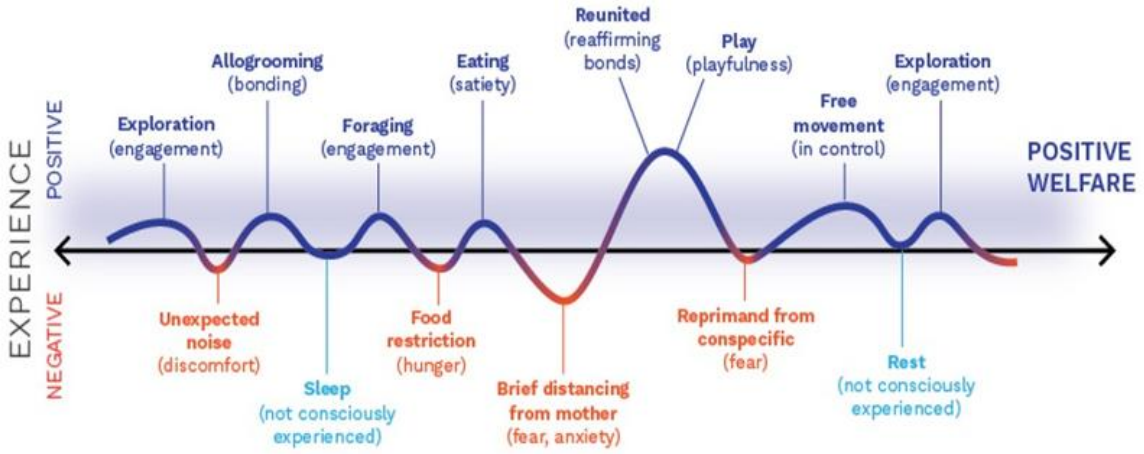


Figure 1. 2. Animals experience alternating positive and negative affective states over time, with overall welfare determined by the balance and frequency of positive experiences (ZAA, 2025).

1.3 Felid welfare in captivity

1.3.1 Evaluating welfare in captive felids

Felidae comprise at least 36 non-domestic species, ranging in size from the tiger (*Panthera tigris:* around 300 kg) to the rusty-spotted cat (*Prionailurus rubiginosus*; around 1 kg), and are distributed across all continents except Australia and Antarctica (Cuff et al., 2016; Jackson & Nowell, 1996). Many felid species are now threatened by human population growth, expanding settlements, habitat disturbance, prey loss, and direct persecution. As a result, some species are at risk of extinction, while others are experiencing population declines (Jackson & Nowell, 1996). Because of these threats, captive populations have become increasingly important as sites for insurance populations, research, and education (Law et al., 1997). However, as the need for captive felid populations grows, so does the need to establish tools for evaluating and enhancing their welfare in these environments.

Behaviour is one of the most used indicators in welfare assessment because it indicates how animals respond to their environment. Based on Broom's (1986) definition of animal welfare, behavioural observations can help assess whether animals are coping well with their conditions. Observation can also show whether certain behaviours are expressed too often, too rarely, or in abnormal ways. In captive settings, behavioural monitoring can help identify signs of stress, discomfort, frustration, or reduced behavioural opportunity. It can also indicate

whether animals are able to express behaviours that are normal and important for their species. When captive environments do not provide enough stimulation, animals may have fewer opportunities or less motivation to perform species-typical behaviours (Boere, 2001; McPhee, 2002). Under these conditions, individuals may develop abnormal behaviours not observed in the wild, including increased inactivity, excessive aggression, poor social interaction, self-injury, and stereotypies (Boere, 2001).

A stereotypy is defined as a repetitive and invariant behaviour pattern that appears to serve no clear purpose or function (Mason, 1991). Such behaviours are frequently linked to past or present environmental deficiencies and are commonly used as indicators of compromised animal welfare (Mason, 1991). In felids, stereotypic behaviours most commonly take the form of repetitive pacing, head weaving and over grooming (Lyons et al., 1997). Those behaviours are rarely observed in free ranging individuals. Pacing is the most prevalent stereotypy among carnivores and has been reported as accounting for the majority of stereotypic behaviour in this group (Clubb & Mason, 2003). Captive conditions can limit an animal's need or opportunity to engage in the behaviours that are critical to success in its natural habitat (e.g., hunting or territorial marking).

An enclosure's environment and set-up include factors such as its type, size, placement, and design, together with the way animals are managed within it. When these elements are poorly planned or maintained, stereotypic behaviours may develop as a result of multiple contributing factors (Lowe, 2013). Limited retreat space can also affect welfare, with the inability to withdraw, avoid, or escape from stressful situations being associated with distress and behavioural frustration in several species (Morgan & Tromborg, 2007). This is particularly relevant in zoo environments, where animals may be exposed to visitors, keepers, conspecifics, and routine husbandry activities. Limited space can also contribute to stereotypic behaviour because animals may be unable to obtain sufficient exercise or express natural hunting and exploratory behaviours (Gregory, 2004).

In addition to the assessment of abnormal behaviours, welfare evaluation in zoo housed felids also considers the extent to which individuals can express species specific behaviours. In the wild, felids spend a large part of their active time on behaviours that are important for survival, such as hunting, territorial marking, and exploring their environment (Lindburg,

1988). In captivity, the expression of these behaviours is often used as an indicator of behavioural and psychological wellbeing (Miller et al., 2020). These behaviours require both physical activity and mental engagement. If zoo environments limit these behaviours, this may indicate reduced welfare.

Territorial marking behaviours, such as scratching and urine spraying, are an important part of the felid behavioural repertoire. These behaviours are used for communication, spatial organisation, and social signalling (Cederlund, 2018). A reduction or absence of marking activity may indicate behavioural inhibition or suboptimal environmental conditions. Welfare problems may also be reflected in changes to natural activity rhythms, such as prolonged inactivity or excessive resting during normally active periods. This is particularly relevant for species whose natural lifestyles may be difficult to fully support in captivity (Clubb & Mason, 2003).

The way felids move through and occupy different parts of an enclosure can provide useful information about how they respond to their surroundings. Enclosure use may reflect the effects of enclosure design, structural complexity, husbandry conditions, and access to retreat areas (Mallapur & Chellam, 2002). Captive felids may show different patterns of enclosure use under different environmental conditions. Uneven use of available space, repeated avoidance of certain areas, or strong preference for only a small part of the enclosure may indicate suboptimal enclosure conditions or limited opportunities for normal behavioural expression (Mallapur et al., 2002). In captive Indian leopards (*Panthera pardus*), higher activity was associated with greater structural complexity, whereas stereotypic pacing was more common when opportunities for normal activity were more restricted (Mallapur et al., 2002).

Activity patterns and time budgets can also provide valuable insight into welfare in captive felids. The activity pattern of a species is shaped by its environment and can help explain its behaviour (Beltrán & Delibes, 1994). Studying these patterns can also support better captive management (Kleiman, 1994). In wild populations, felids show species-specific distributions of resting, locomotion, hunting related behaviour, exploration, and other daily activities (Lindburg, 1988). Captive felids may retain some of these natural activity patterns, but changes can also occur. For example, captive ocelots (*Leopardus pardalis*) showed a daily activity pattern that was broadly similar to wild ocelot but they were less active overall, more active

during the day, and showed stereotypic pacing (Weller & Bennett, 2001). Therefore, examining activity patterns and time budgets can provide a broader understanding of how felids use their time and environment.

Beyond the evaluation of individual behaviours, behavioural diversity has been proposed as a useful measure of welfare (Miller et al., 2020). Behavioural diversity refers to the range of species-typical behaviours expressed by an animal over a given period, rather than the frequency of a single behaviour (Miller et al., 2020). Approaches to welfare assessment have also emphasised the importance of identifying positive behavioural indicators, rather than focusing only on the absence of negative behaviours (Miller et al., 2020). In cheetahs, individuals with higher behavioural diversity had lower faecal glucocorticoid metabolite levels, which suggests a link between greater behavioural diversity and lower physiological stress (Miller et al., 2016).

1.3.2 Factors affecting felid welfare

1.3.2.1 Enrichment

Addressing welfare challenges in captive felids requires strategies that reduce stereotypic, undesirable, or overly aggressive behaviours while encouraging activity and species typical behaviours (Bashaw et al., 2003). Environmental enrichment is one important approach. Enrichment can help reduce boredom and stress in captivity and provide animals with more opportunities to express natural behaviours (Shepherdson et al., 1993). In felids, enrichment may include changes to enclosure design, feeding methods, sensory stimulation, and the introduction of novel objects (Mellen & Shepherson, 1997).

Feeding enrichment provides a clear example of how welfare can be improved in captivity. In the wild, animals spend a large part of their active time searching and consuming food (Herbers, 1981). In large felids, feeding typically consists of four elements: locating prey, capturing it, performing the kill, and behaviour at the carcass. Each stage requires physical effort and involves specialised adaptations, using senses, limbs, claws, teeth, and jaws (Lindburg, 1988). By contrast, many captive species are often given prepared diets, which reduces the need for natural foraging and hunting-related behaviours. This may limit appetitive behaviours linked to food acquisition and may contribute to abnormal behaviours (Lindburg, 1988). Felids also possess a strong motivation to perform predatory behaviours, regardless of

hunger. When opportunities are restricted, abnormal patterns can emerge. For example, captive felids are frequently observed pacing repetitively near enclosure areas where they can see the keeper approaching before feeding (Shepherdson et al., 1993). Therefore, food and water should be provided in ways that better reflect the species' natural feeding patterns and motivations (Rees, 2011).

Bashaw et al. (2003) investigated feeding enrichment in African lions (*Panthera leo*) and Sumatran tigers (*Panthera tigris sumatrae*) by providing horse bones and live fish. They found that the provision of bones increased consumptive behaviours such as chewing and gnawing, while live fish stimulated hunting-related behaviours, including crouching, pouncing and swiping, and reduced stereotypic behaviour in tigers from 60% to 30%. Similarly, providing intact calf carcasses to large felids stimulated feeding and other natural behaviours while reducing stereotypic behaviours in off-exhibit holding areas (McPhee, 2002). Carcass feeding requires animals to tear, chew, gnaw and manipulate food, which can extend feeding time and encourage more natural feeding behaviour. In cheetahs, carcass feeding has also been linked with improvements in oral health and psychological well-being, showing that naturalistic feeding programmes may provide species-relevant welfare benefits (Bond & Lindburg, 1990).

Structural enrichment also plays an important role. Modifications to enclosures can provide opportunities for climbing, resting, or hiding, thereby stimulating activity and reducing abnormal behaviours. Mallapur et al. (2002) reported that leopards (*Panthera pardus*) housed in enriched enclosures containing logs, trees, and sleeping platforms displayed more activity compared with those in barren environments. Designing biologically appropriate environments therefore requires knowledge of species' natural habitats and behavioural ecology.

Sensory enrichment further enhances welfare by stimulating felids through novel objects. The introduction of olfactory stimuli into enclosures can help to relieve boredom (Powell, 1995). Skibiel et al. (2007) showed that spices increased activity and reduced stereotypic pacing in several species, with particularly strong effects in ocelots (*Leopardus pardalis*). Other forms of enrichment can also encourage natural behaviours. For example, an "acoustic prey" system for an African leopard (*Panthera pardus pardus*), allowed the animal to trigger bird sounds and receive food rewards. This system increased locomotion and predatory behaviours and reduced stereotypic pacing (Markowitz et al., 1995). However, no single enrichment

strategy is likely to remain effective in the long term, because felids can quickly habituate to stimuli such as toys, scents, or structural changes (Mellen, 1998). Consequently, enrichment programmes should be varied and regularly reviewed to maintain their effectiveness.

1.3.2.2 Visitors

The visitor effect is another important aspect in felid welfare assessment. Modern zoos have important roles in conservation, research, and education, and visitor income helps support these functions (Reade & Waran, 1996). For example, captive Indian leopards (*Panthera pardus fusca*) spent approximately 12% of their time active when visitors were present, compared with 54% when visitors were absent (Mallapur & Chellam, 2002). Similarly in jaguars (*Panthera onca*), higher visitor density was associated with increased stereotypic, aggressive, and hiding behaviours (Sellinger & Ha, 2005). A broader study found that several felid species showed reduced activity and avoidance of visitor facing areas when zoos were open. Interestingly, jaguars (*Panthera onca*) in the same study displayed greater activity and occasional interactions with visitors (Suárez et al., 2017). Thus, while visitor presence may sometimes act as enrichment, it can also cause stress and negatively affect animal welfare (Chamove et al., 1988; Davey, 2005).

The fact that some individual animals' respond adversely to high visitor numbers emphasises the need for the availability of hiding or retreat areas. In the leopard study, retreat spaces were not mentioned (Mallapur & Chellam, 2002), while only small hiding places were reported in the jaguar study (Sellinger & Ha, 2005). This lack of adequate hiding opportunities likely intensified negative visitor effects. Evidence from domestic cats (*Felis catus*) shows that hiding is a common strategy used by felids to cope with environmental change and avoid unwanted interactions with other animals or humans (Rochlitz, 1999).

1.3.2.3 Keepers

Alongside visitors, keepers are another important source of human influence on felids in zoos. Zoo animals usually interact with their keepers more often than with other humans. Keepers are involved in daily care, including feeding, cleaning, health monitoring, and enrichment. Therefore, the way felids respond to keeper interactions can have an important effect on their welfare.

Mellen (1991) first reported that felids whose keepers spent more time interacting with

them, such as by soliciting contact or speaking to them, showed higher reproductive success. Later research further supported these findings, with felids showing less pacing when keepers spent more time interacting with them (Mellen, 1998). In cheetahs, public exposure and the number of caregivers have also been linked with reproductive and behavioural outcomes, with male cheetahs housed off-display or managed by fewer caregivers showing higher motile sperm output and higher overall behavioural activity (Koester et al., 2015).

1.3.2.4 Social management

Social management can also influence welfare in captive felids, although its effects vary across species and housing conditions. Most felid species are solitary in the wild, although social behaviour can vary between species, sex and social context (Stanton et al., 2015). In captive tigers, social housing conditions were associated with differences in behavioural expression. Paired tigers showed a wider range of behaviours than singly housed tigers, and pair housed tigers without neighbouring tigers showed less pacing than those housed next to other tigers (De Rouck et al., 2005). In cheetahs, group membership and visual contact with neighbouring cheetahs have also been identified as factors influencing stereotypic behaviour, and solitary housing was associated with higher levels of stereotypy (Quirke et al., 2012). These observations are consistent with evidence that male cheetahs naturally form stable coalitions and can maintain proximity, frequent affiliative behaviour, and low levels of overt aggression under captive conditions (Chadwick et al., 2013) .

1.4 Behavioural monitoring in captive felids

As mentioned previously, behavioural monitoring is an important part of welfare assessment in captive felids because behavioural indicators are among the most widely used tools in zoo animal welfare assessment (Binding et al., 2020). Behavioural monitoring can include the recording of abnormal repetitive behaviours, especially stereotypic behaviours, because these behaviours are commonly used as indicators of compromised welfare (Mason, 1991). It can also include normal activity patterns and species-specific behaviours, which provide information on how animals use their environment and respond to management conditions. Behavioural observations are non-invasive, usually do not require highly

specialised equipment, and can often be applied at relatively low cost (Salas et al., 2024). In behavioural studies, observations may be conducted at either the individual or group level, depending on the research question. The most common methods of behavioural observation used in zoos are all-occurrence, interval, and continuous sampling (Altmann, 1974). Different methods are suited to different types of behaviour. Brief behaviours may be recorded using all-occurrence sampling, whereas longer behaviours may be more appropriately assessed using interval or continuous recording (Salas et al., 2024).

Behavioural monitoring also requires the development of an ethogram. Ethograms are fundamental to behavioural research because they allow observed behaviours to be documented and measured accurately. Differences in behavioural definitions across studies can make comparisons difficult, whereas standardised terminology can improve consistency and inter-observer reliability (Stanton et al., 2015). This is especially important in studies of captive felids, where welfare assessment often depends on distinguishing between behaviours such as resting, locomotion, pacing, marking, and social interactions. Stanton et al. (2015) developed a standardised ethogram for the family Felidae to provide clear and consistent behavioural definitions for felid research.

Direct behavioural observations, such as all-occurrence, interval and continuous sampling, also have some limitations. The importance of a behaviour is not always determined only by its frequency or duration. In some cases, intensity may also be important, especially for behaviours such as aggression (Manteca et al., 2016). Behavioural assessment may also be affected by observer bias, time of observation, enclosure conditions, and other external factors. In addition, direct observation is usually limited to selected observation periods. This may make it difficult to record behaviours that are rare, very brief, or hidden from the observer's view (Salas et al., 2024).

These limitations are especially important in captive felids. Many felids spend long periods resting, may move through large or complex enclosures, and may show behavioural changes across the day. These patterns can be difficult to capture consistently through scheduled observations alone. Longitudinal observations can reveal how changes in environment, social group, husbandry, or health influence behaviour and welfare. Therefore, reliable behavioural monitoring is essential if welfare assessment is to support management decisions in a

systematic way (Salas et al., 2024). More broadly, major knowledge gaps remain in zoo animal welfare research, which further supports the need for evidence-based zoo animal management and more continuous, standardised, and objective behavioural data collected over extended periods (Melfi, 2009).

Challenges with visual behavioural observations have driven the development of monitoring methods that extend beyond traditional direct observation. Technologies such as video recording, GPS tracking, magnetometers, and accelerometers can provide more continuous information on animal movement, posture, and activity patterns (Brown et al., 2013; Grünewälder et al., 2012; McGowan et al., 2022). Among these approaches, inertial sensors (primarily accelerometery) are particularly relevant for behavioural monitoring because they can record movement data over extended periods and can be paired with video observations to allow behavioural validation.

1.5 The use of triaxial accelerometers in animal behaviour research

Accelerometer devices can be classified according to the data they collect. Monoaxial accelerometers collect data along a single axis, normally the cranio-caudal axis of the animal (Brown et al., 2013). While these devices are useful in some cases, they do not effectively capture the three-dimensional movement of felids. Omnidirectional accelerometers provide a better understanding of an animals overall three-dimensional movement; however, they record this three-dimensional movement as a single datapoint, limiting the results to the assessment of overall activity. For this reason, triaxial accelerometers now dominate the field of wearable technologies.

Triaxial accelerometers are small motion sensors that measure acceleration along three independent orthogonal body axes: craniocaudal/surge (forwards and backwards); mediolateral/sway (left and right); and dorsoventral/heave (up and down) (Brown et al., 2013). In animal behaviour research, these devices are increasingly used because they can record body movement continuously and provide information on both activity and posture (through the effects of gravity on each independent axis) (McClune et al., 2014; McGowan et al., 2022; Redmond et al., 2024; Smit et al., 2023). This makes them useful for studying behaviours that

may be difficult to capture consistently through direct observation alone. Research in non-domestic felids, however, is limited.

In cheetahs, movement data have been used to understand locomotion during hunting in wild individuals (Wilson & Lowe et al., 2013). Accelerometer-based movement data have also been used to classify broader behavioural states in free-ranging cheetahs, including feeding, mobile and stationary behaviour (Grünewälder et al., 2012). When devices are attached to animals, it is also important to consider whether collar-mounted equipment may affect behaviour or movement, especially when devices are relatively large or heavy for the animal (Coughlin & Van Heezik, 2015). This issue may decrease as devices become smaller and lighter.

Alone, raw triaxial acceleration data provided limited information outside of quantifying overall activity or overall dynamic body acceleration (ODBA). Thus, machine learning is frequently used to convert raw triaxial acceleration data into behavioural classifications, which are more valuable from the animal welfare perspective. To do this, triaxial accelerometer data are often combined with concurrent video-labelled behavioural observations according to a defined ethogram. The corresponding accelerometer data are then matched with these behavioural labels for the training and validation on machine learning models (McGowan et al., 2022; Redmond et al., 2024; Smit et al., 2023).

Machine learning methods are useful for analysing accelerometer data because they can identify complex patterns in movement signals and classify behaviours based on multiple features (Valletta et al., 2017). Raw triaxial acceleration data are usually converted into behavioural categories using a trained classification model. Random forests (RF) are machine learning algorithms that combine predictions from multiple decision trees to produce a final classification through majority voting. They are widely used because they can handle complex datasets and are flexible and relatively simple to implement (Breiman, 2001). In accelerometer-based behavioural studies, model performance can be affected by the quality of the labelled training data, the selection of relevant features, and the balance between behavioural classes (Redmond et al., 2024; Smit et al., 2023). Behavioural data are often imbalanced because some behaviours, such as resting or standing, occur more frequently than short-duration behaviours such as feeding, drinking or grooming (McGowan et al., 2022; Redmond et al., 2024). This can make rare behaviours more difficult to classify and may require behaviours to be removed,

grouped or redefined during model development (Redmond et al., 2024). Feature selection can also improve model performance by reducing irrelevant or low-quality variables (Cadenas et al., 2013). Therefore, model development needs to balance behavioural detail with classification performance, especially when several behaviours have similar movement patterns or low sample sizes.

For captive cheetahs, accelerometer-based behavioural classification may provide a useful method for generating more continuous and objective behavioural data. Once a reliable model has been developed and validated, the classified behaviours can be applied to examine enclosure use, behavioural distribution across zones, and time budgets. This creates a direct link between model development and welfare related behavioural assessment.

1.6 Study site and species

This study focused on two adult female cheetahs, Qia and Quartz, housed at Auckland Zoo, Auckland, New Zealand. At the start of the study, both cheetahs were approximately eight years old, with bodyweights of approximately 38 and 40 kg, respectively. Qia and Quartz were born at Cango Wildlife Ranch in South Africa and have been housed at Auckland Zoo since November 2017. These individuals were suitable for this study because they were already habituated to routine keeper management and regularly participated in keeper-led walks. As part of this routine, they were accustomed to wearing collars and showed no obvious avoidance of collar use. This made them appropriate candidates for collar-based behavioural monitoring. During video scoring, Qia and Quartz were distinguished using visible tail markings. Qia was identified by the larger white marking near the distal end of the tail, whereas Quartz was identified by the smaller white marking in the same area (Figure 1.3). The cheetah enclosure is shown in Figure 1.4, with the indoor habitat, public viewing area and main structural features labelled.

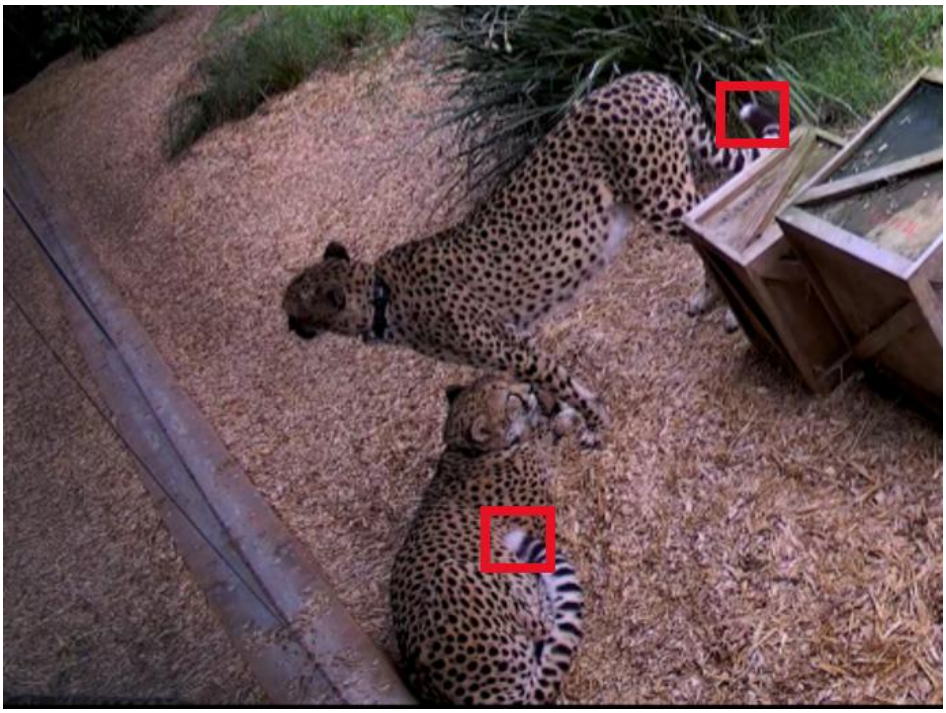


Figure 1. 3. Visible tail markings used to distinguish Qia and Quartz during video scoring. Qia had a larger white marking near the distal end of the tail, whereas Quartz had a smaller white marking in the same area.

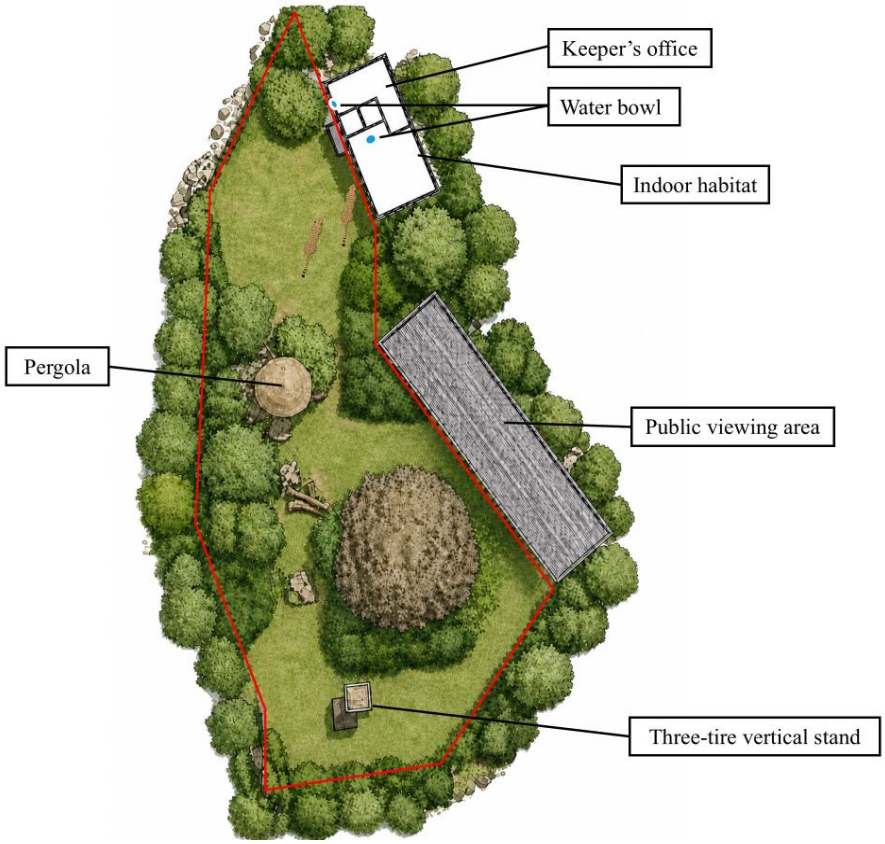


Figure 1. 4. Overview of the cheetah enclosure at Auckland Zoo, showing the indoor habitat, public viewing area and main enclosure features. The red line indicates the main outdoor enclosure boundary.

1.7 Thesis aim and objectives

Aim: This thesis aimed to demonstrate how wearable technologies and machine learning can be used to monitor behaviour and support welfare assessment in captive cheetahs in a zoo environment. The study used two adult female cheetahs at Auckland Zoo as an example of how collar-derived acceleration data can be linked with video observations to assess behaviour, time budgets and enclosure use.

Objectives:

Objective 1: To develop a machine learning model to classify cheetah behaviour from triaxial acceleration data collected using collar-mounted accelerometer devices;

Objective 2: The use of machine learning model, in combination with video footage, to evaluate the behavioural time budgets and enclosure utilisation of the cheetah at Auckland Zoo.

Chapter 2 – Development of machine learning models for classifying captive cheetah behaviour from collar-based triaxial accelerometer data

2.1 Introduction

Behavioural assessment is an important part of welfare evaluation in captive animals because behaviour provides information on how animals respond to their environment and management conditions (Manteca et al., 2016). In captive cheetahs, behavioural records can be used to describe daily activity patterns, resting behaviour, locomotion and feeding-related behaviour. Previous studies of zoo-housed cheetahs have shown that they spend a large proportion of time resting or sleeping, while activity patterns can vary between day and night and between housing conditions (Seyrling et al., 2022). To use these behaviours in welfare assessment, they need to be recorded and classified reliably. Direct behavioural observation remains valuable, but it can be time-consuming, labour-intensive and affected by observer bias (Redmond et al., 2024). It may also be limited when animals are out of sight, inactive for long periods, or active outside selected observation periods. These limitations make it useful to develop methods that can record behaviour over longer periods and provide more objective behavioural data.

Triaxial accelerometers provide a practical method for recording animal movement and posture over extended periods. Unlike direct behavioural observation, accelerometers can collect movement data without the need to have an observer present, when the animal is not visible and during periods of darkness (Brown et al., 2013). Many modern wearable accelerometers measure acceleration across three independent orthogonal body axes: craniocaudal or surge, mediolateral or sway, and dorsoventral or heave. Measuring acceleration across these axes allows both dynamic movement and static posture to be described, providing a basis for identifying behaviours that differ in movement pattern or body position (Brown et al., 2013).

To classify behaviours from accelerometer data, machine learning techniques are often used to build behavioural identification models. These methods are suited to large and high-

dimensional behavioural datasets where relationships among variables may be nonlinear or difficult to describe using traditional statistical approaches (Valletta et al., 2017). Of the machine learning models used, random forests have been most consistently applied and have been used to identify behaviours from accelerometer data in domestic cats, domestic dogs and cheetahs (McGowan et al., 2022; Redmond et al., 2024; Smit et al., 2023).

The level of behavioural resolution is an important consideration when developing behaviour identification models from accelerometer data. Fine scale categories may provide greater behavioural detail, whereas broader categories may improve classification consistency when behaviours have similar movement patterns or limited observations. In cheetahs, previous accelerometer research has compared fine, medium and coarse behavioural resolutions using collar mounted devices and random forest models (McGowan et al., 2022). Unlike previous work based on short exercise bouts, the present study focused on behaviours occurring during normal enclosure use. Therefore, this study developed a behaviour identification model based on the daily behaviour of zoo-housed cheetahs, providing a behaviour structure that is more directly relevant to routine welfare monitoring in a captive zoo setting.

The aim of the study was to develop machine learning algorithms for the identification of captive cheetah behaviour from triaxial acceleration data. Behavioural classification models with different levels of complexity were developed and evaluated to compare classification performance across behavioural structures.

2.2 Materials and Methods

2.2.1 Study location and Animals

This study was conducted at Auckland Zoo, Auckland, New Zealand. Animal ethics approval was obtained from the Auckland Zoo Animal Ethics Committee. Two adult female cheetahs (*Acinonyx jubatus*) housed at Auckland Zoo were used in this study. The cheetahs were both approximately eight years of age and weighed approximately 38 and 40 kg at the start of the study. The animals remained in their usual indoor and outdoor habitats throughout the study and continued to be managed under their normal husbandry routines by regular zoo staff. All handling associated with collar fitting and removal was carried out by experienced

keepers familiar to the animals. Note, theses cheetahs were already accustomed to wearing normal collars for the purpose routine keeper-led walks used for enrichment.

2.2.2 Experimental design

Prior to the study, the cheetahs' exhibit was fitted with four Reolink Duo 2 surveillance cameras (Reolink Technology Co., Wilmington, Delaware, USA). These cameras provided a 180° panoramic view, 4K resolution at 20 FPS, and 30 m infrared night vision. The cameras were connected to a Reolink RLN36 36-channel NVR (Reolink, Hong Kong, China), and video footage was recorded continuously from 1 Apr 2025 to 9 May 2025. The camera system covered both the indoor and outdoor habitats, allowing the cheetahs to be monitored across the main areas used during the study (Figure 2.1).

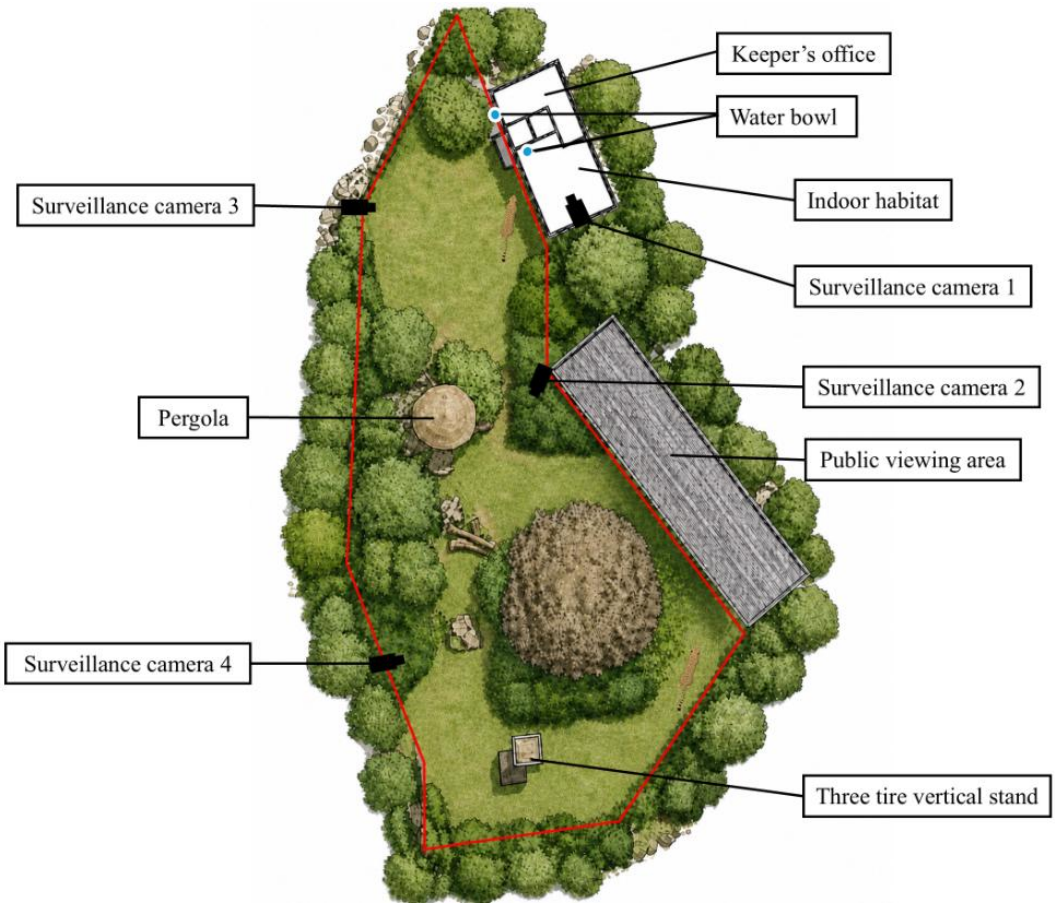


Figure 2. 1. Diagram of the indoor and outdoor habitats, showing the locations of the four surveillance cameras used to monitor the animals.

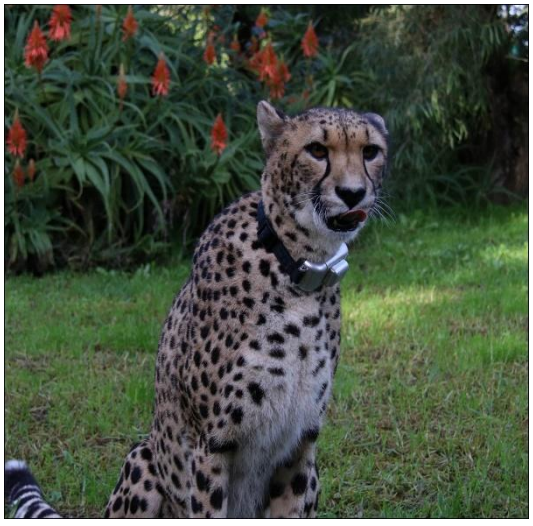
Each cheetah was then gradually habituated to a proprietary remote releasing collar (i.e.,

enabled remote collar removal via a key fob button with a range of ~15-20 m; Moritech® Engineering Ltd., Palmerston North, New Zealand) containing an ActiGraph® wGT3X-BT (ActiGraph LLC, Pensacola, Florida, USA) triaxial accelerometer unit (Figure 2.2 A), that continuously logged acceleration data along three independent axes (lateral, cranio-caudal, and dorso-ventral) at 30 Hz with a dynamic range of ±8 g (where g represents gravitational acceleration, approximately 9.81 m/s²). The full collar assembly weighed 331 g and measured 50 mm × 120 mm × 40 mm (Figure 2.2 B), which was equivalent to 0.85% of the cheetahs' bodyweight. When fitted, the collar was naturally position ventrally under the animals' chins (Figure 2.2 C).



(A)





(B) (C)

Figure 2. 2. (A) ActiGraph® wGT3X-BT unit. (B) ActiGraph® wGT3X-BT device set in the Collar. (C) Cheetah with collar.

Following collar attachment, the cheetahs were monitored closely by zoo staff to assess collar acceptance. Monitoring focused on potential signs of collar-related irritation or distress, including repeated scratching or rubbing at the collar, frequent head shaking, attempts to remove the collar, visible skin irritation around the neck, and changes in normal movement. No adverse reactions to the collars were observed during the study. Collar acclimatisation occurred in stages prior to the continuous data collection period. From 7 to 13 Apr 2025, the collars were worn for approximately 2 h per day in the morning while the exhibit was cleaned. From 14 to 27 Apr 2025, excluding the Easter weekend (20 and 21 Apr), the collars were worn for approximately 6 h per day while staff were present at the zoo. From 28 Apr to 3 May 2025, the collars were worn overnight from approximately 15:00 h to 9:00 h the following morning (18 h). After the acclimatisation period, the collars were fitted at approximately 10:30 h on 4 May 2025 and remained on the cheetahs until removal at 10:30 h on 10 May 2025. For the present analysis, matched accelerometer and video data from 00:00 h on 5 May 2025 to 23:59 h on 9 May 2025 were used because this period provided a complete continuous dataset suitable for behavioural classification. Therefore, all model training, validation, and behavioural analyses reported in this chapter were based on the dataset of matched accelerometer and video data from these five days.

2.2.3 Scoring behaviour from video footage

The observed behaviour of the cheetahs was scored continuously from the recorded video footage at 1 s intervals using Boris® (Friard & Gamba, 2016). Twenty-six behaviours were included in the analysis (Table 2.1). For model development, these behaviours were organised into three broad classes: active, inactive, and maintenance. The “Active” class comprised raking, stretching, shaking, scratching, scent marking, jumping/climbing, running, trotting, walking, rolling, and sniffing. The “Inactive” class included crouching, lying alert on the left side, lying alert on the right side, lying alert ventrally, lying asleep on the left side, lying asleep on the right side, sitting, and standing. “Maintenance” behaviours included allogrooming, grooming, drinking, eating, and toileting. In addition, an ‘Other’ category was used to account for behaviours not included in the ethogram, and an ‘Out of sight’ category was used when the cheetahs could not be seen clearly in the video footage.

Table 2. 1. A standardised ethogram for the Felidae including definitions for all base behaviours derived from Stanton et al. (2015).

Cat.	Behaviour	Definition
Active	Raking	Kicking motions with either one or both of their hindlimbs.
	Stretching	Extends its forelegs while curving its back inwards.
	Shake	Rotates its abdomen from side to side.
	Scratching	Scratches its body using the claws of its hind feet.
	Scent mark	While standing with tail raised vertically, cat releases a jet of urine backwards against a vertical surface or object. The tail may quiver as urine is discharged.
	Jump/ Climb	Ascends and/or descends an object or structure
	Run	Forward locomotion in a rapid gait, which is faster than walking or trotting.
	Trot	Forward locomotion at a swift gait performed with alternating steps. Movement is faster than walking but slower than running.
	Walk	Forward locomotion at a slow gait.
	Roll	While lying on the ground, rotates body from one side to another. During the roll, the back is rubbed against ground, the belly is exposed and all paws are in the air. May continue rolling repeatedly from side to side.
	Sniff	Smells (modifier) by inhaling air through the nose.
Inactive	Crouching	Alert and positions the body close to the ground, whereby all four legs are bent, and the belly is touching the ground.
	Lying alert left	Lying on the left side of their body, either extended or curled. Their head is raised off the ground.
	Lying alert right	Lying on the right side of their body, either extended or curled. Their head is raised off the ground.
	Lying alert ventral	Lying with their abdomen in contact with the ground. Their head is raised off the ground.
	Lying asleep left	Lying on the left side of their body, either extended or curled. Their head resting against (modifier). They are not easily disturbed.
	Lying asleep right	Lying on the right side of their body, either extended or curled. Their head resting against (modifier). They are not easily disturbed.
	Sit	In an upright position, with the hind legs flexed and resting on the ground, while front legs are extended and straight
	Stand	In an upright position and immobile, with all four paws on the ground and legs extended, supporting the body
Maintenance	Allogroom	Cats will briefly push and/or rub their head forcefully against another.
	Groom	Cleans itself by licking, scratching, biting or chewing the fur on its body. May also include the licking of a front paw and wiping over head.
	Drink	Ingests water (or other liquids) by lapping up with the tongue.
	Eat	Ingests food (or other edible substances) by means of chewing with the teeth and swallowing.
	Toileting	Release faeces and urine on the ground while in a squatting position.
	Out of sight	Not visible to the observer.
	Other	Any behaviour that does not fit into one of the behaviours included in this ethogram.

Abbreviation: Cat, Category.

2.2.4 Processing the acceleration data

At the end of the study, raw acceleration data from each axis were downloaded using ActiLife 6® software (ActiGraph, Pensacola, FL, USA) and exported as ".csv" files for subsequent processing. The acceleration data were then processed using R statistical software version 4.5.2 (R Foundation for Statistical Computing, Vienna, Austria). From the raw accelerometer data, a total of 45 features were derived and summarised into 1 s epochs. These included descriptive statistics for each axis and vector magnitude (n = 28), derived kinematic features (n = 3), gravitational/postural features (n = 6), inter-axis correlations (n = 3), and rolling-window features derived from ODBA, vector magnitude, roll and pitch (n = 5). A detailed description of all calculated features is provided in Appendix 1. The 1 s epoch feature dataset was then used for the machine learning process.

2.2.5 Machine learning using random forests

2.2.5.1 Training

Initially, the calculated feature data set (1s epoch) from the acceleration data and scored observed behaviour dataset (1s epoch) from Boris® were time synchronised to within 1-2 seconds. This was achieved by using a clearly identifiable reference event in which each collar was placed on the ground and remained static for a short period. The static period lasted approximately five seconds for Qia (Figure 2.3 A) and 15 seconds for Quartz (Figure 2.3 B). This event was visible in the video footage and was also identified in the triaxial acceleration profiles as a stable period with minimal variation in acceleration. The timing of this static-collar period was compared between the video footage and the accelerometer record, and the difference between the two records was used to adjust the accelerometer timestamps. Based on this procedure, the accelerometer data were adjusted by +55 s for Qia and +39 s for Quartz, allowing the acceleration-derived feature dataset to be aligned with the Boris® behavioural dataset.

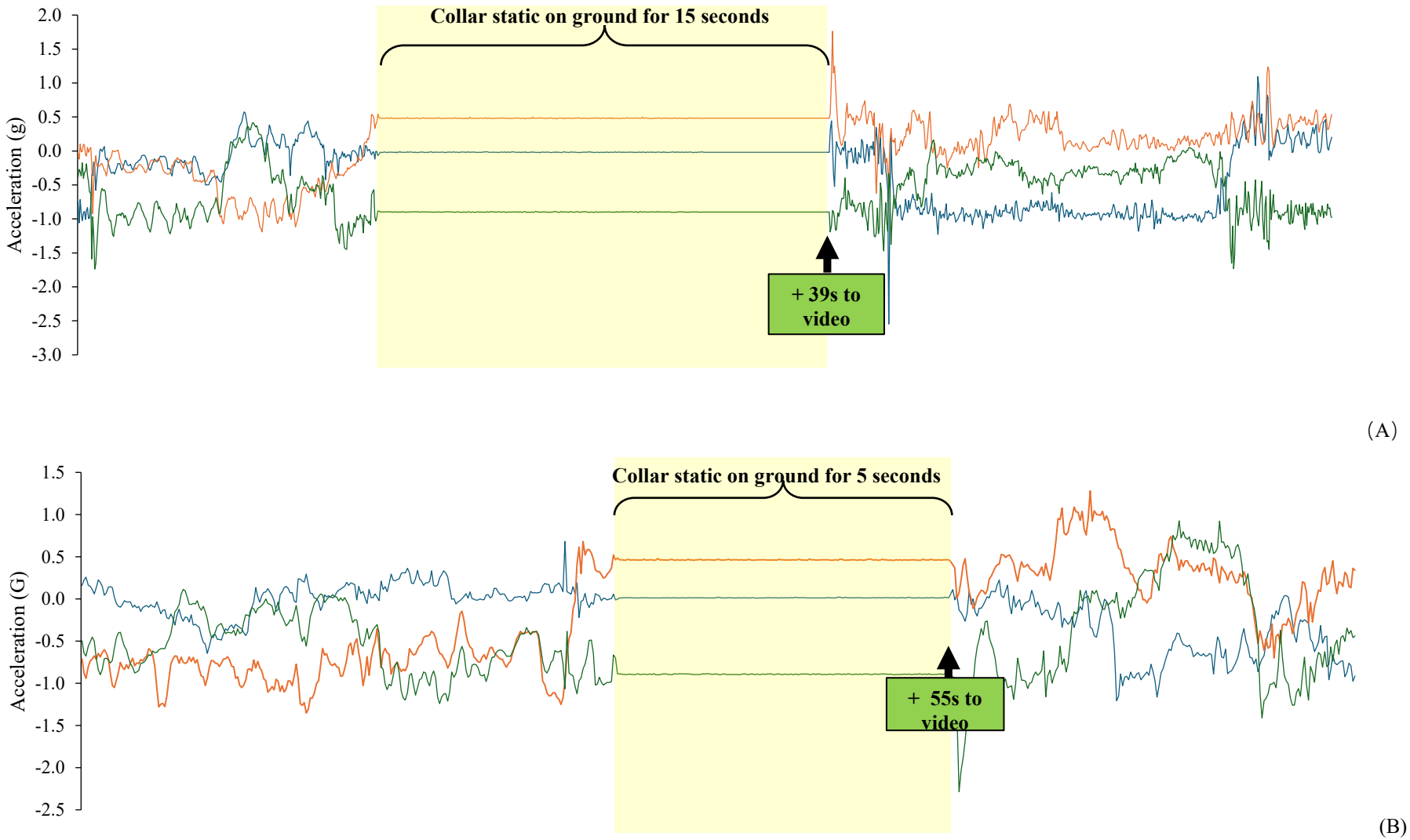


Figure 2. 3. Synchronisation of accelerometer (30 Hz) and video data. The highlighted sections indicate the period when the collar was static on the ground, which was used as a reference point for aligning the accelerometer data with the video footage. Based on this reference event, the accelerometer data were adjusted by +39 s for Quartz (A) and +55 s for Qia (B) to synchronise the collar-derived data with video records. Note: Timestamps are not shown on the x-axis because the raw data were sampled at 30 Hz, resulting in a very fine time scale

Before model development, behaviour categories that were not suitable for reliable classification were excluded from the dataset, including “out of sight” and “other”. Records that could not be observed clearly or could not be assigned to a specific behavioural category were excluded to ensure that the models were trained using consistently annotated behavioural records. Thus, the final dataset consisted of 223,909 s of concurrent labelled behaviour and 1 s feature data.

Random forest (RF) models were trained to classify cheetah behaviour from the feature data derived from the triaxial acceleration outputs, following previous accelerometer-based behavioural classification approaches (Redmond et al., 2024; Smit et al., 2023). Each model consisted of multiple decision trees, with the final behavioural classification determined by majority voting across the trees (Breiman, 2001). The models were built in R using the packages ‘caret’ and ‘randomForest’, with the default setting of 500 trees. The complete dataset was divided into training and validation datasets using a random stratified split, with 70% of the data used for model training and the remaining 30% retained as a test dataset. A full external validation (e.g., a leave-one-animal-out approach) was not feasible because the dataset included only two animals. Although such an approach would provide a more rigorous assessment of model generalizability and is recommended for larger datasets, the limited number of animals would have resulted in an unreliable evaluation. However, additional validation was conducted in Chapter 3 by evaluating model predictions generated from completely unseen accelerometer data against independently and visually scored behavioural time budgets.

Using the training dataset, a total of six RF models were developed with varying levels of behavioural complexity (Figure 2.4). Behaviours were progressively merged or removed in subsequent models based on initial model performance and the available sample size for each behavioural class. Model 1 retained the full set of behavioural classes after this initial filtering step. Model 2 removed several very weak or rarely observed classes. Model 3 reduced behavioural complexity by combining locomotive behaviours into a single locomotion class. Model 4 merged a broader set of active and grooming related classes while retaining ‘drink’, ‘eat’, ‘sit’, ‘stand’, and the two lying classes as separate outputs. Model 5 further simplified the rest related classes into ‘resting alert’ and retained ‘drink’ as an independent behaviour. At this stage, feature selection (FS) and down sampling (DS) were applied to improve model performance. Model 6 also used FS and DS and grouped behaviours into the three broad categories ‘active’, ‘inactive’, and ‘maintenance’ (Figure 2.4).

For Models 5 and 6, feature selection and down sampling were also applied to the training

dataset using the varImp function in the "caret" package in R to aid optimisation of model performance. Feature selection was achieved by retaining features that together explained 80% of the cumulative variable importance based on variance analysis (Cadenas et al., 2013). Down sampling was applied to address class imbalance by reducing over-represented classes within the training data to a maximum of 7,000 s, as previously done for domestic cats (Smit et al., 2023).

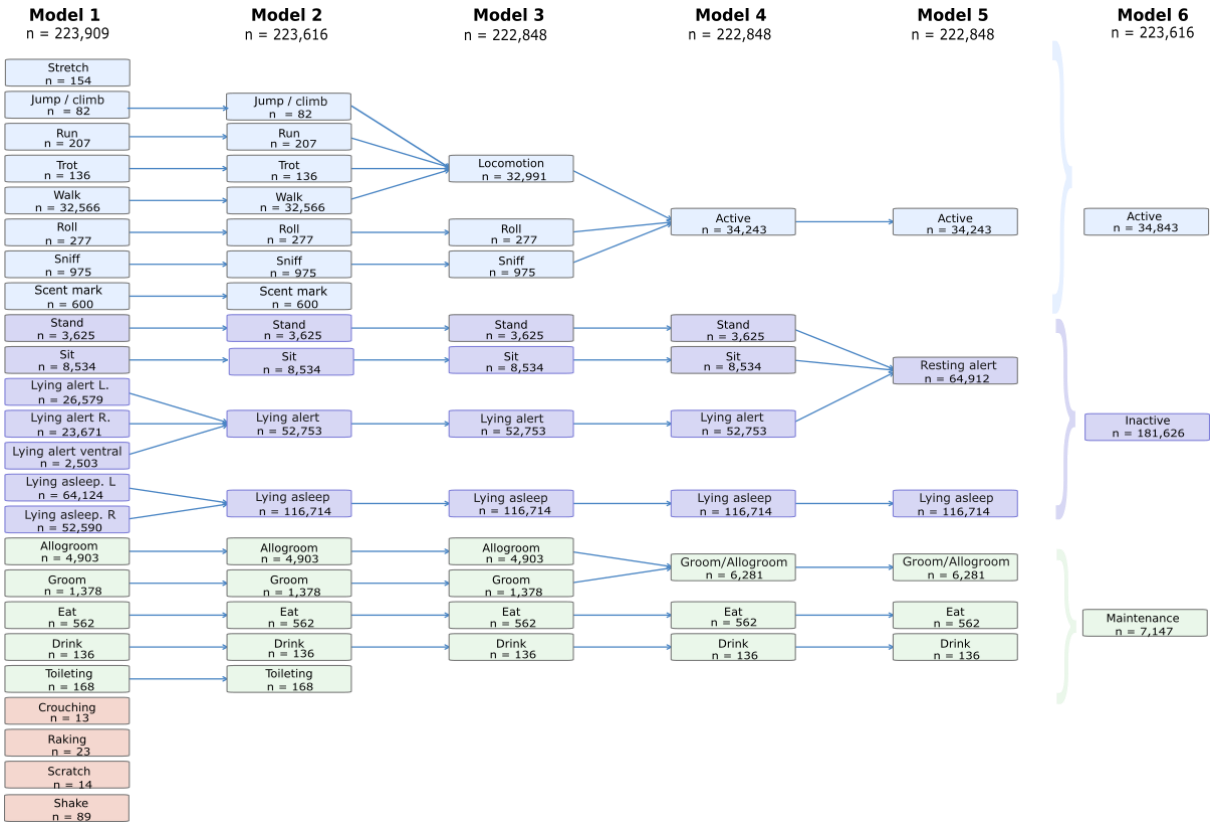


Figure 2. 4. Number of training observations for each behavioural category included across the six RF model development stages. After excluding the other and out of sight categories, the full dataset consisted of 223,909 observations with 70% (156,736 observations) used for training and 30% (67,173 observations) for test. The last four behaviours listed in Model 1 were removed from subsequent models due to limited sample size and/or poor model performance. Abbreviation: L, left; R, right.

2.2.5.2 Evaluating model performance

Model performance was evaluated on the unseen test data set (30% of the data). The performance of the six RF models was evaluated using confusion matrices generated by comparing predicted behaviours with observed behaviours in the test dataset. For each behavioural class, the confusion matrix was used to derive the numbers of true positives (TP; behaviours correctly identified by the model), true negatives (TN; behaviours not observed and correctly identified by the model as absent), false positives (FP; behaviours predicted by the model but not observed), and false negatives (FN; behaviours observed but not identified by

the model). These values were then used to calculate specific behaviour classification statistics included sensitivity, specificity, precision, negative predictive value, balanced accuracy, and F1 score (Table 2.2). Sensitivity was used to assess the ability of the model to correctly identify a behaviour when it occurred, whereas specificity described the ability of the model to correctly identify when that behaviour was absent. Precision reflected the proportion of positive predictions that were correctly classified, while negative predictive value represented the proportion of negative predictions that were correctly classified. Balanced accuracy was also calculated to account for differences in class size by averaging sensitivity and specificity. In addition, F1 score was used as a combined measure of precision and sensitivity and was considered particularly informative where class imbalance was present. Overall model performance was assessed using overall accuracy and Cohen's kappa coefficient. Overall accuracy represented the proportion of all validation observations that were correctly classified by the model, whereas Cohen's kappa coefficient quantified the level of agreement between observed and predicted classifications after accounting for agreement expected by chance (Table 2.2). In general, values greater than 0.70 for sensitivity, precision, balanced accuracy, and F1 score were considered indicative of acceptable classification performance, with higher values reflecting stronger model performance (Smit et al., 2023). In addition to the standard metrics, the coefficient of variance (CV%) between the observed and predicted prevalence of each behaviour class was also calculated, as this reflects how the outputs were used in an applied context (e.g., time budget).

Table 2. 2. Calculations for the parameters used to evaluate model performance.

Parameter	Calculation
Sensitivity / recall	TP / (TP + FN)
Specificity	TN / (TN + FP)
Balanced accuracy	(Sensitivity + Specificity) / 2
Precision	TP / (TP + FP)
Negative predictive value	TN / (TN + FN)
Precision Recall (F1 score)	2 × (Precision × Sensitivity) / (Precision + Sensitivity)
Observed prevalence	(TP + FN) / (TP + TN + FP + FN)
Detected prevalence	(TP + FP) / (TP + TN + FP + FN)
Detection rate	TP / (TP + TN + FP + FN)
Overall accuracy	(TP + TN) / (TP + TN + FP + FN)
Kappa coefficient (κ)	$\kappa = (N \times \Sigma_{i=1}^{k} x_{ii} - \Sigma_{i=1}^{k} (x_{i+} \times x_{+i})) / (N^2 - \Sigma_{i=1}^{k} (x_{i+} \times x_{+i}))$ Where: N = total number of observations across all behavioural categories k = number of behavioural categories i = behavioural category i xii = number of observations assigned to category i by both the visual observations and the model xi+ = number of observations assigned to category i by the visual observations x+i = number of observations assigned to category i by the model

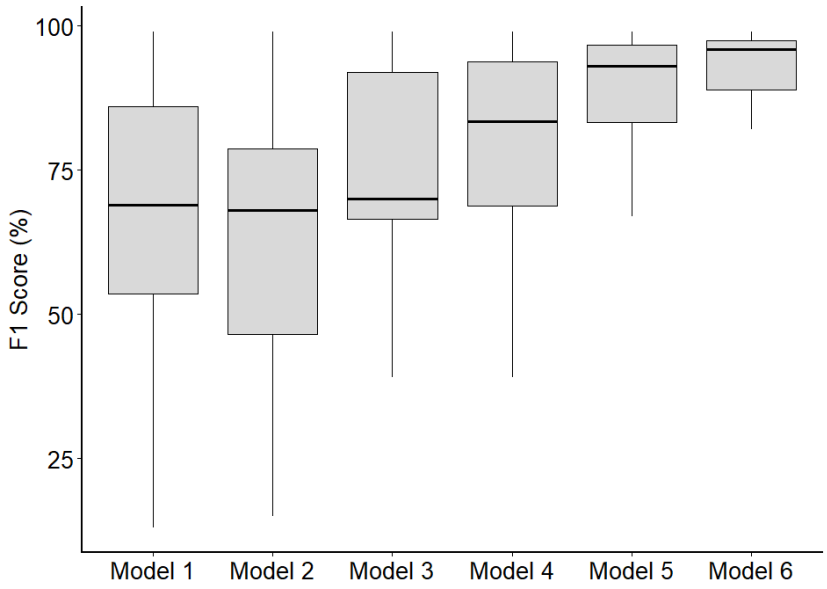
Abbreviations: TP, true positive; TN, true negative; FP, false positive; FN, false negative.

2.3 Result

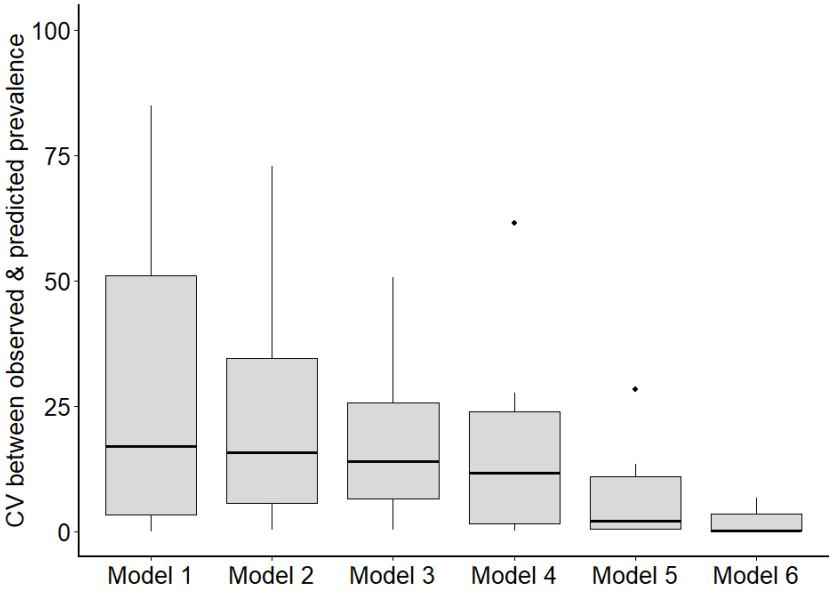
The full matched accelerometer and video dataset initially contained 864,000 observations. After records labelled as “Other”, “Not relevant” (i.e., when collar was off) and “Out of Sight” were excluded, Model 1 contained 223,909 observations across 24 behavioural classes. The number of observations and behavioural classes differed across the six modelling rounds because each model used a different behavioural structure (See Figure 2.4).

The reduction in behavioural classes was associated with improved model performance. Earlier models with more behavioural classes showed more variable F1 scores and higher CV% between observed and detected prevalences (Figure 2.5). This indicated that rare behaviours and/or behaviours with similar movement patterns were difficult to classify reliably. Later models produced more stable performance after poorly classified behaviours were removed or combined into broader classes, although this also reduced the behavioural resolution. Model 5 was selected as the optimal model for subsequent behavioural analysis because it retained six behavioural classes while achieving high agreement, a higher F1 score, and lower CV% than the earlier models. Model 6 had the lowest CV%, but its three-category structure provided less

behavioural detail than Model 5.



(A)



(B)

Figure 2. 5. Comparison of model performance across the six random forest models. (A) Distribution of F1 scores across behavioural classes for each model. (B) CV% between observed and predicted behavioural prevalence for each model. CV% represents the variation between observed and detected prevalence.

2.3.1 Model 1 (24 behavioural classes)

Model 1, which had 24 behavioural classes, had a kappa coefficient of 0.907 and an average F1 score of 0.7 ± 0.23. Across the behavioural classes, model performance was uneven. Higher F1 scores were observed for 'walk', 'lying asleep', 'lying alert', and 'drink'. The mean CV% was 42.4% ± 54.46%, indicating poor agreement between observed and detected prevalence for some behavioural classes. The highest CV% values occurred for behaviours with no detected prevalence, including crouching, raking, scratch, and shake, mainly reflecting failed detection. Other behaviours with large differences between observed and detected prevalences included 'stretch' (135.40%), 'jump/climb' (84.85%), 'toileting' (39.89%), and 'scent mark' (39.12%) (Table 2.3).

The confusion matrix showed that several behaviours were commonly misclassified as 'walk', including 'shake' (69.23%), 'jump/climb' (54.17%), 'stretch' (43.48%), 'sniff' (35.62%), 'raking' (33.33%), 'scent mark' (28.33%), 'trot' (27.50%) and 'scratch' (25.00%). 'Jump/climb' and 'trot' were also misclassified as 'run' in 20.83% and 15.00% of observations respectively, while 'groom' and 'stretch' were frequently misclassified as 'allogroom' (24.94% and 19.57% of observations, respectively). 'Scent mark' was misclassified as 'stand' in 18.33% of observations, while 'toileting' was misclassified as 'sit' in 16.00% of observations (Appendix 2).

2.3.2 Model 2 (16 behavioural classes)

Model 2 had 16 behavioural classes after the three lying alert classes were merged into 'lying alert', and the two lying asleep classes were merged into 'lying asleep'. 'Crouching', 'raking', 'scratch', 'shake', and 'stretch' were removed from the subsequent model structure because they were either not identified or showed poor performance metrics in Model 1. Model 2 had a kappa coefficient of 0.909 and an average F1 score decreased slightly to 0.6 ± 0.24. Compared with Model 1, the kappa coefficient remained similar, while the mean CV% decreased from 42.4% ± 54.46% to 27.6% ± 32.07%. These shifts indicated improved agreement between observed and detected prevalence for some behavioural classes, despite a drop in macro F1 score. However, several problematic classes remained. The lowest sensitivity values were observed for 'jump/climb' (0.08), 'stand' (0.29), 'scent mark' (0.31), 'sniff' (0.40), 'toileting' (0.22), and 'trot' (0.48). The largest prevalence differences were found for 'jump/climb', 'scent mark', 'stand', and 'toileting', with CV% values of 119.66%, 51.43%, 48.48%, and 72.8%, respectively (Table 2.3).

Table 2. 3. Performance characteristics of random forest Models 1 and 2, including sensitivity, specificity, balanced accuracy, F1 score, observed prevalence, detection prevalence and coefficient of variation (CV%) and for each behaviour. CV% represents the variation between observed and detected prevalence. Values in the Average row are presented as mean ± SD across behavioural classes. Behaviours highlighted in red have a CV% > 20, while cells highlighted in orange indicate performance values < 0.70.

Model 1							
Behaviour	**Sensitivity**	**Specificity**	**Balanced Accuracy**	**F1**	**Prevalence (Observed)**	**Detection Prevalence**	**CV%**
Allogroom	0.82	0.99	0.91	0.75	0.02	0.03	13.13
Crouching	0.00	1.00	0.50	NaN	0.00	0.00	141.42
Drink	0.78	1.00	0.89	0.86	0.00	0.00	15.71
Eat	0.62	1.00	0.81	0.69	0.00	0.00	15.40
Groom	0.57	1.00	0.78	0.65	0.01	0.00	21.50
Jump/climb	0.08	1.00	0.54	0.13	0.00	0.00	84.85
Lying alert L	0.89	0.98	0.93	0.87	0.12	0.12	3.14
Lying alert R	0.86	0.98	0.92	0.86	0.11	0.10	0.55
Lying alert ventral	0.70	1.00	0.85	0.79	0.01	0.01	19.04
Lying asleep L	0.99	1.00	0.99	0.99	0.29	0.28	0.38
Lying asleep R	0.99	1.00	0.99	0.99	0.23	0.24	0.07
Raking	0.00	1.00	0.50	NaN	0.00	0.00	141.42
Roll	0.66	1.00	0.83	0.68	0.00	0.00	3.49
Run	0.65	1.00	0.82	0.66	0.00	0.00	3.51
Scent mark	0.29	1.00	0.64	0.37	0.00	0.00	39.12
Scratch	0.00	1.00	0.50	NaN	0.00	0.00	141.42
Shake	0.00	1.00	0.50	NaN	0.00	0.00	141.42
Sit	0.74	0.99	0.87	0.75	0.04	0.04	2.13
Sniff	0.37	1.00	0.68	0.44	0.00	0.00	27.13
Stand	0.44	0.99	0.72	0.49	0.02	0.01	18.29
Stretch	0.00	1.00	0.50	NaN	0.00	0.00	135.40
Toileting	0.42	1.00	0.71	0.54	0.00	0.00	39.89
Trot	0.55	1.00	0.77	0.53	0.00	0.00	5.11
Walk	0.98	0.99	0.98	0.96	0.15	0.15	3.06
Average	*0.5±0.35*	*1.0±0.01*	*0.7±0.17*	*0.7±0.23*	*Σ=1.00*	*Σ=1.00*	*42.4±54.46*
Model 2							
Allogroom	0.81	0.99	0.90	0.75	0.02	0.02	7.00
Drink	0.88	1.00	0.94	0.90	0.00	0.00	3.63
Eat	0.64	1.00	0.82	0.69	0.00	0.00	12.36
Groom	0.53	1.00	0.76	0.64	0.01	0.00	29.86
Jump/climb	0.08	1.00	0.54	0.15	0.00	0.00	119.66
Lying alert	0.96	0.96	0.96	0.92	0.24	0.26	5.67
Lying asleep	0.99	0.99	0.99	0.99	0.52	0.52	0.29
Roll	0.59	1.00	0.80	0.67	0.00	0.00	19.37
Run	0.76	1.00	0.88	0.70	0.00	0.00	10.55
Scent mark	0.31	1.00	0.65	0.42	0.00	0.00	51.43
Sit	0.61	1.00	0.80	0.71	0.04	0.03	22.36
Sniff	0.40	1.00	0.70	0.48	0.00	0.00	27.47
Stand	0.29	1.00	0.64	0.38	0.02	0.01	48.48
Toileting	0.22	1.00	0.61	0.33	0.00	0.00	72.85
Trot	0.48	1.00	0.74	0.49	0.00	0.00	5.51
Walk	0.98	0.99	0.98	0.96	0.15	0.15	2.21
Average	*0.60±0.29*	*1.0±0.01*	*0.8±0.14*	*0.6±0.24*	*Σ=1.00*	*Σ=1.00*	*27.6±32.07*

Abbreviation: NaN, not a number, indicates that the metric could not be calculated because there were no predicted observations for that behavioural class.

The confusion matrix showed that the main errors were now more concentrated among behaviours with similar posture or movement characteristics. Again, several behaviours were commonly misclassified as 'walk', including 'jump/climb' (33.3%), 'trot' (35.0%), 'sniff' (31.2%), 'scent mark' (26.1%). Several other behaviours were frequently misclassified as 'lying alert', including 'stand' (39.2%), 'sit' (31.3%), 'toileting' (30.0%), 'scent mark' (21.7%), and roll (15.7%). 'Eat' and 'groom' were also misclassified as 'allogroom' in 35.71% and 21.8% of observations, respectively. In addition, jump/climb was misclassified as run in 37.5% of observations, while 'toileting' was misclassified as 'sit' in 32.0% of observations (Appendix 3)

2.3.3 Model 3 (11 behavioural classes)

Model 3 had 11 behavioural classes. 'Jump/climb', 'run', 'trot', and 'walk' were merged into a 'locomotion' class. 'Scent mark' and 'toileting' were removed from the subsequent model structure because both continued to show weak classification performance and large prevalence differences in Model 2. Model 3 had a kappa coefficient of 0.913 and an average F1 score of 0.7 ± 0.21. Compared with Model 2, the kappa coefficient increased slightly and the mean CV% decreased from 27.6% ± 32.07% to 19.4% ± 17.09%. Class level performance improved more clearly than in the earlier iteration (Table 2.4). Mean sensitivity rose from 0.6 ± 0.29 in Model 2 to 0.7 ± 0.25 in Model 3, though F1 scores remained low for 'stand' (0.39), 'sniff' (0.41), 'eat' (0.63). The lowest sensitivity values in Model 3 were observed for 'eat' (0.54), 'groom' (0.59), 'sit' (0.60), 'sniff' (0.31), and 'stand' (0.29). The largest differences between observed and detected prevalence were found for 'stand', 'sniff', 'groom', and 'eat', with CV% values of 50.73, 46.71, 26.66, and 24.72 respectively (Table 2.4).

The confusion matrix showed that the remaining misclassifications in Model 3 mainly involved a smaller number of behavioural classes. Several behaviours were commonly misclassified as 'lying alert', including 'stand' (41.31%), 'sit' (33.20%), and 'roll' (15.66%). In addition, 'sniff' was misclassified as 'locomotion' in 35.96% of observations, 'eat' was misclassified as 'allogroom' in 35.71% of observations, 'groom' was misclassified as 'allogroom' in 21.79% of observations, and 'roll' was misclassified as 'locomotion' in 20.48% of observations (Table 2.5).

Table 2. 4. Performance characteristics of random forest Models 3 to 6. The table shows sensitivity, specificity, balanced accuracy, F1 score for each behaviour, observed prevalence, detection prevalence and coefficient of variation (CV%). CV% represents the variation between observed and detected prevalence. Values in the Average row are presented as mean ± SD across behavioural classes. Behaviours highlighted in red have a CV% > 20, while orange shading indicates performance values < 0.70.

Behaviour	Sensitivity	Specificity	Balanced Accuracy	F1	Prevalence (Observed)	Detection Prevalence	CV%
Model 3							
Allogroom	0.80	0.99	0.90	0.76	0.02	0.02	7.17
Drink	0.85	1.00	0.93	0.92	0.00	0.00	11.47
Eat	0.54	1.00	0.77	0.63	0.00	0.00	24.72
Groom	0.59	1.00	0.79	0.70	0.01	0.00	26.66
Locomotion	0.98	0.99	0.99	0.97	0.15	0.15	1.93
Lying alert	0.96	0.96	0.96	0.92	0.24	0.26	5.86
Lying asleep	0.99	0.99	0.99	0.99	0.52	0.52	0.30
Roll	0.64	1.00	0.82	0.70	0.00	0.00	14.05
Sit	0.60	1.00	0.80	0.70	0.04	0.03	23.48
Sniff	0.31	1.00	0.65	0.41	0.00	0.00	46.71
Stand	0.29	1.00	0.64	0.39	0.02	0.01	50.73
Average	*0.7±0.25*	*1.0±0.01*	*0.8±0.12*	*0.7±0.21*	*Σ=1.00*	*Σ=1.00*	*19.4±17.09*
Model 4							
Active	0.97	0.99	0.98	0.96	0.15	0.16	1.75
Drink	0.78	1.00	0.89	0.87	0.00	0.00	17.93
Eat	0.52	1.00	0.76	0.62	0.00	0.00	27.68
Groom/Allogroom	0.80	0.99	0.90	0.80	0.03	0.03	1.38
Lying alert	0.96	0.96	0.96	0.93	0.24	0.26	5.41
Lying asleep	0.99	0.99	0.99	0.99	0.52	0.52	0.25
Sit	0.62	1.00	0.81	0.71	0.04	0.03	22.66
Stand	0.27	1.00	0.63	0.39	0.02	0.01	61.52
Average	*0.7±0.26*	*1.0±0.01*	*0.9±0.13*	*0.8±0.21*	*Σ=1.00*	*Σ=1.00*	*17.3±20.77*
Model 5							
Active	0.97	0.99	0.98	0.97	0.15	0.15	0.28
Drink	0.83	1.00	0.91	0.90	0.00	0.00	13.56
Eat	0.56	1.00	0.78	0.67	0.00	0.00	28.28
Groom/Allogroom	0.79	1.00	0.89	0.81	0.03	0.03	3.07
Lying asleep	0.99	0.99	0.99	0.99	0.52	0.52	0.43
Resting alert	0.97	0.98	0.97	0.96	0.29	0.30	1.13
Average	*0.9±0.16*	*1.0±0.01*	*0.9±0.08*	*0.9±0.12*	*Σ=1.00*	*Σ=1.00*	*7.8±11.23*
Model 6							
Active	0.96	0.99	0.98	0.96	0.16	0.16	0.01
Inactive	0.99	0.95	0.97	0.99	0.81	0.82	0.25
Maintenance	0.78	1.00	0.89	0.82	0.03	0.03	6.81
Average	*0.9±0.11*	*1.0±0.03*	*1.0±0.05*	*0.9±0.09*	*Σ=1.00*	*Σ=1.00*	*2.4±3.86*

2.3.4 Model 4 (eight behavioural classes)

Model 4 provided a further improvement by reducing the classification structure to eight behavioural classes. 'Allogroom' and 'groom' were merged into 'groom/allogroom' due to their consistent misclassification in previous models. 'Locomotion', 'roll', and 'sniff' were grouped into a broader 'active' class, because the confusion matrix showed persistent overlap among these behaviours in Model 3. 'Drink', 'eat', 'lying alert', 'lying asleep', 'sit', and 'stand' were

retained separately. Compared with Model 3, Model 4 showed a slight improvement in overall performance, with a kappa coefficient of 0.919 and average F1 score of 0.8 ± 0.21. The Mean CV% also decreased from 19.4% ± 17.09% to 17.3% ± 20.77%, although the reduction was smaller than that observed between Models 2 and 3. Weak F1 scores remained for 'stand' (0.39), 'eat' (0.62). The lowest sensitivity values in Model 4 were observed for 'eat' (0.52), 'sit' (0.62), and 'stand' (0.27). The largest differences between observed and detected prevalence were found for 'stand', 'eat', and 'sit', with CV% values of 61.52, 27.68, and 22.66 respectively (Table 2.4). Thus, based on all metrics, the 'stand', 'eat' and 'sit' were a consistent source of error in Model 4.

The confusion matrix showed that 'stand' and 'sit' were confused with 'lying alert', with 37.81% and 31.64% of their observed records misclassified as 'lying alert', respectively. In addition, 'eat' was misclassified as 'groom/allogroom' in 28.57% of observations and as 'active' in 19.05% of observations, while 'stand' was misclassified as 'active' in 20.52% of observations. Compared with the earlier models, Model 4 reduced the spread of large errors, but it did not eliminate the persistent overlap among rest related classes (e.g., 'stand' and 'active'), or between 'eat' and 'groom/allogroom' (Table 2.5).

2.3.5 Model 5 (six behavioural classes)

The strongest performance was observed in the final two models, where feature selection and down-sampling were also applied. Model 5 reduced the behavioural classes to six and retained the biologically important categories 'active', 'drink', 'eat', 'groom/allogroom', 'lying asleep', and 'resting alert'. 'Sit', 'stand' and 'lying alert' were merged into a 'resting alert' class, because these classes were a major source of misclassification in Model 4.

Model 5 showed clear improvement, with a kappa coefficient of 0.955, an average F1 score of 0.9 ± 0.12, and a mean sensitivity of 0.9 ± 0.16. Mean CV% also dropped substantially from 17.3% ± 20.77% to 7.8% ± 11.23%. 'Eat' remained the weakest class in Model 5, with the lowest F1 score (0.67), the lowest sensitivity value (0.56), and the largest prevalence difference (CV% = 28.28%) (Table 2.4). The confusion matrix showed that the remaining major errors were limited to two behaviour pairs. 'Eat' was misclassified as 'groom/allogroom' in 35.12% of observations, and 'groom/allogroom' was misclassified as 'resting alert' in 16.45% of observations (Table 2.6).

Table 2. 5. Confusion matrices showing the percentage of predicted and observed behaviours for Models 3 and 4. Correct classifications are highlighted in yellow, while red cells indicate misclassifications greater than 15%.

Model 3

Prediction	Observed behaviour										
	Allogroom	Drink	Eat	Groom	Locomotion	Lying alert	Lying asleep	Roll	Sit	Sniff	Stand
Allogroom	79.93	2.50	35.71	21.79	0.67	0.54	0.02	0.00	1.99	14.73	4.60
Drink	0.00	85.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Eat	1.43	0.00	53.57	0.24	0.04	0.01	0.00	0.00	0.00	0.34	0.00
Groom	1.02	0.00	1.79	58.60	0.02	0.01	0.01	0.00	0.16	2.40	0.37
Locomotion	3.20	7.50	7.74	5.57	98.02	0.26	0.01	20.48	1.95	35.96	15.36
Lying alert	11.56	5.00	0.00	8.47	0.46	96.33	1.03	15.66	33.20	7.88	41.31
Lying asleep	0.41	0.00	0.00	0.48	0.00	1.35	98.91	0.00	0.04	2.05	0.46
Roll	0.00	0.00	0.00	0.24	0.04	0.06	0.00	63.86	0.00	0.34	0.00
Sit	1.22	0.00	0.00	0.48	0.06	1.07	0.01	0.00	60.27	1.37	7.73
Sniff	0.48	0.00	0.00	1.21	0.13	0.06	0.00	0.00	0.35	30.82	1.20
Stand	0.75	0.00	1.19	2.91	0.56	0.32	0.01	0.00	2.03	4.11	28.98
Total Observations(s)	1,470	40	168	413	9,897	15,825	35,014	83	2,560	292	1,087

Model 4

Prediction	Observed behaviour							
	Active	Drink	Eat	Groom/Allogroom	Lying alert	Lying asleep	Sit	Stand
Active	97.50	12.50	19.05	5.94	0.44	0.01	2.66	20.52
Drink	0.00	77.50	0.00	0.00	0.00	0.00	0.00	0.00
Eat	0.03	0.00	51.79	1.22	0.00	0.00	0.00	0.00
Groom/Allogroom	1.22	5.00	28.57	80.36	0.68	0.01	2.54	5.24
Lying alert	0.63	5.00	0.60	9.71	96.43	1.01	31.64	37.81
Lying asleep	0.06	0.00	0.00	0.32	1.39	98.94	0.31	0.55
Sit	0.12	0.00	0.00	1.11	0.90	0.01	61.52	9.02
Stand	0.45	0.00	0.00	1.33	0.16	0.01	1.33	26.86
Total Observations(s)	10,272	40	168	1,884	15,825	35,014	2,560	1,087

Table 2. 6. Confusion matrices showing the percentage of predicted and observed behaviours for Models 5 and 6. Correct classifications are highlighted in yellow, while red cells indicate misclassifications greater than 15%.

Model 5

Prediction	Observed behaviour					
	Active	Drink	Eat	Groom/Allogroom	Lying asleep	Resting alert
Active	96.88	12.50	7.74	3.56	0.01	1.40
Drink	0.00	82.50	0.00	0.00	0.00	0.00
Eat	0.01	0.00	55.95	0.90	0.00	0.00
Groom/Allogroom	0.92	0.00	35.12	78.87	0.01	0.83
Lying asleep	0.00	0.00	0.00	0.21	98.82	1.02
Resting alert	2.19	5.00	1.19	16.45	1.17	96.74
Total Observations(s)	10,272	40	168	1,884	35,014	19,472

Model 6

Prediction	Observed behaviour		
	Active	Inactive	Maintenance
Active	96.10	0.56	4.94
Inactive	2.94	99.13	16.84
Maintenance	0.97	0.31	78.22
Total Observations(s)	10,452	54,487	2,144

2.3.6 Model 6 (three behavioural classes)

Model 6 represented the most simplified classification structure and grouped behaviours into the three broad classes: 'active', 'inactive', and 'maintenance'. 'Groom/allogroom', 'eat', 'toileting', and 'drink' were grouped into 'maintenance', while 'lying asleep' and 'resting alert' were grouped into inactive. In addition, 'scent mark' and 'stretch' were added back into 'active' class to test a broader structure. Model 6 had an average F1 score of 0.9±0.09, a mean balanced accuracy of 1.0 ± 0.05, and mean CV% of 2.4% ± 3.86% (Table 2.4). However, kappa decreased from 0.955 in Model 5 to 0.936 in Model 6. The confusion matrix showed that only one major error remained in this model. Maintenance was misclassified as inactive in 16.84% of observations (Table 2.6).

2.4 Discussion

Wearable technologies, such as accelerometers, can record animal movement and posture over long periods, especially when direct observation is limited or not possible (e.g., during

night-time activity or when animals are out of sight) (Brown et al., 2013; Valletta et al., 2017). Given that felids typically exhibit a crepuscular and cryptic lifestyle, wearable technologies may offer insight into the behaviour and welfare of this species in captivity. Thus, this study aimed to develop and validate random forest models for classifying captive cheetah behaviour from triaxial accelerometer data.

The present study developed several RF models to classify cheetah behaviour, each with varying levels of behavioural resolution. Model 5 was selected as the optimal model because it provided the best balance between classification performance, behavioural resolution (six behavioural classes) and the retention of interpretable behavioural categories. This model retained six behavioural categories: 'active', 'drink', 'eat', 'groom/allogroom', 'lying asleep' and 'resting alert'. Retaining 'drink' and 'eat' as separate categories was important as these behaviours have been treated as distinct behavioural classes in previous cheetah accelerometer classification studies (Giese et al., 2021; Grünewälder et al., 2012). Compared with the earlier models, Model 5 showed stronger and more stable behaviour performance, with a F1 score of 0.90 ± 0.12 and a mean CV% of $7.8 \pm 11.23\%$. Although Model 6 had the lowest mean CV%, this was achieved by reducing the behavioural repertoire to three broad categories. Model 5 was therefore considered optimal because it retained greater behavioural detail while maintaining strong classification performance. The pattern of stronger performance after behaviour grouping was broadly consistent with previous cheetah accelerometer research, in which classification performance was generally higher when behaviours were grouped into broader categories (McGowan et al., 2022). Direct comparison between studies is limited by differences in behavioural categories, devices and validation datasets.

Behavioural imbalance and similar acceleration patterns likely contributed to the weaker performance of the earlier models. Indeed, the current dataset was heavily skewed towards resting and low-activity behaviours (81.1% observations), whereas several short duration behaviours, such as scratch, raking, shake, toileting and scent marking, had fewer observations. Imbalanced data are common for studies that evaluate animal behaviour, as animals express some behaviours much more frequently than others, and behaviours with fewer observations or that have similar acceleration patterns are often more difficult to classify (Grünewälder et al., 2012; McGowan et al., 2022; Redmond et al., 2024; Smit et al., 2023). Whether down sampling should be applied remains context-dependent and is often determined empirically based on model performance (Sakai et al., 2019). In behavioural datasets, the probability of a given behavioural class occurring reflects its natural frequency, resulting in inherently imbalanced class distributions. Although down sampling can improve the representation of

infrequent behaviours during model training, it may also reduce overall model performance by diminishing the influence of behaviours that naturally comprise the majority of an animal's activity budget. In the present study, however, down sampling enhanced the performance for Model 5 and 6.

Regardless of whether down sampling is done or not, there is often still an imbalance for important but infrequent/short duration behaviours like eating and drinking. It is therefore important to evaluate model performance based on metrics that account for imbalance. Performance metrics such as sensitivity, precision and F1 score were therefore considered, alongside the agreement between observed and detected prevalence. This approach follows recommendations that precision-recall based metrics are more informative than accuracy alone for imbalanced classification problems (Redmond et al., 2024; Saito & Rehmsmeier, 2015). The agreement between observed and detected prevalence was also important, because such models are primarily used to estimate how much time an animal spends exhibiting each behavioural class. To achieve good model performance, however, it is often necessary to either remove infrequent behaviours or merge them with similar (e.g., similar intensity or motion) behaviours.

In the current study, rare behaviours that could not be reliably classified were removed from later models, while other behaviours with similar acceleration patterns were combined into broader categories. This was particularly important for ‘active’ behaviours, where ‘jump/climb’, ‘run’, ‘trot’ and ‘walk’ were first combined into locomotion, and then later grouped with roll and sniff into an active category. This greatly improved model performance, especially when low movement postural behaviour classes were merged into resting alert in Model 5. Previous cheetah accelerometer studies have shown that model performance is influenced by behavioural resolution, sample size and behaviour type (Giese et al., 2021; Grünewälder et al., 2012; McGowan et al., 2022). Similar approaches have also been used in other accelerometer classification studies, where low sample or similar behaviours were removed or grouped to reduce confusion and improve model performance (Redmond et al., 2024; Smit et al., 2023). The improvement in Model 5 was linked to the use of categories that were more clearly separated in the acceleration data.

The confusion matrices showed that several active behaviours were difficult to separate in the earlier models. ‘Walking’, ‘trotting’, ‘running’, ‘jump/climb’ and ‘sniffing’ were distinguishable in the video observations, but all involved some degree of movement and may not have produced clearly separable acceleration patterns from a collar-mounted device. Grouping these behaviours into ‘active’ reduced this source of misclassification. In captive

cheetah studies, locomotion has been defined broadly as movement from one location to another at any speed, and enrichment treatments have been shown to affect locomotory and exploratory behaviours (Quirke & O'Riordan, 2011a). After running enrichment, captive cheetahs have also been described using different gaits, including walking, trotting, sprinting and galloping, while chasing a lure (Quirke et al., 2013). Ideally, different gaits would be classified as separate behavioural classes within the machine learning models. However, the cheetahs in the present study did not perform high-intensity gaits (e.g., running) with sufficient frequency to provide enough training data for the model to reliably learn these behaviours. Future studies should therefore aim to encourage and capture a broader range of locomotive intensities while collecting concurrent video and accelerometer data. This would provide sufficient labelled examples for the model to distinguish between different gaits and improve the classification of locomotor behaviours.

Some classification errors may have reflected overlap between posture-based and action-based behaviours. 'Stand' represented a static posture, but other behaviours, such as eating, drinking or walking, could occur while the animal was standing or moving from a standing position. Therefore, these behaviours may have shared similar collar-derived acceleration patterns with 'stand', especially when whole-body movement was limited. The static and resting behaviours also showed overlap in the earlier models. 'Stand', 'sit', 'lying alert' and 'lying asleep' were distinguishable in the video observations, but these behaviours involved little movement and may have produced only subtle differences in collar orientation. As a result, their acceleration signals were likely more difficult to separate than those of dynamic behaviours. Similar difficulties have been reported in collar-mounted accelerometer studies, where static postures such as 'stand', 'sit' and 'lying' were difficult to distinguish because changes in neck or body orientation are very small (Kumpulainen et al., 2021; Redmond et al., 2024). Previous models have also used broader resting categories to improve the classification of low-movement behaviours (Den Uijl et al., 2017; Smit et al., 2023). In Model 5, merging 'stand', 'sit' and 'lying alert' into 'resting alert' reduced this source of confusion while still retaining a distinction between 'resting alert' and 'lying asleep'. The confusion between 'groom/allogroom' and 'resting alert' is more difficult to explain, though is most likely because grooming and allogrooming occurred while the animal was in a resting alert state (e.g., sitting or lying). These errors suggest that some labels in the earlier models captured both posture and action, making them difficult to separate using collar-derived acceleration data.

'Eat' remained one of the weaker categories in Model 5 and was most often confused with 'groom/allogroom'. This likely reflects similarities in head and neck movement between these

behavioural categories, particularly when the body was relatively still. Despite this, eat was retained because feeding-related behaviour can provide useful information on feeding events detection in cheetahs (Giese et al., 2021; Grünewälder et al., 2012). From an animal heath perspective, the accurate detection of drinking behaviour is promising as polydipsia can be an early indicator of health concerns like chronic kidney diseases, a condition which is prevalent among cat species (Mitchell et al., 2018).

Collar placement may also influence behaviour detection, because collar-derived acceleration mainly reflects movement at the neck and collar rather than all aspects of whole-body movement. Collar-mounted accelerometers are practical for cheetahs because they can be fitted within existing management procedures and can record movement data without continuous direct observation. This approach has been used in previous cheetah studies to classify behaviour, identify feeding events and record movement dynamics in captive and free-ranging individuals (Giese et al., 2021; Grünewälder et al., 2012; McGowan et al., 2022; Wilson & Mills et al., 2013). Movement of the collar around the neck, changes in device orientation and differences in collar fit may reduce the separation between behaviours with similar postures or movement patterns (Redmond et al., 2024; Smit et al., 2023). This is particularly relevant for low movement behaviours, where small differences in head or neck position may produce similar acceleration signals. Device specification may also influence performance, particularly when behaviours are classified at finer behavioural resolutions (McGowan et al., 2022).

Future work could include more labelled examples of uncommon behaviours, such as drinking, eating, grooming, scent marking and toileting, to improve the model's ability to classify these categories. In addition, extra sensor outputs, including gyroscope data, could be assessed to determine whether they improve the separation of behaviours involving subtle postural changes or head and neck movement (Kumpulainen et al., 2021; Redmond et al., 2024). Sampling frequency and epoch length are also important because brief behaviours may be missed or averaged out when data are summarised over longer time windows (Hounslow et al., 2019). Any increase in sampling frequency would need to be balanced against battery life, memory capacity and data download time from the devices, especially if the model is applied over longer monitoring periods (Giese et al., 2021). Future modelling could also consider methods that use information from neighbouring time points, as some behaviours may occur in short sequences rather than as isolated observations.

2.5 Conclusion

This chapter developed and validated RF models for classifying captive cheetah behaviour from triaxial accelerometer data. Six models were compared across different levels of behavioural resolution, before Model 5 was selected as the optimal model. This model retained six behavioural categories: active, drink, eat, groom/allogroom, lying asleep and resting alert, while maintaining strong classification performance. The results showed that behavioural resolution had an important effect on model performance. More detailed behavioural categories were more difficult to classify when they were rare, short in duration or similar in their acceleration patterns. Grouping similar behaviours improved model stability and reduced misclassification, while still retaining behaviourally meaningful categories. Model 5 offered the most appropriate balance between behavioural detail and classification reliability, making it a suitable model structure for behavioural monitoring in captive cheetahs.

Chapter 3 – Enclosure use and time budget analysis of two captive cheetahs *(Acinonyx jubatus*)

3.1 Introduction

Behavioural observations are widely used in zoo animal studies because they provide information on how captive animals interact with their managed environments. Behavioural data can be useful for captive management because they describe activity patterns and may provide early information about how animals respond to their environment (Mohapatra et al., 2014). Felids may spend long periods resting or hidden, and their activity patterns can be influenced by enclosure design, feeding routines and daily husbandry practices (Weller & Bennett, 2001). Therefore, behavioural time budgets are useful for describing how captive felids allocate observed time among major behavioural states. In captive felid studies, activity budgets have been used to assess day and night activity patterns, feeding behaviours and the proportion of time spent in active, resting and stereotypic behaviours (Mohapatra et al., 2014; Weller & Bennett, 2001). This information is particularly valuable for captive facilities when used in connection with space utilisation data.

Enclosure use is also an important part of behavioural assessment. Measuring space use is a common method for identifying positive and negative aspects of captive environments (Ross & Lukas, 2006). A clearer understanding of space use may also help inform exhibit design and animal welfare assessment (Estevez & Christman, 2006). Enclosure assessment should consider not only the amount of space available, but also the quality of that space (Hunter et al., 2014). Structural features may influence how animals use an enclosure (Mallapur et al., 2002). For example, raised areas have been linked with exploratory behaviour and the use of vantage points in captive cheetahs (Quirke & O'Riordan, 2015). When behaviour and location are recorded together, they can provide a clearer understanding of how animals use the available space and which enclosure features may support different behaviours. Enclosure use studies provide a way to assess how animals use different zones or resources within a captive environment. Previous work has used zone occupancy, spread of participation index and electivity index to describe animal space use, and has also suggested that behavioural records should be combined with enclosure use data (Brereton, 2020). Studies on captive felids have also shown that enclosure design and structural features can influence activity, resting behaviour and space use (Lyons et al., 1997; Mallapur et al., 2002).

Direct observation can provide detailed behavioural and location records, but it is time consuming and can be limited by visibility and observation period. Collar-based triaxial accelerometer classification can support long term behavioural monitoring, but it does not directly show where in the enclosure behaviours occur or which habitat features are being used. The behavioural model developed in Chapter 2 therefore provides a basis for comparing collar-derived behaviour with video observations, while the video records provide the spatial information needed to interpret enclosure use.

The aim of this chapter is to quantify the observed behavioural time budget for two captive cheetahs in Auckland Zoo. In this chapter, the six behavioural categories used in Model 5 ('active', 'drink', 'eat', 'groom/allogroom', 'lying asleep' and 'resting alert') were used to structure an observation-based behavioural analysis. This allowed observed behaviours to be compared with collar behavioural classifications during the matched observation period, serving as a validation for the model developed in Chapter 2. The same observed dataset was also used to record enclosure location, allowing behaviour and space use to be analysed together.

3.2 Materials and Methods

3.2.1 Study location and animals

The study location and animals were the same as those described in Chapter 1 and 2. Animal ethics approval was obtained from the Auckland Zoo Animal Ethics Committee. Briefly, the study was conducted at Auckland Zoo, Auckland, New Zealand, and included two adult female cheetahs (*Acinonyx jubatus*). Both individuals were healthy, approximately eight years of age, and weighed approximately 38 and 40 kg. The cheetahs remained in their usual indoor and outdoor habitats throughout the study and continued to be managed under their normal husbandry routines by regular zoo staff. Both cheetahs were accustomed to wearing normal collars during routine keeper-led walks. All handling associated with collar fitting and removal was carried out by experienced keepers with whom the animals were familiar. This chapter relied on observed behavioural and spatial data. During video scoring, the two cheetahs were distinguished using the visible tail markings described in Section 1.6 and shown in Figure 1.3, with Qia having the larger white marking near the distal end of the tail and Quartz having the smaller marking in the same area.

3.2.2 Habitat management

The two cheetahs were housed together in a landscaped, naturalistic enclosure that

included both indoor and outdoor habitat areas. For this study, the enclosure was divided into eight analytical zones to allow the location of the cheetahs and their behaviours to be recorded systematically during video scoring. These zones were defined based on visible enclosure features and the physical layout of the habitat. The position of each zone is shown in Figure 3.1, and the main features of each zone are summarised in Table 3.1.

The indoor habitat and keeper access areas were located at the upper end of the enclosure (Zone 1; Figure 3.1). The keeper access area was separated from the indoor habitat by a metal gate, allowing keepers to safely manage access to the indoor area. The indoor and outdoor areas were also connected by a controlled metal slide, which allowed staff to manage the cheetahs' access between the two areas. Water was available *ad libitum* in both the indoor area and the outdoor habitat. The Zone 2 water bowl was fully obscured by vegetation in the video footage. The outdoor habitat consisted of grassed areas, dense vegetation, mature trees, rocks, logs, shaded areas, and several enrichment related structures. These included a pergola in the central area of the enclosure (Zone 5) and a three-tier vertical stand in the lower section of the outdoor habitat (zone 6). A public viewing area was positioned along one side of the enclosure (Zone 3).

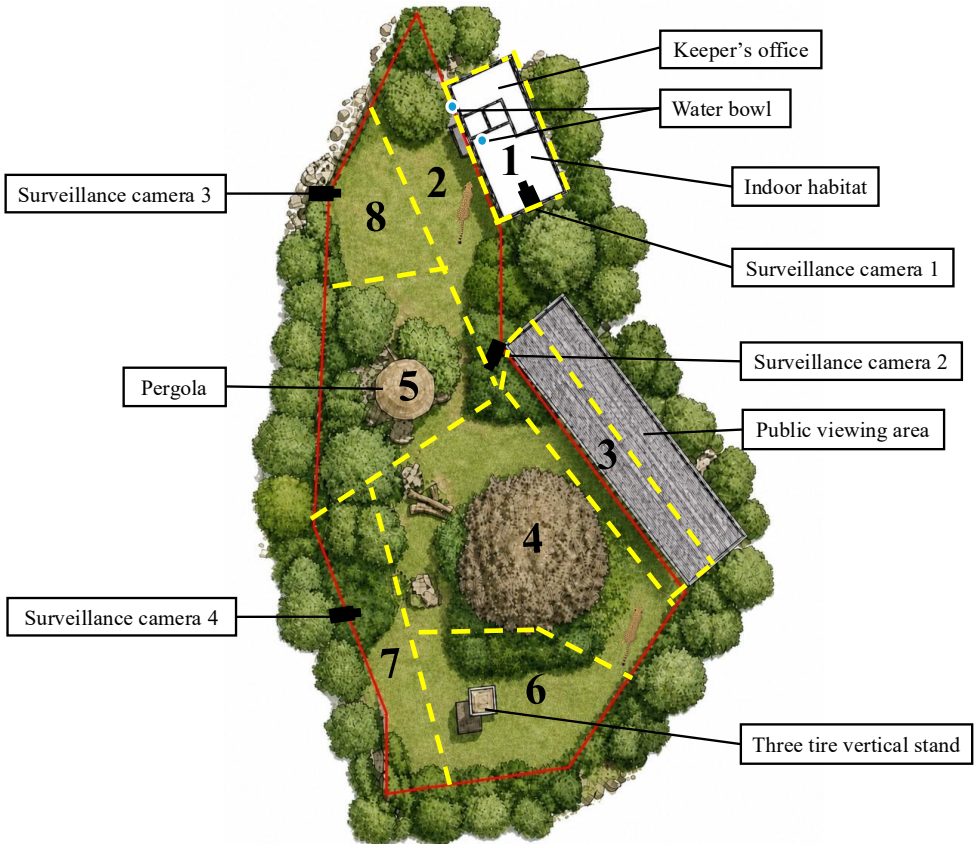


Figure 3. 1. Layout of the cheetah enclosure showing the eight analytical zones, four surveillance cameras, and main habitat features used for behavioural and spatial analysis, including the indoor habitat, keeper's office, water bowl, pergola, public viewing area, and three-tier vertical stand.

Table 3.1. Description of the eight zones used for enclosure use analysis.

Zone	Description
Zone 1	Indoor habitat; keeper access area; indoor water bowl
Zone 2	Open grassed area; large accessible shrub area; outdoor water bowl obscured by vegetation; boundary fence with outside view
Zone 3	Public viewing side; roof overhang; straw-covered ground; lower shrub
Zone 4	Central tree; horizontal logs; large rocks
Zone 5	Pergola; surrounding shrub vegetation
Zone 6	Three-tier vertical stand; Open grassed
Zone 7	Small shrubs; small tree; boundary fence with outside view
Zone 8	Flat open grassed

3.2.3 Video recording and behavioural scoring

Video footage was obtained from four Reolink Duo 2 surveillance cameras (Reolink Technology Co., Wilmington, Delaware, USA). These cameras provided a 180° panoramic view, 4K resolution at 20 FPS, and 30 m infrared night vision. The cameras were connected to a Reolink RLN36 36-channel NVR (Reolink, Hong Kong, China), which recorded continuously the cheetahs' indoor and outdoor habitat areas from 1 April until the end of 9 May 2025. For the present chapter, footage recorded from 5 May 2025 at 01:00:00 to 10 May 2025 at 00:00:00 was selected for behavioural and spatial analysis. This period corresponded with the available collar-derived behavioural data.

The four surveillance cameras provided overlapping coverage of the indoor and outdoor habitat areas (Figure 3.1). Surveillance camera 1 covered the indoor habitat and the section of Zone 2 closest to the metal gate between the indoor and outdoor areas. Surveillance camera 2 covered Zone 3, including the transparent public viewing barrier, the shrub at the end of Zone 3, the front pergola area in Zone 5, and the area in front of the large central tree in Zone 5. Surveillance camera 3 covered Zone 2, Zone 8, and the shrub vegetation in Zone 5. Surveillance camera 4 covered Zones 6 and 7, the rear section of the pergola in Zone 5, and the horizontal logs and rocks in Zone 4.

The behavioural categories used for video scoring followed the six-class structure of Model 5, which was selected as the optimal model in Chapter 2. The selected video footage was reviewed using one minute focal sampling and Boris® software (Friard & Gamba, 2016). Behavioural scoring was conducted using the categories listed in Table 3.2. The six categories were 'active', 'drink', 'eat', 'groom/allogroom', 'lying asleep', and 'resting alert'. As per Chapter 2, the 'active' category included 'walk', 'run', 'trot', 'jump/climb', 'roll', and 'sniff'; 'groom' and 'allogroom' were combined as 'groom/allogroom'; 'lying asleep left' and 'lying

asleep right' were included under 'lying asleep'; and 'lying alert left', 'lying alert right', 'lying alert ventral', 'stand' and 'sit' were included under 'resting alert'. 'Drink' and 'eat' were retained as separate categories. Although the camera system provided overlapping coverage of the enclosure, some areas could be partially obscured by vegetation, enclosure structures or low light conditions. Thus, an additional category of "out of sight" was added to the behavioural scoring structure.

Table 3.2. Behavioural categories and corresponding behaviours used for video scoring. Base behavioural definitions were adapted from Stanton et al. (2015). Out of sight and keeper-led walking were recorded as additional observed categories and were not part of the six Model 5 behavioural categories.

Behavioural Cat.	Definition
Active	
Walk	Forward locomotion at a slow gait.
Run	Forward locomotion in a rapid gait, which is faster than walking or trotting.
Trot	Forward locomotion at a swift gait performed with alternating steps. Movement is faster than walking but slower than running.
Jump/Climb	Ascends and/or descends an object or structure
Roll	While lying on the ground, rotates body from one side to another. During the roll, the back is rubbed against ground, the belly is exposed and all paws are in the air. May continue rolling repeatedly from side to side.
Sniff	Smells (modifier) by inhaling air through the nose.
Drink	Ingests water (or other liquids) by lapping up with the tongue.
Eat	Ingests food (or other edible substances) by means of chewing with the teeth and swallowing.
Groom/allogroom	
Groom	Cleans itself by licking, scratching, biting or chewing the fur on its body. May also include the licking of a front paw and wiping it over one's head.
Allogroom	Cats will briefly push and/or rub their head forcefully against another.
Lying asleep	
Lying asleep L	Lying on the left side of their body, either extended or curled. Their head resting against (modifier). They are not easily disturbed.
Lying asleep R	Lying on the right side of their body, either extended or curled. Their head resting against (modifier). They are not easily disturbed.
Resting alert	
Lying alert L	Lying on the left side of their body, either extended or curled. Their head is raised off the ground.
Lying alert R	Lying on the right side of their body, either extended or curled. Their head is raised off the ground.
Lying alert ventral	Lying with their abdomen in contact with the ground. Their head is raised off the ground.
Stand	In an upright position and immobile, with all four paws on the ground and legs extended, supporting the body
Sit	In an upright position, with the hind legs flexed and resting on the ground, while front legs are extended and straight
Out of sight	Cheetah was not visible; Additional video-based category
Keeper-led walking	Additional video-based category

Abbreviations: Cat, category; L, left; R, right.

3.2.4 Processing the video data

The Boris® output files were processed using R statistical software version 4.5.2 (R Foundation for Statistical Computing, Vienna, Austria). All exported comma-separated value files (.csv) were imported and combined into a single dataset. The dataset was cleaned to retain only the variables required for behavioural, spatial, and temporal analysis. A real timestamp was generated for each behavioural record by extracting the recording date and hour from the observation ID and calculating the minute position within each video file using the Boris® time and time offset. Data quality checks were then conducted to identify missing minutes and to confirm that records were present for both cheetahs in each recorded minute.

Zone information was recorded during video scoring using the eight zones defined for this study in Figure 3.1. When a cheetah was located within an identifiable zone, but its behaviour could not be clearly determined, the behavioural category was coded as 'Out of sight' but the record was included in the total time for that zone. When both behaviour and location could not be determined, the behaviour was coded as 'Out of sight' and the zone was recorded as 'Unknown'. Records with Unknown zone were retained as a separate category in the full enclosure use summary. 'Keeper-led' walking was retained as a separate observed category because it occurred during the observation period but was not part of the Model 5 behavioural categories.

For the observed behavioural time budget analysis, only the six behavioural categories represented in Model 5 were retained. Behaviours or categories not represented in Model 5, including 'Out of sight', 'keeper-led walking', 'shake', 'scent mark', 'toileting' and 'scratch', were excluded from this retained behavioural dataset. Percentages for the behavioural time budget were calculated from the retained Model 5 behavioural dataset for each cheetah. Enclosure use was calculated using the full observed dataset to account for the complete observation period. 'Keeper-led walking' was recorded using start and stop events, and its duration was calculated from the difference between the stop and start times.

3.2.5 Collar-video behavioural validation

Collar-derived behavioural data were obtained from the Model 5 classifications generated from the accelerometer data in Chapter 2. Before validation, the Boris® scored video data were summarised to match the temporal resolution of the collar derived dataset. Hourly summaries were generated for each cheetah by calculating the number of minutes and percentage of time spent in each behavioural category within each hour. Daily summaries were then generated by summing the hourly values for each cheetah and date.

Validation was conducted using three days of time-aligned video-scored and collar-derived behaviour data. The first and final days were excluded to avoid potential instability associated with the beginning and end of the collar recording period. The middle three days therefore provided the most stable matched dataset for direct comparison between observed and collar-derived behavioural percentages. Daily summaries were used to compare mean behavioural percentages between observed and collar data, while hourly summaries were used for regression analysis. Paired t-tests were used to compare daily video-observed and collar-derived behavioural percentages for each behavioural category in each cheetah. The calculations used for behavioural time-budget, enclosure use and collar-video validation analyses are summarised in Table 3.3.

Table 3.3. Calculations used for behavioural time-budget; enclosure use and collar-video validation analyses.

Analysis	Description	Formula / Calculation
Behavioural time budget		
Behavioural time budget (%)	Percentage of retained behavioural time spent in each behaviour.	$P_b = \frac{m_b}{M_r} \times 100$
Enclosure use analysis		
Enclosure zone use (%)	Percentage of total observation time spent in each zone or category.	$P_z = \frac{m_z}{M_t} \times 100$
Validation analysis		
Mean (%)	Mean behavioural percentage across the validation period.	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
Standard error of the mean (SEM)	Standard error of the daily mean behavioural percentage.	$SEM = \frac{SD}{\sqrt{n}}$
CV%	Coefficient of variation between collar-derived and observed behavioural percentages.	$CV\% = \frac{SD(C,V)}{V} \times 100$
Difference	Difference between collar derived and observed behavioural percentages.	$D = C - V$
Linear regression	Regression between hourly observed and collar derived behavioural percentages.	$y = ax + b$
Coefficient of determination	Strength of the linear relationship in the hourly regression.	R^2
Regression p value	Significance level of the linear relationship in the hourly regression.	$p = 2 \times P\left(T_{N-2} \geq \left\lvert\frac{a}{SE(a)}\right\rvert\right)$
Paired t-test	Comparison between paired daily collar-derived and observed behavioural percentages.	$p = 2 \times P\left(T_{n-1} \geq \left\lvert\frac{\bar{d}}{SD_d/\sqrt{n}}\right\rvert\right)$

Note: P_b = behavioural time budget percentage; P_z = zone-use percentage; m_b = minutes recorded for a behaviour; m_z = minutes recorded in a zone or category; M_r = total retained behavioural minutes; M_t = total observation minutes; x_i = daily behavioural percentage; n = number of days; SD = standard deviation; D = difference; C= collar derived percentage; V = observed percentage; x= observed hourly percentage; y = collar-derived hourly percentage; a = regression slope; b = intercept. p = p value for the regression slope; SE(a) = standard error of the regression slope; N = number of hourly paired observations used in the regression; T_{N-2}= t distribution with N-2 degrees of freedom; $\bar{d}$= mean paired difference; SD_d= standard deviation of paired differences; T_{n-1}= t distribution with $n-1$degrees of freedom. Statistical significance was set at $p<0.05$ for paired t-tests and regression analyses.

3.3 Results

Both cheetahs spent most time exhibiting ‘resting alert’ and ‘lying asleep’. Collar data showed similar broad time budget patterns to the video observations, although agreement varied among behaviours. Enclosure use was unevenly distributed, with Zones 3 and 6 accounting for most observed time.

3.3.1 Video observed behavioural time budget

The observed dataset contained 7,140 mins of observation for each cheetah. For the behavioural time budget analysis, behaviours that were not represented in the Model 5 behavioural categories from Chapter 2 were excluded. These included ‘out of sight’, ‘keeper-led walking’, ‘shake’, ‘scent mark’, ‘toileting’ and ‘scratch’. These excluded behaviours accounted for 389 min in Qia and 342 min in Quartz. After these exclusions, the retained observed behavioural dataset contained 6,751 min for Qia and 6,798 min for Quartz.

Based on this retained dataset, ‘resting alert’ was the most frequently recorded behaviour in both cheetahs, accounting for 47.4% of retained behavioural time in Qia and 44.8% in Quartz (Figure 3.2). ‘Lying asleep’ was the second most common behaviour, accounting for 33.2% in Qia and 38.2% in Quartz. ‘Active’ behaviour accounted for 15.5% of retained behavioural time in Qia and 13.2% in Quartz. ‘Groom/allogroom’, ‘eat’ and ‘drink’ were recorded at lower proportions in both individuals, with each behaviour accounting for less than 4% of retained behavioural time.

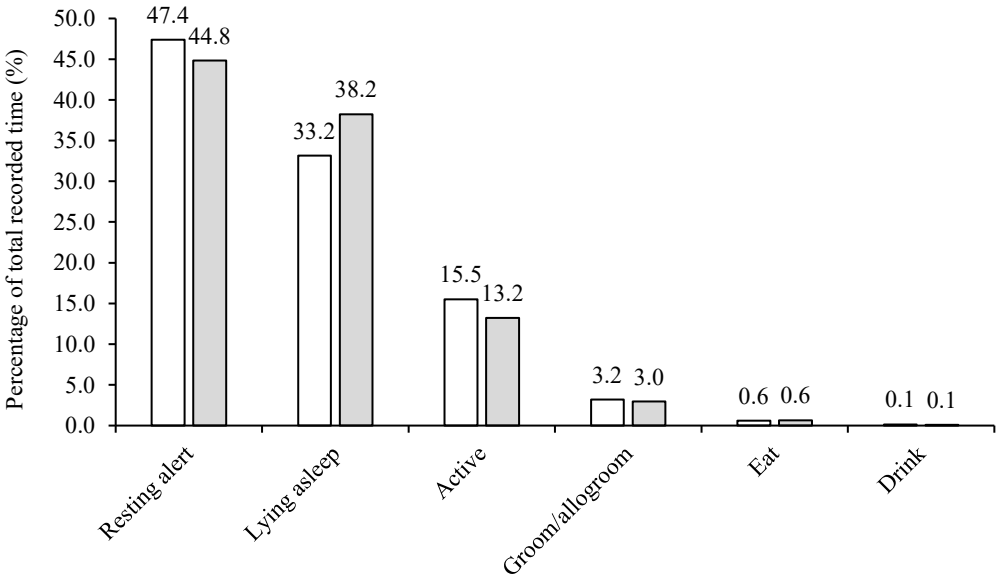


Figure 3. 2. Video observed behavioural time budget for Qia (white) and Quartz (grey) based on Model 5 behavioural categories. Percentages were calculated from the retained behavioural dataset after non-model behaviours were excluded.

3.3.2 Collar versus video behavioural validation

Daily comparisons between observed and collar data behaviours showed similar overall time budget patterns. 'Resting alert' and 'lying asleep' were the dominant behaviours in both video and collar, while 'eat' and 'drink' were recorded at the lowest proportions. The main difference was observed between 'resting alert' and 'lying asleep'. Collar-derived time spent displaying 'resting alert' behaviour was higher for the collar-predicted behaviour than observed for both Qia and Quartz, with differences of 5.42% and 2.67%, respectively (Table 3.4). In contrast, collar-derived 'lying asleep' was lower than observed in both individuals, with differences of -6.76% in Qia and -3.27% in Quartz. 'Active' was slightly higher in the collar datasets than in the video observations for both individuals, with differences of 2.25% in Qia and 1.14% in Quartz. While 'eat' and 'drink' accounted for small proportions of the daily time budget, both behaviours were lower in the collar data than in the video observations (Table 3.4). Paired t-tests showed no significant differences between video-observed and collar-derived daily percentages for four of the six behavioural categories in both cheetahs, including 'drink', 'groom/allogroom', 'resting alert' and 'lying asleep' ($p > 0.05$). Significant differences were detected for 'active' and 'eat' ($p < 0.05$).

Table 3.4. Daily comparison between video observed behaviours and collar datasets percentages for Qia and Quartz. Values are presented as mean ± SEM; Difference and CV% are reported as percentage points.

Subject	Qia				Quartz			
Behavioural categories	**Video**	**Collar**	**Difference**	**CV%**	**Video**	**Collar**	**Difference**	**CV%**
Active	14.03 ± 0.88	16.28 ± 1.08	2.25	11.35	12.28 ± 1.07	13.42 ± 1.09	1.14	6.55
Drink	0.15 ± 0.03	0.05 ± 0.01	-0.10	47.23	0.12 ± 0.02	0.07 ± 0.01	-0.05	31.38
Eat	0.59 ± 0.03	0.06 ± 0.01	-0.53	63.37	0.54 ± 0.05	0.14 ± 0.01	-0.40	52.30
Groom/allogroom	2.16 ± 0.26	1.88 ± 0.11	-0.28	9.20	2.19 ± 0.19	2.10 ± 0.06	-0.09	2.80
Resting alert	49.31 ± 3.86	54.73 ± 2.70	5.42	7.78	44.21 ± 1.99	46.88 ± 2.95	2.67	4.28
Lying asleep	33.76 ± 2.74	27.00 ± 1.93	-6.76	14.16	40.65 ± 2.99	37.38 ± 4.04	-3.27	5.68

After the daily comparison, hourly regression analyses were used to further assess agreement between observed and collar behavioural percentages. Hourly regression results showed clearer agreement for 'active', 'lying asleep' and 'resting alert' than for 'groom/allogroom', 'eat' and 'drink'. 'Active' had the highest R^2 value (0.93), followed by lying asleep ($R^2 = 0.89$) and resting alert ($R^2 = 0.84$). All regression slopes were statistically significant ($p < 0.001$) (Figure 3.3). This shows a significant relationship between collar-derived and video-observed percentages for all behaviours. Collar-derived 'lying asleep' tended to be lower than observed 'lying asleep', whereas collar-derived resting alert tended to be higher than observed resting alert. 'Groom/allogroom' and 'eat' had lower R^2 values of 0.60 and 0.61, respectively, showing weaker correspondence between hourly observed and collar-

derived percentages. 'Eat' also had a low regression slope, and collar values remained low when observed eat increased. Drink showed a moderate R^2 value of 0.7587, but its low regression slope indicated weaker agreement in estimated percentages.

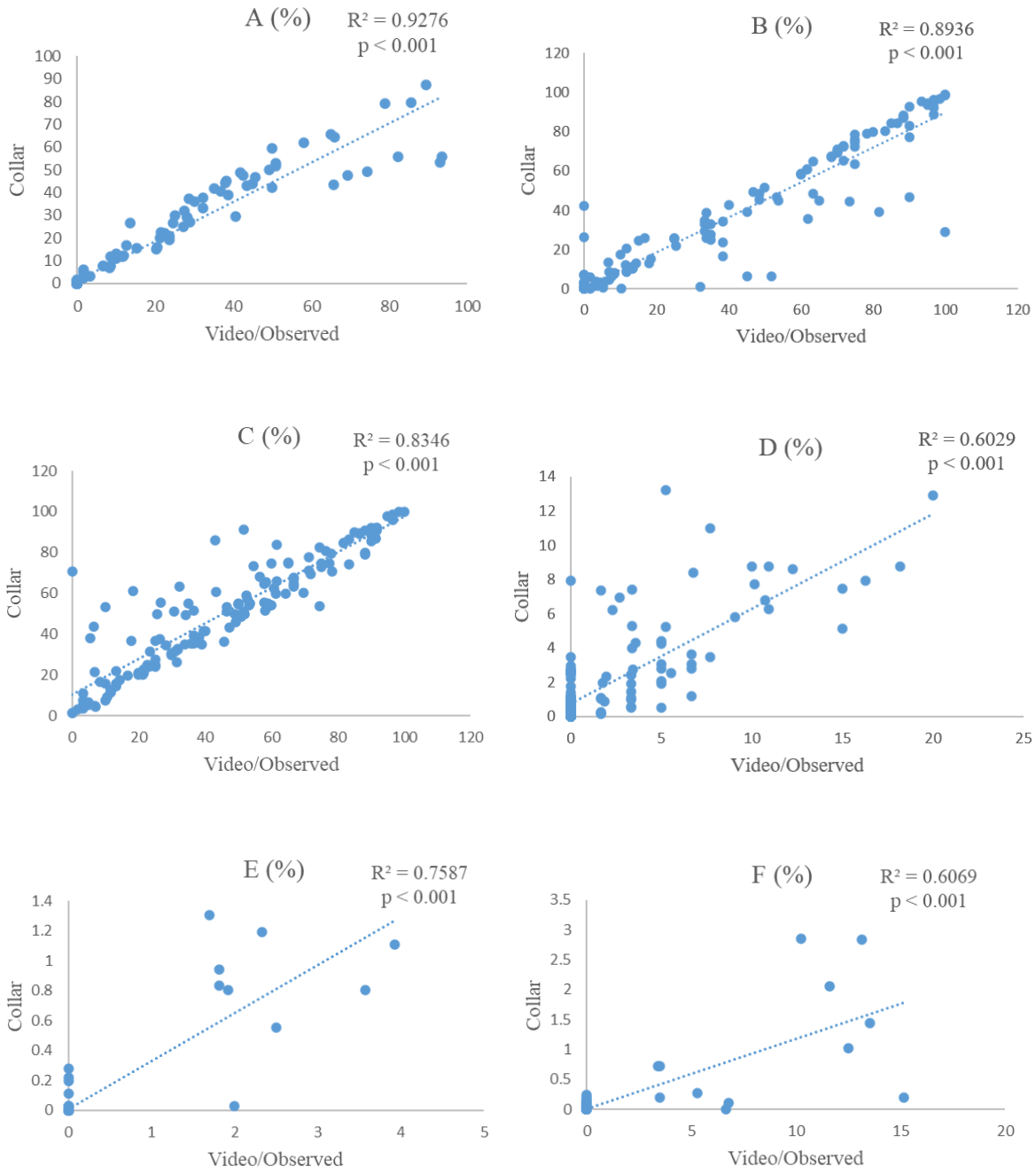


Figure 3. 3. Hourly regression relationships between video-observed and collar-classified behavioural percentages for six behaviours (A) Active, (B) Lying asleep, (C) Resting alert, (D) Groom/allogroom, (E) Drink, and (F) Eat. Each data point represents one hourly comparison between video-observed and collar-classified percentages. Regression lines, R^2 values and p values are shown for each behaviour.

3.3.3 Enclosure zone uses and behavioural distribution by zone

Unlike the behavioural time budget analysis, enclosure use was calculated from the full five days in observed dataset to represent the complete observation period. 'Unknown' and 'keeper-led walking' records were therefore retained, and percentages were calculated from 7,140 mins of observation for each cheetah.

Enclosure use was concentrated in a small number of zones. Both cheetahs used Zone 3 the most during the observation period, accounting for 67.1% of Qia's time and 63.7% of Quartz's time. Zone 6 was the second most used area, accounting for 14.3% in Qia and 18.4% in Quartz, although its use remained substantially lower than that of Zone 3. Use of the remaining zones was comparatively low. Zone 2 was used for 6.6% of the time by Qia and 5.9% by Quartz. Zone 8 was the least used area in both individuals, accounting for less than 0.5% of their time. 'Unknown' and 'keeper-led walking' were retained to account for the complete observation period, but each accounted for less than 2% of both cheetahs' time (Table 3.5).

Table 3.5. Total enclosure zone use. Values are presented as minutes (% of total observation time), calculated from 7,140 min of observation for each cheetah. Brief zone descriptions are provided for reference.

Zone / category	Description	Qia		Quartz	
		Min	%	Min	%
Zone 1	Indoor habitat	255	3.57	344	4.82
Zone 2	Open grassed area	468	6.55	419	5.87
Zone 3	Public viewing side	4,791	67.10	4,550	63.73
Zone 4	Central tree	93	1.30	88	1.23
Zone 5	Pergola	110	1.54	60	0.84
Zone 6	Three-tier vertical stand	1,018	14.26	1,315	18.42
Zone 7	Small shrubs	217	3.04	176	2.46
Zone 8	Flat open grassed	21	0.29	15	0.21
Unknown	Location and behaviour unknown	58	0.81	64	0.90
Keeper-led walking	Walking led by keeper	109	1.53	109	1.53
Total		7,140	100.00	7,140	100.00

Behavioural distribution by zone showed that 'resting alert' and 'lying asleep' accounted for the largest proportions of observed time across zones in both cheetahs. These behaviours were mainly recorded in Zone 3 and Zone 6. For Qia, 'resting alert' and 'lying asleep' in Zone 3 accounted for 29.27% and 28.82% of the total observed time. For Quartz, 'lying asleep' and 'resting alert' in Zone 3 accounted for 34.09% and 23.67%. 'Resting alert' was also frequently recorded in Zone 6, accounting for 11.53% in Qia and 15.35% for Quartz. 'Active' was distributed across several zones rather than concentrated in one area. For Qia, 'Active' was mainly recorded in Zone 3 (5.73%), Zone 2 (4.22%) and Zone 1 (2.07%). For Quartz, 'Active' was mainly recorded in Zone 2 (3.63%), Zone 3 (3.39%) and Zone 1 (3.35%).

‘Groom/allogroom’ was mostly recorded in Zone 3 for both cheetahs, accounting for 2.59% and 2.38% of Qia and Quartz’s time, respectively. ‘Eat’ and ’drink’ were recorded at low proportions across the observation period. Observed ‘drink’ was only recorded in Zone 1 in the observed data for both cheetahs, as the Zone 2 water bowls were outside the field of view (Table 3.6).

3.4 Discussion

This chapter compared collar-derived behavioural classifications with video-observed time budgets and examined behavioural distribution across enclosure zones. The collar data showed a similar broad time-budget pattern to the video observations, although agreement varied among behaviours. The classified collar data provided a way to compare behavioural time budget and enclosure use during the video observation period. For both cheetahs, the observed records were dominated by ‘resting alert’ and ‘lying asleep’, while ‘active’ behaviour made up a smaller proportion of the time budget. The collar data showed a similar broad pattern, although agreement with the video observations varied among behaviours. Collar-based classifications of ‘active’, ‘resting alert’ and ‘lying asleep’ matched the video data more closely than ‘eat’, ‘drink’ and ‘groom/allogroom’, which occurred less often. Enclosure use was also uneven, with both cheetahs recorded most often in Zone 3 and Zone 6.

The high proportion of time spent inactive is not unexpected in zoo-housed cheetahs, as resting and sleeping have been recorded as large parts of captive cheetah time budgets in previous studies (Munita et al., 2016; Seyrling et al., 2022; Skibiel et al., 2007). Wild cheetahs also spend a substantial proportion of their time resting (70.2%), with activity generally concentrated during daylight hours and reduced during hotter periods (Hetem et al., 2019; Mills & Mills, 2017). Cheetahs are adapted for brief, high-intensity pursuits, which may be followed by extended periods of rest or inactivity (Hetem et al., 2019). In captivity, where natural hunting is absent, activity patterns may also be influenced by environmental temperature, feeding schedule, environmental enrichment (Quirke & O’Riordan, 2011b; Seyrling et al., 2022). The difference between ‘resting alert’ and ‘lying asleep’ was important for interpreting this large inactive component. Both behaviours involved little movement, but they represented different levels of alertness. ‘Resting alert’ was defined as when the cheetah was inactive (lying or sitting) but still visually responsive with their head up. In contrast, ‘lying asleep’ was identified when the cheetah was lying with the head lowered and showed no visible alertness; eye closure was used as a supporting cue where it could be seen (Stanton et al., 2015). Although resting alert

and lying asleep made up a large part of the time budget, this alone should not be read as direct evidence of welfare status. Welfare interpretation would need broader evidence, including abnormal behaviours, enclosure use, longer term behavioural patterns and other health or physiology measures (Rose & Riley, 2021). Moreover, past research in captive felids has shown that their activity budgets are influenced by time of day, feeding schedule and housing conditions(Bashaw et al., 2007; Weller & Bennett, 2001).

Table 3.6. Behavioural distribution across enclosure zones for Qia and Quartz based on the full video-observed dataset. Values are presented as minutes (% of total observation time), calculated from 7,140 mins of observation for each cheetah. Cells with no recorded observations are shown as 0. Highlighted cells indicate the main behavioural records in that zone.

Subject	Behaviour categories	Zone 1 (%)	Zone 2 (%)	Zone 3 (%)	Zone 4 (%)	Zone 5 (%)	Zone 6 (%)	Zone 7 (%)	Zone 8 (%)	Unknown (%)	Total (%)
Qia	Active	2.07%	4.22%	5.73%	0.64%	0.99%	0.35%	0.45%	0.21%	0	14.70%
	Drink	0.14%	0	0	0	0	0	0	0	0	0.02%
	Eat	0.13%	0	0.13%	0.20%	0.11%	0	0	0	0	0.60%
	Groom/Allogroom	0.01%	0	2.59%	0	0.03%	0.31%	0.08%	0	0	3.03%
	Resting alert	0.92%	0.36%	29.27%	0.38%	0.25%	11.53%	2.00%	0.08%	0	44.80%
	Lying asleep	0	0	28.82%	0	0.01%	2.02%	0.50%	0	0	31.40%
	Shake	0	0	0.01%	0	0	0	0	0	0	0.01%
	Toileting	0.03%	0	0	0	0.03%	0	0	0	0	0.06%
	Scent Mark	0.27%	0.06%	0.03%	0.03%	0.04%	0	0	0	0	0.42%
	Out of sight	0	1.92%	0.52%	0.06%	0.07%	0.06%	0	0	0.81%	3.43%
	Keeper-led walking	0	0	0	0	0	0	0	0	1.53%	1.53%
	Total (%)	3.57%	6.56%	67.10%	1.31%	1.53%	14.27%	3.03%	0.29%	2.34%	100.00%
Quartz	Active	3.35%	3.63%	3.39%	0.70%	0.56%	0.34%	0.43%	0.20%	0	12.60%
	Drink	0.08%	0	0	0	0	0	0	0	0	0.08%
	Eat	0.21%	0.27%	0.01%	0	0.13%	0	0	0	0	0.62%
	Groom/Allogroom	0.03%	0	2.38%	0	0	0.31%	0.10%	0	0	2.82%
	Lying asleep	0	0.01%	34.09%	0	0	2.30%	0	0	0	36.40%
	Resting alert	1.05%	0.20%	23.67%	0.46%	0.10%	15.35%	1.86%	0	0	42.69%
	Scent Mark	0.07%	0	0.04%	0.03%	0	0	0	0	0	0.14%
	Scratch	0	0	0	0	0	0.01%	0	0	0	0.01%
	Out of sight	0.03%	1.76%	0.14%	0.03%	0.06%	0.11%	0.07%	0.01%	0.90%	3.11%
	Keeper-led walking	0	0	0	0	0	0	0	0	1.53%	1.53%
	Total (%)	4.82%	5.87%	63.72%	1.22%	0.85%	18.42%	2.46%	0.21%	2.43%	100.00%

In the current study, most active behaviour occurred around keeper feeding, feeding enrichment and before 'keeper-led walking' (i.e., the cheetah is walked on a lead around the zoo prior to its opening to visitors). Feeding periods and feeding-related enrichment can be associated with activity peaks in captive cheetahs (Quirke & O'Riordan, 2011b; Seyrling et al., 2022). In this study, active behaviour was partly linked to predictable daily routines, rather than simply reflecting general activity. 'Eat' and 'drink' made up only a small part of the time budget for both cheetahs. 'Drinking' was recorded rarely and was only observed in the morning during the video observation period. 'Eating' was also limited because feeding was controlled by the daily husbandry schedule. Low percentages of feeding and drinking have also been recorded in captive cheetah time-budget studies, where eating and drinking occupied only a small proportion of observed time (Munita et al., 2016). This is exacerbated by the fact that many short duration behaviours, such as 'drinking', are often missed by focal point sample (e.g., 1 minute sampling like in the present study). Wearable technologies that provide continuous behavioural classification (1 second) may offer a more reliable assessment of these short duration behaviours, provided they have a sufficiently unique acceleration profile.

The distinction between 'resting alert' and 'lying asleep' was biologically useful, but the difference between collar-derived and video-observed estimates should be interpreted cautiously, as these were also the most prevalent behaviours in the time budget. In the video observations, these behaviours could be separated using visual cues such as body posture, head position, eye closure and responsiveness. The collar-derived data could reflect some aspects of posture and head or neck position through changes in collar orientation and acceleration, but it could not directly record visual cues such as eye closure or responsiveness. The collar recorded total acceleration across the craniocaudal or surge, mediolateral or sway, and dorsoventral or heave axes, including both static acceleration related to body orientation and dynamic acceleration related to movement (Brown et al., 2013; Shepard et al., 2008). 'Resting alert' and 'lying asleep' may have overlapped in the collar data because both behaviours involved little dynamic acceleration and a broadly similar lying posture. Differences in eye state, head position or responsiveness may have been clear in the video but may not have produced a consistent difference in the three-axis acceleration features. This may explain why collar estimates of resting alert were higher than video observations, while collar estimates of lying asleep were lower.

Active behaviour showed the strongest collar-video agreement because it produced clearer

changes in the triaxial acceleration data. Compared with resting behaviours, active behaviour involved more body movement and larger changes across the acceleration axes, which made it easier for the model to separate from inactive states. This pattern matches the general strength of accelerometer-based classification, where behaviours with distinct and repeatable movement signals are usually identified more reliably than behaviours defined by subtle posture, low-amplitude movement or behavioural context (Valletta et al., 2017). The classification model was better at distinguishing active from inactive periods than at distinguishing different resting behaviours.

'Eat', 'drink' and 'groom/allogroom' were a different problem from the three main behavioural states. They made up only a small part of the video record, so the collar-video comparison was based on far fewer minutes than for 'active', 'resting alert' and 'lying asleep'. Under these conditions, a missed event, a short timing mismatch, or a small classification error could have a larger effect on the percentage agreement. This is especially relevant for captive cheetah time-budget data, where rare or short-duration behaviours can be affected by the observation method and by how visible the animal is during recording (Munita et al., 2016). For these behaviours, the weaker agreement is better read as a combined effect of low frequency, short duration, video visibility and model classification difficulty, rather than as a simple failure of the collar model.

The 'drink' data needs caution in interpreting. In the video data, drinking was recorded only a few times each day and only in the indoor zone. The outdoor water bowl in Zone 2 was not visible in the video footage because it was obscured by vegetation. Therefore, the observed drinking records may only represent the events that were visible to the camera. The weaker collar-video agreement for drink may partly reflect this limitation in the video record, rather than the collar model alone.

The cheetahs used different zones within the enclosure for different behavioural functions. The clearest rest-related feature was the combination of shelter, resting substrate and partial cover in Zone 3. This zone was the main high-use area for both cheetahs and contained a roof overhang, straw-covered ground and lower shrub. The provision of suitable bedding and resting areas is recognised as an important consideration in captive cheetah management (Woc Colburn et al., 2018). Although it was located along the public viewing side, these features may have provided a suitable resting area while still allowing the cheetahs to see across the enclosure. In the wild, cheetahs have been found to select semi-closed habitats, where vegetation may provide concealment from larger predators (Klaassen & Broekhuis, 2018). The combination of shelter and visibility in Zone 3 may have allowed the cheetahs to monitor their surroundings

while resting.

A different function was provided by the raised structure in Zone 6. This zone contained the three-tier vertical stand and was used frequently for 'resting alert' or surveillance behaviour. The use of this area may be partly related to the raised view provided by the stand. Raised areas have been reported as important features in captive felid enclosures, where they can be used for resting and observation (Lyons et al., 1997). Raised areas have also been linked with exploratory behaviour and the use of vantage points to observe the surrounding environment (Quirke & O'Riordan, 2015). In wild cheetahs, elevated landscape features such as rocky outcrops have been linked with a vantage-point function (Pettorelli et al., 2009).

Access and boundary features also appeared to shape enclosure use. Zone 2 connected directly with the indoor area (Zone 1), and it contained a large shrub that the cheetahs could move into. It was close to the boundary fence with adjacent animal areas on the other side. The use of Zone 1 and Zone 2 was linked mainly with 'active' movement, indoor access and boundary monitoring. This interpretation fits with previous work on captive felids, where enclosure edges were described as areas that can provide external visual stimuli (Lyons et al., 1997).

The lower-use areas seemed to have limited behavioural functions. Zone 4 contained a central tree, but the cheetahs mainly used this as a transitional area. Zone 5 was also a boundary and connecting area; although it contained a pergola structure, the cheetahs did not appear to use it as a main resting area, possibly because it did not provide the same combination of shelter, resting substrate and partial cover as Zone 3. Zone 7 was mostly associated with resting alert and may have been used for watching the surrounding area from the enclosure boundary. Zone 8 was also along the boundary, but it was more open and had little cover or structure, which may explain why it had the lowest recorded use. Captive environments are usually smaller and less structurally complex than natural habitats, and previous studies on captive felids have shown that enclosure complexity and structural features can influence activity, resting behaviour and space use (Lyons et al., 1997; Mallapur & Chellam, 2002). In this study, lower-use zones such as Zone 4, Zone 5 and Zone 8 contributed to enclosure complexity by providing movement routes, boundary access and visual cover, even if they were not used as main resting areas. Their role should be interpreted by how they contributed to enclosure function, not only by the percentage of time spent in them.

These enclosure-use patterns show how the cheetahs used the available space. Features that appear attractive or practical from a human perspective were not always the areas most used by the animals. In this study, the highest use was linked with areas that provided shelter,

resting substrate, partial cover or a raised viewing point. This suggests that enclosure design should be based on how animals use different areas, rather than on human-centred design priorities alone (Ross & Lukas, 2006)

3.5 Conclusion

By linking triaxial accelerometer collar-derived behavioural classifications with video-based enclosure-use analysis, this chapter showed how captive cheetah used different parts of their habitat during the matched observation period. The results suggest that enclosure features influenced how the cheetahs used space, with sheltered resting areas and raised viewing points being used most often. This provides useful information for interpreting enclosure function and for understanding how enclosure design may be relevant to captive cheetah welfare. Future research should include more individuals, longer observation periods and improved camera coverage. Longer-term accelerometer monitoring could help detect broad behavioural patterns that are difficult to record continuously through direct observation. It also supports future enclosure-use and welfare assessments

Chapter 4 - General Discussion

The overall objective of this thesis was to develop a triaxial accelerometer approach for long-term behavioural monitoring in captive cheetahs, as a case study on how wearable technologies can be implemented as a captive management tool. This is driven by a key problem in zoo animal welfare assessment: behaviour is an important welfare indicator, but direct observation is limited by observer time, visibility and the difficulty of recording animals continuously over long periods (Melfi, 2009). Previous cheetah studies have shown that accelerometer data can be used to classify behaviour and detect feeding-related events (Giese et al., 2021; McGowan et al., 2022). However, this thesis also shows the challenge of balancing computational performance with biological interpretation. Behaviours with similar movement patterns may need to be grouped to improve model performance, even when they have different biological meanings. This was evident in the present study, where broader categories improved classification stability but also reduced behavioural resolution. Therefore, the selected model needed to balance statistical performance with biologically interpretable behavioural categories. This thesis extends previous work by applying the selected behavioural structure to time-budget analysis and video-based enclosure-use records. In doing so, the accelerometer model was linked with questions about captive cheetah management, habitat use and welfare-related interpretation. The behavioural and spatial records help interpret how Qia and Quartz interacted with their managed environment. The time-budget analysis showed how the cheetahs allocated time among active, inactive and maintenance behaviours.

The enclosure-use analysis showed where these behaviours occurred, and which habitat features were associated with different behavioural functions. Previous studies have shown that space use can help assess enclosure function and inform welfare interpretation (Brereton, 2020; Ross & Lukas, 2006). Studies on captive felids have also shown that enclosure design, structural complexity and specific features can influence ‘activity’, ‘resting behaviour’ and space use (Lyons et al., 1997; Mallapur et al., 2002). Areas used mainly for ‘resting alert’ or ‘lying asleep’ may indicate suitable sheltering locations, while areas used for active behaviour may be more related to movement routes, routine husbandry or access between enclosure features. The use of raised viewing points also fits with cheetah-specific evidence that raised areas can support exploratory behaviour and provide vantage points for observing the surrounding environment (Quirke & O'Riordan, 2015).

The findings also indicate several areas where this approach could be further developed. The study was based on two adult female cheetahs housed in one zoo environment. This

provided a controlled setting for developing and applying the collar-based behavioural model, while future studies could expand the dataset by including more individuals, different age groups, different sex classes and cheetahs housed in different zoo environments. This would help determine how far one behavioural model can be applied across individuals, and whether small adjustments are needed for different animals or institutions. Longer monitoring periods would further improve the value of this approach. Behavioural patterns may change with weather conditions, daylight duration, enrichment schedules, husbandry routines and health status. By collecting collar data over longer periods, future studies could develop normal behavioural baselines for each animal and identify changes in activity, resting, feeding-related behaviour or enclosure use over time.

The collar-based triaxial accelerometer classification model developed in this thesis provides a basis for recording behavioural patterns across longer periods. This would not replace keeper knowledge or direct observation, but it could help zoo staff identify changes in activity, resting, feeding-related behaviour or daily rhythms that may not be obvious from limited observation alone. This approach may also have practical value for zoo management. Long-term behavioural records could help zoos make more evidence-based decisions about enclosure design, enrichment schedules and husbandry changes. For example, if the data show that a certain enclosure feature is rarely used, or that activity increases after a particular enrichment programme, management decisions can be based on measured behavioural patterns rather than assumption. This could help zoos use staff time and management resources more effectively while supporting welfare assessment in a more systematic way.

Future research could expand this approach by testing the Model 5 collar-based triaxial accelerometer classification model with more individuals, longer monitoring periods and other zoo environments. This would help assess how well the model works between individuals and institutions. Similar methods could also be applied to other felid species or other captive mammals. The main value is not only the specific Model 5 categories used here, but also the process of combining video-labelled behaviours, accelerometer data and machine learning to build a practical monitoring model. This would make the method useful for wider zoo welfare research, especially where continuous direct observation is difficult.

In the longer term, this approach could also contribute to monitoring non-captive animals. Direct observation is often more difficult in free-ranging animals, and GPS collars mainly show where an animal moves rather than what it is doing. Accelerometers can add behavioural information to movement data by identifying activity, resting and other broad behavioural states. Models developed in zoos would need further testing before field application because

wild animals experience more variable habitats and behaviours. Therefore, captive studies provide a useful starting point because behaviours can be labelled against video observations under more controlled conditions before being applied more widely in the field.

4.1 Conclusions

Welfare assessment should consider both behaviour and the physical environment. Although this study did not directly measure welfare, the behavioural and spatial records help interpret how the captive cheetahs interacted with different enclosure features. Repeated use of certain areas for resting, sleeping, observing the surroundings or moving through the habitat suggests that these features may support different behavioural functions. Triaxial accelerometer behavioural monitoring and enclosure-use analysis can help identify how captive animals use their habitat and which features may be relevant to enclosure design, husbandry decisions and longer-term welfare monitoring.

5.0 References

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, *49*(3-4), 227-266.

Backues, K., Clyde, V., Denver, M., Fiorello, C., Hilsenroth, R., Lamberski, N., Larson, S., Meehan, T., Murray, M., & Ramer, J. (2011). Guidelines for zoo and aquarium veterinary medical programs and veterinary hospitals. *Journal of Zoo and Wildlife Medicine*, *42*(1), 176-192.

Bashaw, M. J., Bloomsmith, M. A., Marr, M., & Maple, T. L. (2003). To hunt or not to hunt? A feeding enrichment experiment with captive large felids. *Zoo Biology*, *22*(2), 189-198.

Bashaw, M. J., Kelling, A. S., Bloomsmith, M. A., & Maple, T. L. (2007). Environmental effects on the behavior of zoo-housed lions and tigers, with a case study to the effects of a visual barrier on pacing. *Journal of Applied Animal Welfare Science*, *10*(2), 95-109.

Beltrán, J. F., & Delibes, M. (1994). Environmental determinants of circadian activity of free-ranging Iberian lynxes. *Journal of Mammalogy*, *75*(2), 382-393.

Benkendorf, D. J., Schwartz, S. D., Cutler, D. R., & Hawkins, C. P. (2023). Correcting for the effects of class imbalance improves the performance of machine-learning based species distribution models. *Ecological Modelling*, *483*, 110414.

Binding, S., Farmer, H., Krusin, L., & Cronin, K. (2020). Status of animal welfare research in zoos and aquariums: Where are we, where to next? *Journal of Zoo and Aquarium Research*, *8*(3), 166-174.

Boere, V. (2001). Environmental enrichment for neotropical primates in captivity. *Ciência Rural*, *31*, 543-551.

Bond, J. C., & Lindburg, D. G. (1990). Carcass feeding of captive cheetahs (*Acinonyx jubatus*): the effects of a naturalistic feeding program on oral health and psychological well-being. *Applied Animal Behaviour Science*, *26* (4), 373-382.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5-32.

Brereton, J. E. (2020). Directions in animal enclosure use studies. *Journal of Zoo and Aquarium Research*, *8*(1), 1-9.

Broom, D. M. (1986). Indicators of poor welfare. *British Veterinary Journal*, *142*(6), 524-526.

Broom, D. M. (2002). Welfare in wildlife management and zoos. *Advances in Ethology, Berlin*, *37*, 4-6.

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.

Cadenas, J. M., Garrido, M. C., & Martínez, R. (2013). Feature subset selection filter–wrapper based on low quality data. *Expert Systems with Applications*, *40*(16), 6241-6252.

Cederlund, J. (2018). *Behavioural responses of Amur tigers (Panthera tigris altaica) and African lions (Panthera leo) to conspecific urine and to a component of*

tiger marking fluid [Master's thesis, Linköping University].

Chadwick, C. L., Rees, P. A., & Stevens-Wood, B. (2013). Captive-housed male cheetahs (*Acinonyx jubatus soemmeringii*) form naturalistic coalitions: Measuring associations and calculating chance encounters. *Zoo Biology*, *32*(5), 518-527.

Chamove, A. S., Hosey, G. R., & Schaetzel, P. (1988). Visitors excite primates in zoos. *Zoo Biology*, *7*(4), 359-369.

Clubb, R., & Mason, G. (2003). Captivity effects on wide-ranging carnivores. *Nature*, *425*(6957), 473-474.

Coughlin, C. E., & Van Heezik, Y. (2015). Weighed down by science: do collar-mounted devices affect domestic cat behaviour and movement? *Wildlife Research*, *41*(7), 606-614.

Cuff, A. R., Sparkes, E. L., Randau, M., Pierce, S. E., Kitchener, A. C., Goswami, A., & Hutchinson, J. R. (2016). The scaling of postcranial muscles in cats (*Felidae*) I: forelimb, cervical, and thoracic muscles. *Journal of Anatomy*, *229*(1), 128-141.

Davey, G. (2005). The visitor effect. *Zoos'Print Journal*, *20*(6), 1900-1903.

De Rouck, M., Kitchener, A., Law, G., & Nelissen, M. (2005). A comparative study of the influence of social housing conditions on the behaviour of captive tigers (*Panthera tigris*). *Animal Welfare*, *14*(3), 229-238.

Den Uijl, I., Gómez Álvarez, C. B., Bartram, D., Dror, Y., Holland, R., & Cook, A. (2017). External validation of a collar-mounted triaxial accelerometer for second-by-second monitoring of eight behavioural states in dogs. *PLOS ONE*, *12*(11), e0188481.

Estevez, I., & Christman, M. C. (2006). Analysis of the movement and use of space of animals in confinement: The effect of sampling effort. *Applied Animal Behaviour Science*, *97*(2-4), 221-240.

European Zoo Nutrition Centre. (2010). *European Zoo Nutrition Centre feeding guidelines.*http://www.eznc.org/PrimoSite/show.do?ctx=7795,21671&anav=21550#21698

Farm Animal Welfare Council. (1993). *Report on Priorities for Animal Welfare Research and Development*. London, UK Retrieved from https://edepot.wur.nl/134980

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11), 1325-1330.

Giese, L., Melzheimer, J., Bockmühl, D., Wasiolka, B., Rast, W., Berger, A., & Wachter, B. (2021). Using machine learning for remote behaviour classification-Verifying acceleration data to infer feeding events in free-ranging cheetahs. *Sensors*, *21*(16), 5426.

Green, T., & Mellor, D. J. (2011). Extending ideas about animal welfare assessment to include ‘quality of life’ and related concepts. *New Zealand Veterinary Journal*, *59*(6), 263-271.

Gregory, N. G. (2004). Emotional numbness and deprivation. In UFAW animal welfare

series: *Physiology and Behaviour of Animal Suffering* (pp. 32-45). Oxford, United Kingdom: Blackwell Science.

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.

Herbers, J. M. (1981). Time resources and laziness in animals. *Oecologia*, *49*(2), 252-262.

Hetem, R. S., Mitchell, D., de Witt, B. A., Fick, L. G., Maloney, S. K., Meyer, L. C. R., & Fuller, A. (2019). Body temperature, activity patterns and hunting in free-living cheetah: Biologging reveals new insights. *Integrative Zoology*, *14*(1), 30-47.

Hounslow, J., Brewster, L., Lear, K., Guttridge, T., Daly, R., Whitney, N., & Gleiss, A. (2019). Assessing the effects of sampling frequency on behavioural classification of accelerometer data. *Journal of Experimental Marine Biology and Ecology*, *512*, 22-30.

Hunter, S. C., Gusset, M., Miller, L. J., & Somers, M. J. (2014). Space use as an indicator of enclosure appropriateness in African wild dogs (*Lycaon pictus*). *Journal of Applied Animal Welfare Science*, *17*(2), 98-110.

Jackson, P., & Nowell, K. (1996). *Wild Cats: Status Survey and Conservation Action Plan*. IUCN.

Jensen, S. P., Gray, S. J., & Hurst, J. L. (2003). How does habitat structure affect activity and use of space among house mice? *Animal Behaviour*, *66*(2), 239-250.

Kershaw, T., Hall, E. J., Dobbs, P., Liptovszky, M., & Strong, V. (2020). An exploration of the value of elective health checks in UK zoo-housed gibbons. *Animals*, *10*(12), 2307.

Klaassen, B., & Broekhuis, F. (2018). Living on the edge: Multiscale habitat selection by cheetahs in a human-wildlife landscape. *Ecology and Evolution*, *8*(15), 7611-7623.

Kleiman, D. G. (1994). Foreword: Animal behavior studies and zoo propagation programs. *Zoo Biology*, *13*(5), 411-412.

Koester, D. C., Freeman, E. W., Brown, J. L., Wildt, D. E., Terrell, K. A., Franklin, A. D., & Crosier, A. E. (2015). Motile sperm output by male cheetahs (*Acinonyx jubatus*) managed ex situ is influenced by public exposure and number of care-givers. *PLOS ONE*, *10*(9), e0135847.

Kumpulainen, P., Cardó, A. V., Somppi, S., Törnqvist, H., Väätäjä, H., Majaranta, P., Gizatdinova, Y., Antink, C. H., Surakka, V., & Kujala, M. V. (2021). Dog behaviour classification with movement sensors placed on the harness and the collar. *Applied Animal Behaviour Science*, *241*, 105393.

Law, G., Macdonald, A., & Reid, A. (1997). Dispelling some common misconceptions about the keeping of felids in captivity. *International Zoo Yearbook*, *35*(1), 197-207.

Lindburg, D. G. (1988). Improving the feeding of captive felines through application of field data. *Zoo Biology*, *7*(3), 211-218.

Lowe, L. R. (2013). *Stereotypic behaviour of captive New Zealand zoo mammals: Attitudes of zoo staff, prevalence, and effectiveness of short-term environmental enrichment* [Master's thesis, Massey University].

Lyons, J., Young, R. J., & Deag, J. M. (1997). The effects of physical characteristics of the environment and feeding regime on the behavior of captive felids. *Zoo Biology*, *16*(1), 71-83.

Mallapur, A., & Chellam, R. (2002). Environmental influences on stereotypy and the activity budget of Indian leopards (*Panthera pardus*) in four zoos in Southern India. *Zoo Biology*, *21*(6), 585-595.

Mallapur, A., Qureshi, Q., & Chellam, R. (2002). Enclosure design and space utilization by Indian leopards (*Panthera pardus*) in four zoos in southern India. *Journal of Applied Animal Welfare Science*, *5*(2), 111-124.

Manteca, X., Amat, M., Salas, M., & Temple, D. (2016). Animal-based indicators to assess welfare in zoo animals. *CABI Reviews*, *11*(010), 1-10.

Markowitz, H., Aday, C., & Gavazzi, A. (1995). Effectiveness of acoustic "prey": Environmental enrichment for a captive African leopard (*Panthera pardus*). *Zoo Biology*, *14*(4), 371-379.

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.

Mason, G. J. (1991). Stereotypies: a critical review. *Animal Behaviour*, *41*(6), 1015-1037.

McClune, D. W., Marks, N. J., Wilson, R. P., Houghton, J. D., Montgomery, I. W., McGowan, N. E., Gormley, 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.

McCulloch, S. P. (2013). A critique of FAWC's five freedoms as a framework for the analysis of animal welfare. *Journal of Agricultural and Environmental Ethics*, *26*(5), 959-975.

McGowan, N. E., Marks, N. J., Maule, A. G., Schmidt-Küntzel, A., Marker, L. L., & Scantlebury, D. M. (2022). Categorising cheetah behaviour using tri-axial accelerometer data loggers: a comparison of model resolution and data logger performance. *Movement Ecology*, *10*(1), 7.

McPhee, M. E. (2002). Intact carcasses as enrichment for large felids: Effects on on- and off-exhibit behaviors. *Zoo Biology*, *21*(1), 37-47.

Melfi, V. (2009). There are big gaps in our knowledge, and thus approach, to zoo animal welfare: a case for evidence-based zoo animal management. *Zoo Biology*, *28*(6), 574-588.

Mellen, J. D. (1991). Factors influencing reproductive success in small captive exotic felids (*Felis* spp.): a multiple regression analysis. *Zoo Biology*, *10*(2), 95-110.

Mellen, J. D., Hayes, M. P., & Shepherdson, D. J. (1998). Captive environments for small felids. In D. J. Shepherdson, J. D. Mellen, & M. Hutchins (Eds.), *Second nature: Environmental enrichment for captive animals* (pp. 184-201).

Smithsonian Institution Press.

Mellen, J. D., & Shepherdson, D. J. (1997). Environmental enrichment for felids: an integrated approach. *International Zoo Yearbook*, *35*(1), 191-197.

Mellor, D., Patterson-Kane, E., & Stafford, K. J. (2009). *The sciences of animal welfare*. Chichester, United Kingdom: Wiley-Blackwell.

Mellor, D. J. (2012). Affective states and the assessment of laboratory-induced animal welfare impacts. *Altex Proc*, *1*, 445-449.

Mellor, D. J. (2016). Moving beyond the "five freedoms" by updating the "five provisions" and introducing aligned "animal welfare aims". *Animals*, *6*(10), 59.

Mellor, D. J., & Beausoleil, N. J. (2015). Extending the 'Five Domains' model for animal welfare assessment to incorporate positive welfare states. *Animal Welfare*, *24*(3), 241-253.

Mellor, D. J., Beausoleil, N. J., Littlewood, K. E., McLean, A. N., McGreevy, P. D., Jones, B., & Wilkins, C. (2020). The 2020 five domains model: Including human–animal interactions in assessments of animal welfare. *Animals*, *10*(10), 1870.

Mellor, D. J., & Reid, C. S. W. (1994). Concepts of animal well-being and predicting the impact of procedures on experimental animals. In R. Baker, G. Jenkin, & D. J. Mellor (Eds.), *Improving the well-being of animals in the research environment* (pp. 3-18). Australian and New Zealand Council for the Care of Animals in Research and Teaching.

Mench, J. A., & Mason, G. J. (1997). Behaviour. In M. C. Appleby & B. O. Hughes (Eds.), *Animal welfare* (pp. 127-142). CAB International.

Miller, L., Pisacane, C., & Vicino, G. (2016). Relationship between behavioural diversity and faecal glucocorticoid metabolites: A case study with cheetahs (*Acinonyx jubatus*). *Animal Welfare*, *25*(3), 325-329.

Miller, L. J., Vicino, G. A., Sheftel, J., & Lauderdale, L. K. (2020). Behavioral diversity as a potential indicator of positive animal welfare. *Animals*, *10*(7), 1211.

Mills, M. G. L., & Mills, M. E. J. (2017). Twenty-four-hour activity patterns and distances moved. In *Kalahari cheetahs: Adaptations to an arid region* (pp. 65-74). Oxford, United Kingdom: Oxford University Press.

Mitchell, E. P., Prozesky, L., & Lawrence, J. (2018). A new perspective on the pathogenesis of chronic renal disease in captive cheetahs (*Acinonyx jubatus*). *PLOS ONE*, *13*(3), e0194114.

Mohapatra, R. K., Panda, S., & Acharya, U. R. (2014). Study on activity pattern and incidence of stereotypic behavior in captive tigers. *Journal of Veterinary Behavior*, *9*(4), 172-176.

Moreira, N., Brown, J. L., Moraes, W., Swanson, W. F., & Monteiro-Filho, E. (2007). Effect of housing and environmental enrichment on adrenocortical activity, behavior and reproductive cyclicity in the female tigrina (*Leopardus tigrinus*) and margay (*Leopardus wiedii*). *Zoo Biology*, *26*(6), 441-460.

Morgan, K. N., & Tromborg, C. T. (2007). Sources of stress in captivity. *Applied Animal Behaviour Science*, *102*(3-4), 262-302.

Munita, C., Tadich, T. A., & Briceño, C. (2016). Comparison of 2 behavioral sampling methods to establish a time budget in a captive female cheetah (*Acinonyx jubatus*). *Journal of Veterinary Behavior*, *13*, 1-5.

Pettorelli, N., Hilborn, A., Broekhuis, F., & Durant, S. (2009). Exploring habitat use by cheetahs using ecological niche factor analysis. *Journal of Zoology*, *277*(2), 141-148.

Powell, D. (1995). Preliminary evaluation of environmental enrichment techniques for African lions (*Panthera leo*). *Animal Welfare*, *4*(4), 361-370.

Quirke, T., & O'Riordan, R. M. (2011a). The effect of different types of enrichment on the behaviour of cheetahs (*Acinonyx jubatus*) in captivity. *Applied Animal Behaviour Science*, *133*(1-2), 87-94.

Quirke, T., & O'Riordan, R. M. (2011b). The effect of a randomised enrichment treatment schedule on the behaviour of cheetahs (*Acinonyx jubatus*). *Applied Animal Behaviour Science*, *135*(1-2), 103-109.

Quirke, T., & O'Riordan, R. (2015). An investigation into the prevalence of exploratory behavior in captive cheetahs (*Acinonyx jubatus*). *Zoo Biology*, *34*(2), 130-138.

Quirke, T., O'Riordan, R., & Davenport, J. (2013). A comparative study of the speeds attained by captive cheetahs during the enrichment practice of the "cheetah run". *Zoo Biology*, *32*(5), 490-496.

Quirke, T., O'Riordan, R. M., & Zuur, A. (2012). Factors influencing the prevalence of stereotypical behaviour in captive cheetahs (*Acinonyx jubatus*). *Applied Animal Behaviour Science*, *142*(3-4), 189-197.

Reade, L. S., & Waran, N. K. (1996). The modern zoo: How do people perceive zoo animals? *Applied Animal Behaviour Science*, *47*(1-2), 109-118.

Redmond, C., Smit, M., Draganova, I., Corner-Thomas, R., Thomas, D., & Andrews, C. (2024). The use of triaxial accelerometers and machine learning algorithms for behavioural identification in domestic dogs (*Canis familiaris*): A validation study. *Sensors*, *24*(18), 5955.

Rees, P. A. (2011). Nutrition and feeding. In *An introduction to zoo biology and management* (pp. 151-172). Wiley-Blackwell.

Rochlitz, I. (1999). Recommendations for the housing of cats in the home, in catteries and animal shelters, in laboratories and in veterinary surgeries. *Journal of Feline Medicine and Surgery*, *1*(3), 181-191.

Rose, P. E., & Riley, L. M. (2021). Conducting behavioural research in the zoo: A guide to ten important methods, concepts and theories. *Journal of Zoological and Botanical Gardens*, *2*(3), 421-444.

Ross, S. R., & Lukas, K. E. (2006). Use of space in a non-naturalistic environment by chimpanzees (*Pan troglodytes*) and lowland gorillas (*Gorilla gorilla gorilla*). *Applied Animal Behaviour Science*, *96*(1-2), 143-152.

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, *10*(3), e0118432.

Sakai, K., Oishi, K., Miwa, M., Kumagai, H., & Hirooka, H. (2019). Behavior classification of goats using 9-axis multi sensors: The effect of imbalanced

datasets on classification performance. *Computers and Electronics in Agriculture*, *166*, 105027.

Salas, M., Tallo-Parra, O., & Manteca, X. (2024). Evidence-based zoo animal welfare assessment: Putting science into practice. *Journal of Zoo and Aquarium Research*, *12*(4), 205-211.

Sellinger, R. L., & Ha, J. C. (2005). The effects of visitor density and intensity on the behavior of two captive jaguars (*Panthera onca*). *Journal of Applied Animal Welfare Science*, *8*(4), 233-244.

Seyrling, I., Dierkes, P. W., & Burger, A. L. (2022). Diurnal and nocturnal behaviour of cheetahs (*Acinonyx jubatus*) and lions (*Panthera leo*) in zoos. *Animals*, *12*(18), 2367.

Shepard, E. L., Wilson, R. P., Quintana, F., Laich, A. G., Liebsch, N., Albareda, D. A., Halsey, L. G., Gleiss, A., Morgan, D. T., & Myers, A. E. (2008). Identification of animal movement patterns using tri-axial accelerometry. *Endangered Species Research*, *10*, 47-60.

Shepherdson, D. J., Carlstead, K., Mellen, J. D., & Seidensticker, J. (1993). The influence of food presentation on the behavior of small cats in confined environments. *Zoo Biology*, *12*(2), 203-216.

Shyne, A. (2006). Meta-analytic review of the effects of enrichment on stereotypic behavior in zoo mammals. *Zoo Biology*, *25*(4), 317-337.

Skibiel, A. L., Trevino, H. S., & Naugher, K. (2007). Comparison of several types of enrichment for captive felids. *Zoo Biology*, *26*(5), 371-381.

Smit, M., Ikurior, S. J., Corner-Thomas, R. A., Andrews, C. J., Draganova, I., & Thomas, D. G. (2023). The use of triaxial accelerometers and machine learning algorithms for behavioural identification in domestic cats (*Felis catus*): A validation study. *Sensors*, *23*(16), 7165.

Stanton, L. A., Sullivan, M. S., & Fazio, J. M. (2015). A standardized ethogram for the felidae: A tool for behavioral researchers. *Applied Animal Behaviour Science*, *173*, 3-16.

Suárez, P., Recuerda, P., & Arias-de-Reyna, L. (2017). Behaviour and welfare: The visitor effect in captive felids. *Animal Welfare*, *26*(1), 25-34.

Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour*, *124*, 203-220.

Webster, J. (2008). Animal Welfare: limping towards Eden: A practical approach to redressing the problem of our dominion over the animals. John Wiley & Sons.

Weller, S. H., & Bennett, C. L. (2001). Twenty-four hour activity budgets and patterns of behavior in captive ocelots (*Leopardus pardalis*). *Applied Animal Behaviour Science*, *71*(1), 67-79.

Wielebnowski, N. C., Fletchall, N., Carlstead, K., Busso, J. M., & Brown, J. L. (2002). Noninvasive assessment of adrenal activity associated with husbandry and behavioral factors in the North American clouded leopard population. *Zoo Biology*, *21*(1), 77-98.

Wilson, A. M., Lowe, J., Roskilly, K., Hudson, P. E., Golabek, K., & McNutt, J. (2013).

Locomotion dynamics of hunting in wild cheetahs. *Nature*, *498*(7453), 185-189.

Wilson, J. W., Mills, M. G., Wilson, R. P., Peters, G., Mills, M. E., Speakman, J. R., Durant, S. M., Bennett, N. C., Marks, N. J., & Scantlebury, M. (2013). Cheetahs, *Acinonyx jubatus*, balance turn capacity with pace when chasing prey. *Biology Letters*, *9*(5), 20130620.

Woc Colburn, A. M., Sanchez, C. R., Citino, S., Crosier, A. E., Murray, S., Kaandorp, J., Kaandorp, C., & Marker, L. (2018). Clinical management of captive cheetahs. In P. J. Nyhus (Ed.), *Cheetahs: Biology and conservation* (pp. 335–347). Academic Press.

Würbel, H. (2001). Ideal homes? Housing effects on rodent brain and behaviour. *Trends in Neurosciences*, *24*(4), 207-211.

Zoo and Aquarium Association Australasia. (2025). *The five domains.* https://www.zooaquarium.org.au/public/Public/Animal-Welfare/The-Five-Domains.aspx?hkey=33869d99-0a7c-4616-823b-2032848d0829

Appendix

Appendix 1. Description of identifier variables used for model building.

Feature	#	Description	Formula/calculation
Descriptive statistics (1s)			
Mean	4	Mean value calculated for each second using the raw acceleration data (30 Hz). Where x = X, Y, Z, and VM.	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
Sum	4	Sum of the acceleration values within each second. Where x = X, Y, Z, and VM.	$\sum_{i=1}^{n} x_i$
Minimum (min)	4	Lowest value recorded within each second. Where x = X, Y, Z, and VM.	Min(x)
Maximum (Max)	4	Highest value recorded within each second. Where x = X, Y, Z, and VM.	Max(x)
Standard deviation (SD)	4	Describes the amount of variation within the data. Where x = X, Y, Z, and VM.	$\text{SD} = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n = 1}}$
Skewness (Skew)	4	Indicates the asymmetry of the probability distribution. Where x = X, Y, Z, and VM.	$\text{Skew} = \frac{E[(x-\mu)^3]}{\sigma^3}$
Kurtosis (Kurt)	4	Indicates the relative weight of the tails compared with a normal distribution. Where x = X, Y, Z, and VM.	$\text{Kurt} = \frac{E[(x-\mu)^4]}{\sigma^4}$
Derived kinematic features			
Vector magnitude (VM)	0	Combined magnitude of acceleration across all three axes.	$VM = \sqrt{X^2 + Y^2 + Z^2}$
Static acceleration (X_g, Y_g, Z_g)	0	Gravity component estimated using a centred moving average window (~1 s)	$Axis_g = rollmean(x)$
Dynamic body acceleration	0	Dynamic acceleration after removal of the gravity component.	$Axis_{DBA} = Axis - Axis_g$
Overall dynamic body acceleration (ODBA)	1	Sum of absolute dynamic acceleration across axes.	$ODBA = \sum_{i=1}^{n} \lvert DBA_X \rvert + \lvert DBA_Y \rvert + \lvert DBA_Z \rvert$
Root means square DBA (DBA_{RMS})	1	Root mean square of the dynamic acceleration across axes.	$DBA_{RMS} = \sqrt{X_{DBA}^2 + Y_{DBA}^2 + Z_{DBA}^2}$
Signal magnitude area (SMA)	1	Sum of absolute acceleration values across all axes.	
Gravity / posture features			
Roll (mean)	1	Rotation around the longitudinal axis derived from static acceleration	$Roll = \arctan\left(\frac{-X_g}{\sqrt{Y_g^2 + Z_g^2}}\right)$
Pitch (mean)	1	Rotation around the lateral axis derived from static acceleration	$Pitch = \arctan\left(\frac{Y_g}{Z_g}\right)$
Tilt (mean)	1	Overall inclination angle relative to gravity	$Tilt = \arccos\left(\frac{Z_g}{\sqrt{X_g^2 + Y_g^2 + Z_g^2}}\right)$
Roll speed	1	Rate of change in roll angle between consecutive samples.	$Rollspeed = diff(Roll)$
Pitch speed	1	Rate of change in pitch angle between consecutive samples.	$Pitchspeed = diff(Pitch)$
Tilt speed	1	Rate of change in tilt angle between consecutive samples.	$Tiltspeed = diff(Tilt)$
Inter-axis relationships			
Correlation XY	1	Correlation between axes	$r_{XY} = cor(X, Y)$
Correlation XZ	1	Correlation between axes	$r_{XZ} = cor(X, Z)$
Correlation YZ	1	Correlation between axes	$r_{YZ} = cor(Y, Z)$
Rolling averages			
Rolling ODBA (3s)	1	Three-second centred rolling mean of ODBA.	$Rolling\ ODBA_{3s} = mean\ (ODBA_{W3s})$
Rolling ODBA (5s)	1	Five-second centred rolling mean of ODBA.	$Rolling\ ODBA_{5s} = mean\ (ODBA_{W5s})$
Rolling VM_SD (3s)	1	Three-second centred rolling mean of vector magnitude variability.	$Rolling\ VM_SD_{3s} = SD\ (VM_{W3s})$
Rolling Roll_SD (3s)	1	Standard deviation of roll angle across a 3-second centred window.	$Rolling\ Roll_SD_{3s} = SD\ (Roll_{W3s})$
Rolling Pitch_SD (3s)	1	Standard deviation of pitch angle across a 3-second centred window.	$Rolling\ Pitch_SD_{3s} = SD\ (Pitch_{W3s})$

Appendix 2. Confusion matrix showing the percentage of predicted and observed behaviours for Model 1. Correct classifications are highlighted in yellow, while red cells indicate misclassifications greater than 15%.

Model 1																							
Prediction	**Observed behaviour**																						
	Allogroom	**Crouching**	**Drink**	**Eat**	**Groom**	**Jump/climb**	**Lying alert L**	**Lying alert R**	**Lying alert ventral**	**Lying asleep L**	**Lying asleep R**	**Raking**	**Roll**	**Run**	**Scent mark**	**Scratch**	**Shake**	**Sit**	**Sniff**	**Stand**	**Stretch**	**Toileting**	**Trot**	**Walk**
Allogroom	82.45	0.00	0.00	21.43	24.94	0.00	1.05	0.99	2.53	0.03	0.04	50.00	1.20	0.00	7.78	0.00	3.85	2.46	11.30	4.69	19.57	0.00	0.00	0.61
Crouching	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Drink	0.00	0.00	77.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Eat	1.43	0.00	0.00	61.90	0.73	4.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.68	0.00	2.17	0.00	0.00	0.02
Groom	2.04	0.00	0.00	2.98	56.66	0.00	0.05	0.01	0.00	0.01	0.01	0.00	1.20	0.00	0.00	0.00	0.00	0.23	3.77	0.18	2.17	0.00	0.00	0.06
Jump/climb	0.00	0.00	0.00	0.00	0.00	8.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.61	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.00	0.00	0.01
Lying alert L	3.20	0.00	0.00	0.00	3.63	4.17	89.03	7.52	9.33	0.44	0.18	0.00	8.43	0.00	6.67	0.00	15.38	10.27	1.71	13.43	6.52	14.00	0.00	0.11
Lying alert R	2.65	0.00	5.00	0.00	0.24	0.00	5.74	85.89	6.27	0.21	0.67	0.00	6.02	1.61	3.33	50.00	0.00	6.13	2.74	5.89	6.52	0.00	0.00	0.07
Lying alert ventral	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.15	69.73	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.55	0.34	0.64	0.00	4.00	0.00	0.00
Lying asleep L	0.14	0.00	0.00	0.00	0.24	0.00	0.83	0.18	0.13	98.72	0.29	0.00	1.20	0.00	0.00	0.00	0.00	0.39	0.00	0.28	0.00	2.00	0.00	0.00
Lying asleep R	0.07	0.00	0.00	0.60	0.24	0.00	0.14	1.56	0.13	0.48	98.66	0.00	0.00	0.00	0.00	0.00	0.00	0.04	1.37	0.28	0.00	0.00	0.00	0.00
Raking	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Roll	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.14	0.00	0.00	0.00	0.00	66.27	1.61	0.00	25.00	0.00	0.00	0.00	0.00	0.00	0.00	2.50	0.04
Run	0.00	0.00	0.00	0.00	0.00	20.83	0.00	0.00	0.00	0.00	0.00	0.00	2.41	64.52	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	15.00	0.05
Scent.mark	0.27	33.33	0.00	0.00	0.24	0.00	0.01	0.03	0.00	0.02	0.00	0.00	0.00	0.00	28.89	0.00	0.00	0.12	0.68	1.10	0.00	14.00	0.00	0.13
Scratch	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Shake	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sit	2.45	0.00	0.00	0.00	0.97	0.00	1.93	2.31	9.07	0.04	0.03	16.67	1.20	0.00	6.67	0.00	3.85	73.98	1.71	10.76	6.52	16.00	0.00	0.05
Sniff	0.68	33.33	7.50	0.60	2.18	0.00	0.09	0.06	0.27	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.51	36.64	1.47	0.00	0.00	0.00	0.17
Stand	1.36	33.33	2.50	0.60	3.87	8.33	0.50	0.77	2.27	0.04	0.06	0.00	0.00	0.00	18.33	0.00	7.69	3.28	3.42	43.79	13.04	4.00	0.00	0.57
Stretch	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00
Toileting	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.00	0.18	0.00	42.00	0.00	0.00
Trot	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	19.35	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.00	55.00	0.07
Walk	3.27	0.00	7.50	11.90	6.05	54.17	0.38	0.38	0.13	0.01	0.02	33.33	12.05	11.29	28.33	25.00	69.23	1.84	35.62	16.74	43.48	4.00	27.50	98.00
Total Observations(s)	1470	3	40	168	413	24	7973	7101	750	19237	15777	6	83	62	180	4	26	2560	292	1087	46	50	40	9769

Appendix 3. Confusion matrix showing the percentage of predicted and observed behaviours for Model 2. Correct classifications are highlighted in yellow, while red cells indicate misclassifications greater than 15%.

Model 2																
Prediction	**Observed behaviour**															
	Allogroom	**Drink**	**Eat**	**Groom**	**Jump/climb**	**Lying alert**	**Lying sleep**	**Roll**	**Run**	**Scent mark**	**Sit**	**Sniff**	**Stand**	**Toileting**	**Trot**	**Walk**
Allogroom	80.54	0.00	25.00	24.21	0.00	0.71	0.01	1.20	0.00	7.22	2.30	8.22	5.89	2.00	0.00	0.68
Drink	0.00	87.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.18	0.00	0.00	0.00
Eat	1.50	0.00	63.69	1.21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.37	0.00	0.00	0.00	0.03
Groom	1.09	0.00	4.17	52.78	0.00	0.04	0.01	0.00	0.00	0.56	0.08	2.40	0.09	0.00	0.00	0.08
Jump/climb	0.00	0.00	0.00	0.00	8.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lying alert	10.14	5.00	0.00	9.44	8.33	96.36	0.98	15.66	1.61	21.67	31.29	7.53	39.19	30.00	0.00	0.45
Lying sleep	0.20	0.00	0.00	0.00	0.00	1.26	98.95	0.00	0.00	0.00	0.39	2.40	0.18	0.00	0.00	0.00
Roll	0.00	0.00	0.00	0.24	0.00	0.06	0.00	59.04	0.00	0.00	0.00	0.00	0.09	0.00	2.50	0.02
Run	0.00	0.00	0.00	0.00	37.50	0.00	0.00	3.61	75.81	0.00	0.00	0.00	0.09	0.00	15.00	0.06
Scent mark	0.27	0.00	0.00	0.24	0.00	0.01	0.00	0.00	0.00	30.56	0.20	0.00	0.74	0.00	0.00	0.08
Sit	1.63	0.00	0.00	0.73	0.00	0.97	0.01	0.00	0.00	2.78	61.21	0.68	7.36	32.00	0.00	0.06
Sniff	0.61	2.50	1.19	2.18	0.00	0.04	0.00	0.00	0.00	0.00	0.43	40.41	1.93	0.00	0.00	0.19
Stand	1.16	0.00	0.00	3.15	12.50	0.27	0.02	0.00	0.00	11.11	2.11	5.48	28.52	6.00	0.00	0.48
Toileting	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.08	0.00	0.18	22.00	0.00	0.00
Trot	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.06	0.00	0.00	0.00	0.09	0.00	47.50	0.12
Walk	2.86	5.00	5.95	5.81	33.33	0.27	0.01	20.48	14.52	26.11	1.91	31.16	15.46	8.00	35.00	97.74
Total Observations(s)	1470	40	168	413	24	15825	35014	83	62	180	2560	292	1087	50	40	9769