Importance of group dynamics on female mating behaviour in reindeer *Rangifer tarandus*

Guillaume Body

A Thesis
In the Department
Of
Biology

Presented in Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy (Biology) at
Concordia University
Montreal, Québec, Canada

2014

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: **Guillaume Body**

Entitled: **Importance of group dynamics on female mating behaviour in reindeer *Rangifer tarandus***

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee

____ Chair
Dr L. Kalman

____ External examiner
Dr D. Réale

____ External to Program
Dr J. Pfaus

____ Examiner
Dr G. Brown

____ Examiner
Dr J. Grant

____ Thesis Supervisor
Dr R. Weladji

Approved by ____
Dr S. Dayanandan, Graduate Program Director

February 25, 2014 ____
Professor J. Locke, Interim Dean, Faculty of Arts and Science

ABSTRACT

Importance of group dynamics on female mating behaviour in reindeer *Rangifer tarandus*

Guillaume Body, Ph.D.

Concordia University, 2014

The aim of this work is to investigate female mating behaviour in reindeer *Rangifer tarandus* using a combination of both field and GPS based data. By developing adequate methodology, I was able to remotely estimate individual activity budgets, and to investigate fission-fusion group dynamics in a controlled herd. Variations of the group dynamics revealed an increase in group size at the beginning of the peak rut, resulting from an increase of group cohesiveness. I demonstrated that this increase in cohesiveness resulted from the herding behaviour of males, rather than female mate choice. Harassment avoidance marginally increased the group cohesiveness, as females may have preferred to remain within the group to avoid the costly solitary situation where they were harassed by low quality males during the peak rut. However, forming larger groups appeared to be costly, since increasing group size increased the disturbance level on female activities due to foraging competition. Therefore, I rejected the hypothesis of harassment dilution, which states that females aggregate to decrease the *per capita* level of harassment. Instead, a trade-off appeared on female mating tactic between foraging competition and harassment avoidance; between a large group dominated by an adult male, and a small group dominated by a young harassing male. By demonstrating that a proximate process (males herding females) was responsible for the pattern at the population level (the increase of the average group size), I validated the self-organization theory on ungulate group dynamics. I also used these semi-experimental conditions to correct and validate a field-based measure of the intensity of the group dynamics.

Acknowledgment

A thesis is far from the work of a single person. Yet, it is a personal experience, I learnt a lot in this process and I lived many good experiences. I have many thanks to give, both academic and personal.

First of all, I would like to thank my PhD supervisor, Robert Weladji. Thank you for selecting me and trusting me, I am very glad to have been involved in this project in which you invested so much, way before selecting a PhD student. You supported me at every moment, you answered all my questions, and I was always welcome to your office, which I used and abused. Long walks in the field were privileged moments for exchanges.

My second thank goes to Øystein Holand. I started to think about my PhD theme while with you in a sauna, which is uncommon. Your classical answer "What do you think?" helped me to go further in my research questions. You taught me a lot about reindeer, and field work where no one can follow you, but Hallvard. From you and Knut Røed, to whom my thanks go as well, I've always received confidence in my project and useful advices. If I need your help in the future, for instance to explain reindeer courtship, I know I'll be well received. Thanks

I would also like to thank my committee members who were available and helpful when I needed, James Grant and Grant Brown. Thanks also to the two external examiners who accepted to read and to review my thesis.

In this thesis, I received help and comments from many people: Guillaume Larocque, Gilles N. Yoccoz, Mathieu Garel, Marco Festa-Bianchet, Eric Vander Wal and several anonymous reviewers who help improving the work undertaken in this document. Thanks Stefania Ucchuddu for finishing the work initiated with your MSc project. As English is not my first language I hope this document is understandable due to the corrections I obtained from friends. Special thanks to Queenie Gray who tried to understand what I was saying through texts, but also to Justin Strong, Andrew Harbicht and Michelle LeBlanc. Big thanks also to Sonia Saïd who advised me to keep contact with France on my rare stops there, and who was always able to find time to schedule my presentations.

This research would not have been possible without the participation of the Finnish team, at the Reindeer Research Station. “I have to say” thanks to Mauri Nieminen, of course, for your availability and wish to improve research at the center. Thanks Mika Tervonen, Jukka Siitari, Heikki Törmänen for your work on reindeer, GPS collars and database. Thanks also to Sari Siitari for your help at the research station.

I was not alone to go through this process. I thank my current and former lab mates: Alys, Natalka, Issac, Sacha, Justin, and the undergraduate students who volunteer in the lab. Thanks also to other grad students who started with me: Queenie, Andrew, Seba and Ashley, Shiva, Pierre, Michelle and many others. Thanks to those I met in Finland during field work: Hallvard, Natasha, Stefania, Eliana, and Martin. Merci aussi à mes amis de France qui ont fait ce même parcours et m’ont suivi de plus ou moins loin, G^2, Emmanuelle, Agnès.

Venir à Montréal a été une expérience humaine très enrichissante. Quelques personnes tiennent une place particulière pour moi: Ombeline, Emilie, Ursula et Dédy, j’ai trouvé en vous ma famille québécoise. Il y eu de nombreuses rencontres et j’ai une pensée pour la fine équipe de ces excellents barbecues d’hiver. I’ve also discovered the West of Canada and the Rockies. Thanks Barb, Ryan (who generously gave me some of his pictures) and the others in Waterton. Special thanks to Q-lu for bringing me there, and further.

Malgré la distance, et durant mes rares et trop courts passages en France, ma famille m’a toujours apporté son soutien. Il y a longtemps que je vous ai habitué à donner peu de nouvelles, et le travail intensif de la thèse n’a pas arrangé cette tendance, j’en suis désolé. Vos visites respectives m’ont fait énormément de bien, à chaque fois. Maman, merci pour nos conversations philosophico-politiques et nos balades dans les Calanques. Ces moments me suivent toujours. Merci Papa, pour ton soutien inconditionnel qui m’a permis de survivre à cette thèse, et pour ton intérêt pour mes projets. Merci Brunelle, Neeko, pour être aussi voyageur que moi ; je me suis senti un peu moins fou d’avoir traversé l’Atlantique pour étudier les rennes de Finlande.

Contribution of Authors

I was the principal investigator for all of the research work, but Chapter 5, undertaken in this thesis, under the supervision of Dr Weladji, and with the collaboration of Dr Holand, and Dr Nieminen. I supervised Dr Ucchuddu for Chapter 5, and the work undertaken in that chapter was equally shared. Chapter 2 resulted in a manuscript which has been published in an international peer-reviewed journal, and for which author’s credit is: Body G, Weladji RB, Holand Ø. Chapters 3 has been submitted to international peer-reviewed journals, as will be Chapters 4, 5 and 6. For publication, the authors will be credited in the following order for chapters 3, 4, 6: Body G, Weladji RB, Holand Ø, Nieminen M. For Chapter 5, credits will be given in the following order: Ucchuddu S, Body G, Weladji RB, Holand Ø, Nieminen M.

Table of Content

List of Figures

List of Tables

List of Equations

Female reindeer with Tellus GPS collar. Kaamanen, Finland, 2012 © Body G

Chapter 1 General Introduction

Charles Darwin presented his theory of sexual selection in 1871 (Darwin 1871), completing his theory of natural selection published 10 years earlier (Darwin 1859). He posited sexual selection was a main force driving species evolution and it explained strong morphological differences among closely related species, as well as extravagant individual characteristics (Danchin et al. 2008).

Evolution by sexual selection implies a difference in mating opportunities among individuals of one sex, generally the male. This inequality of mates can be driven by two processes: intra-sexual selection (generally referred to as "male-male competition"), and inter-sexual selection (generally referred to as "female mate choice). Intra-sexual selection often occurs among males and implies a competition for monopolization of mating opportunities. This process usually induces development of armaments (Danchin et al. 2008). Typical examples of armaments are found in ungulates, in which males have big horns or antlers (e.g. Fig 1.1a) and where fights among males are common during the breeding season (Clutton-Brock and McAuliffe 2009). Conversely, inter-sexual selection leads to the development of ornaments (Danchin et al. 2008), and examples of ornaments can easily be found in bird species in which males are highly coloured (e.g. Fig 1.1b, Mays et al. 2008).

Figure 1.1 Armament and ornament exemple in Canadian species. (a) Adult male rocky mountain bighorn sheep *Ovis canadensis*, and (b) colourfull male Western Tananger *Piranga ludoviciana* © Peruniak R

However, this traditional perspective of the independent occurrence of intra-sexual and inter-sexual selection is crumbling (Clutton-Brock 2007). Meanwhile, overlooked mating strategies, such as female mate choice in mammals (Clutton-Brock and McAuliffe 2009) or female-female competition (Stockley and Bro-Jørgensen 2011) are gaining momentum. Sexual conflicts such as coercion (Muller et al. 2011; Bro-Jørgensen 2011), reverse sexual conflict (Bro-Jørgensen 2007), harassment avoidance (Clutton-Brock and McAuliffe 2009) or complex interactions among sex specific mating tactics (Bro-Jørgensen 2011) are increasingly being better described and understood. Revision of ungulate mating tactics, which traditionally focused on male-male competition (Carranza 2000; Isvaran 2005), must therefore include these various aspects. The purpose of this thesis is to study female mating behaviour in an ungulate species, the reindeer *Rangifer tarandus*.

1.1 Mating systems and sexual selection

The main question of the past decades has been to determine the extent of sexual selection. One of the useful measures being the “opportunity for sexual selection” (Shuster and Wade 2003) which measures the variance of fitness among individuals within a sex. High fitness variance induces a high opportunity for selection and one may expect important secondary sexual characters (i.e. ornaments or armaments) or sexual dimorphism (Vanpé et al. 2007b) to occur. This measure mainly explains the reason for the development of secondary sexual characters in males: the variance among male success is generally much higher than the variance among female success. Indeed, it is useful to compare species, populations and mating systems as well.

Mating systems, the outcomes of the reproductive strategies of individuals (Clutton-Brock 1989), have been classified in regard to male and female mating behaviour (Emlen and

Oring 1977; Clutton-Brock 1989). Mating systems range from obligate monogamy to high polygamy over a range of increasing opportunity for sexual selection, as the polygamy of the system increases. The classification of mating systems over taxa allowed researchers to relate environmental and species characteristics to a particular kind of mating system, and therefore to a level of opportunity for sexual selection. Three key parameters have been identified to explain the degree of polygamy: 1) the necessity of bi-parental care, which constraints systems to monogamy (Clutton-Brock 1989), 2) the spatial distribution of females in estrous and 3) the extent of the synchrony of estrous (Emlen and Oring 1977).

The opportunity for sexual selection is high for females exhibiting moderate spatial and temporal clustering of estrous, as few males are able to monopolize the majority of estrous females in these conditions. Spatial distribution of resources is an important parameter influencing the spatial distribution of females, and males adapt their strategy accordingly (Carranza et al. 1995; Carranza and Valencia 1999). Indeed, males may display an array of tactics including: a resource defense tactic when females aggregate at particular places, a harem defense tactic when groups of females are moving, a tending tactic either when females are spread out or when they form groups too large to be defended (Emlen and Oring 1977; Clutton-Brock 1989; Carranza 2000; Isvaran 2005), a defence of small territory which do not contain resources, i.e. leks, when females cluster at these places for various possible reasons (Clutton-Brock et al. 1996), or alternative tactics (Brockmann 2001) when they are not able to dominate the male-male competition. Conversely, female mating tactics may themselves influence their aggregation level, and therefore they may directly influence male mating tactics and the opportunity for sexual selection.

1.2 The female mating tactics

Even though the ultimate goal is the same, i.e. maximizing individual's fitness, female mating tactics cannot be classified in the same way as those of males. In contrast to males, multiple mating opportunities do not increase female mating success, i.e. copulating with many males does not increase the number of offspring produced (the Bateman gradient; Wade and Shuster 2010; but see Kvarnemo and Simmons 2013; Aloise King et al. 2013; Briefer et al. 2013 for advantages provided by multiple mating). To classify female mating tactics, we have to distinguish the ultimate reasons for the selection, by determining the parameter increasing the female fitness, and identify the process of selection by itself, i.e. the actual female behaviour.

1.2.1 The ultimate reasons of the selection

Two non-mutually exclusive reasons for the selection are apparent (Fig 1.2): first, females can adopt a tactic that provides direct benefits, which increases female reproductive success, or females can adopt a tactic that provides indirect benefits, which increases offspring fitness.

By obtaining direct benefits, females will improve their current reproductive success by increasing their productivity (number of offspring produced) and their parental ability (ability to successfully raise their young). Direct benefits can be obtained by selecting males for their parental care ability (Jennions and Petrie 1997; Alonzo 2012) or their reproductive experience (Dubois and Cézilly 2002; Griggio and Hoi 2011), by accepting nutritive gifts males trade for copulation (Velando 2004; Tryjanowski and Hromada 2005; Albo and Costa 2010), by selecting the quality of the territory rather than the male which defends it (Alonso et al. 2012; Hasegawa et al. 2012), or by reducing the cost suffered during the breeding season, for instance by avoiding the harassment from young males (Galimberti et al. 2000; Cappozzo et al. 2008).

Lactation in mammals decreases the need for females to obtain parental care from males (Aloise King et al. 2013), and examples of male parental care are rare in ungulates (Bro-Jørgensen 2011). Conversely, pregnancy and lactation are energetically costly, and female ungulate body condition is consequently a main determinant of their reproductive success (Festa-Bianchet 1998; Ropstad 2000; Hamel et al. 2009). Female body condition also increases the offspring body condition which may impact its quality throughout its life (Forchhammer et al. 2001; Hamel et al. 2009; Fig 1.2), thereby increasing the female fitness in the offspring generation.

The breeding season can influence female body condition (Byers et al. 2005; Holand et al. 2006) or even their survivorship due to male mating behaviours (Réale et al. 1996), enhancing the need for females to care about the direct benefits of their mating tactics.

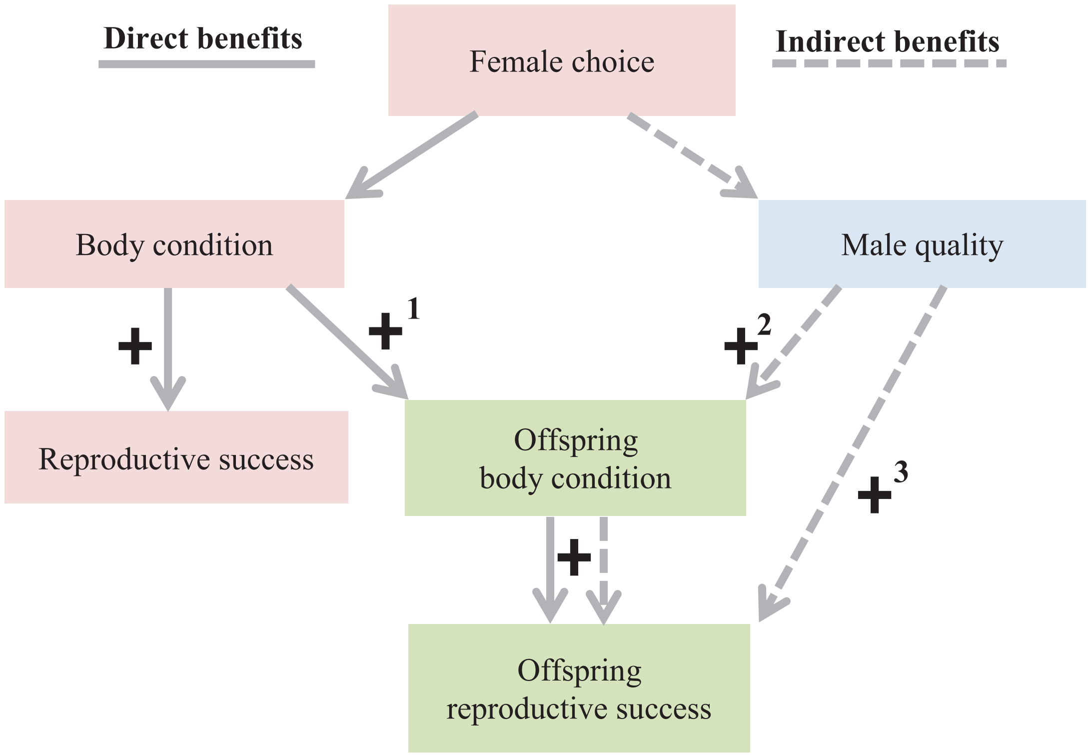


Figure 1.2 Ultimate reasons of the selection on their fitness consequences. [1] Early body condition shapes individual quality for its whole life; [2] Good genes hypothesis; [3] Sexy son hypothesis. Positive correlation are indicated for each relation

Selection for resource-based territories (fallow deer *Dama dama*: Clutton-Brock 1989; pronghorn *Antilopacra americana*: Byers et al. 1994; red deer *Cervus elaphus*: Carranza 1995) and harassment avoidance (fallow deer: Clutton-Brock et al. 1992; red deer: Carranza and Valencia 1999; moose *Alces alces*: Bowyer et al. 2011) seem, unsurprisingly, to be the main mating tactics based on direct benefits described in ungulate species.

The ultimate selection for indirect benefits is also known as the "good genes" hypothesis, whereby the offspring inherit the qualities of their father (Mays et al. 2008). These qualities correspond to any life history trait that is selected by natural selection (Alonso et al. 2010; Fig 1.2), including parasitism resistance, and secondary sexual traits that are selected by sexual selection (the "sexy son" hypothesis, Fig 1.2, Huk and Winkel 2007). Male quality can also influence the offspring sex ratio (Røed et al. 2007), but the extent of the inheritance of father quality can depend on offspring sex (Kokko and Jennions 2008; Mainguy et al. 2009).

Females may assess male heritable quality from secondary sexual traits. For instance, parasitism resistance, a highly genetically determined trait, is honestly represented by ornament size and symmetry in buffalo *Syncerus caffer* (Ezenwa and Jolles 2008). In general, a good immune system, or higher survival ability, may allow males to invest extra-energy into secondary sexual traits, which is known as the handicap principle (Danchin et al. 2008). Parasitism resistance is also known to be negatively influenced by a loss of genetic variability (Hawley et al. 2005; Drury 2010). Consequently, the widely reported inbreeding avoidance in ungulates (reviewed in Clutton-Brock and McAuliffe 2009, but see Holand et al. 2007) may correspond to a mate choice for indirect benefit.

1.2.2 The proximate processes of the selection

Independent of the ultimate reason for sexual selection, we can classify the proximate processes of the selection (the processes by which variation in mating opportunities occur) in three categories (Table 1.1): Coincidental mate choice corresponds to any behaviour that induces a bias in male mating opportunities, but without any influences of males on the system. Female movements are comparable with or without males. This process emphasizes that differences in mating opportunities have not evolved through a process of female choice, but through a process of male-male competition to get access to groups or locations where females are (Clutton-Brock and McAuliffe 2009). Habitat preference, resource based territories and female aggregation in response to predators, induce coincidental mate choices.

Female fallow deer and female red deer, for instance, display coincidental mate choice: they select territory quality rather than male characteristics (Clutton-Brock 1989; Carranza 1995) and show a preference to join female aggregations whether or not a male is present (Clutton-Brock and McComb 1993).

Assortative mating between male and female quality is a particular form of coincidental mate choice that reduces the opportunity for sexual selection (Farrell et al. 2011). This correlation happens because adult dominant males become exhausted at the end of the rut, allowing young males to obtain more mating opportunities (Hirotani 1994) when

Table 1.1 Classification of the proximal processes of the selection

Coincidental mate choice[1]	**Female mate choice**[1]	
	Indirect mate choice[2]	**Direct mate choice**[2]
Natural behaviours of females induce difference of mate opportunities	Females actively increase competition among males	Females evaluate and select males based on criteria
Fallow deer[3] Red deer[4]	Pronghorn[5] Moose[6]	Red deer for roar[7]

[1]Clutton-Brock and McAuliffe 2009; [2]Wiley and Poston 1996; [3]McComb and Clutton-Brock 1994; [4]Byers et al. 1994; [5]Bowyer et al. 2011; [6] Charlton et al. 2007

youngest, lightest and less successful females enter into estrous.

Conversely, female mate choice refers to the selection process for which females change their behaviour according to the presence of males (Clutton-Brock and McAuliffe 2009). Two kinds of female mate choice can be defined in relation to whether or not females assess criteria (Table 1.1, Wiley and Poston 1996). Indirect mate choice refers to processes where females voluntarily increase competition among males. Female behaviour is not optimal in regards to resource distribution as in coincidental mate choice, but their behaviour is not dependent on characteristics of the male.

Female movements between herds are an example of indirect mate choice because female departure attempts enhance male herding ability. A male of high quality will successfully herd females more often than a male of lower quality, resulting in a correlation between male quality and harem size, which increases the opportunity for sexual selection (Wade and Shuster 2004). However, for mate choice to be indirect, the female departure rate should be independent of the male quality. A perfect example of indirect mate choice is found in pronghorns (Byers et al. 1994) where some females actively induce fights between males and then mate immediately with the winner, whoever that is. Another example is found in moose, when females call the dominant male when approached by a subordinate male, which increases the competition among males (Bowyer et al. 2011).

In contrast, direct mate choice requires female cognitive capacities to recognize and to compare criteria about males and consequently to adapt their behaviour to this assessment (Wiley and Poston 1996). Criteria can be assessed by females (reviewed in Clutton-Brock and McAuliffe 2009) based on physical appearance such as body condition, ornament size and symmetry (Markusson and Folstad 1997), vocal performance (Charlton et al. 2007), or behavioural characteristics such as social rank, fighting ability (Byers et al. 2010; Pérez-Barbería and Yearsley 2010), or ability to provide protection against harassment.

This evaluation of male quality requires that females perform a mate sampling strategy and compare available males. Three main models of sampling strategy exist (Luttbeg 2002; Wiegmann et al. 2010; Wiegmann et al. 2013): the fixed sample search strategy, in which females evaluate a fixed number of males, then go back to the best one (best-of-n strategy, e.g. Rintamäki et al. 1995); the sequential search strategy for which females select the first male that reaches their expectation (threshold strategy, e.g. Beckers and Wagner 2011); and the Bayes comparison strategy for which females use the knowledge of male quality they have acquired from the males they have already met to assess the potential gain and cost of sampling another male (see Luttbeg 1996).

Demonstrating direct mate choice in nature is difficult. Indeed, females may not always be able to perform one of the above sampling strategies, these strategies may be too costly compared to the expected benefits, or females may mate even if available males do not reach their expectation. These difficulties highlight the difference between female "preference function", i.e. the order with which a female ranks prospective mates, and female choosiness, i.e. the effort a female is prepared to invest in mate assessment (Jennions and Petrie 1997).

The range of traits selected by female ungulates is expected to be as large as those selected in birds (Byers et al. 2010), but the evidences are lacking (Clutton-Brock and McAuliffe 2009). Female preferences can be experimentally assessed by modifying a male trait, but this is difficult to do in ungulates. Two studies were, nevertheless, able to modify the antler shape or the male vocalization. Female fallow deer did not express a preference between antlered males and antlerless males (McComb and Clutton-Brock 1994) whereas female red deer preferred vocalizations that came from large males over those that came from small males (Charlton et al. 2007).

The costly movements among harems of female pronghorns have also been interpreted as a sign of sampling strategy, and therefore direct mate choice (Byers et al. 1994; Byers et al. 2005). Other studies reported higher movement rate of females during their estrous, and some females were even leaving their usual home range for a short time period (roe deer *Capreolus capreolus*: Lovari et al. 2008; Richard et al. 2008; red deer: Stopher et al. 2011). However, whether these movements correspond to a sampling strategy and a direct mate choice remains unclear as red deer females movements were independent on male quality (Stopher et al. 2011).

1.2.3 The case of harassment avoidance

Harassment avoidance is expected to be a main mating tactic of female ungulates (Clutton-Brock and McAuliffe 2009; Bro-Jørgensen 2011), and is even used to explain the evolution of leks (Clutton-Brock et al. 1992, 1996; Carranza 2000 but see Carbone and Taborsky 1996). Females may avoid harassment by aggregating, i.e. the dilution effect (Clutton-Brock et al. 1992; Carranza and Valencia 1999), or by staying under the protection of an adult dominant male (Holand et al. 2006; Bowyer et al. 2011). Yet, if this mating tactic clearly provides direct benefits to females, it can involve any of the proximal processes described above.

Indeed, female aggregation may be independent of harassment level, but larger aggregation may still dilute the harassment received *per capita*. If females do not increase their aggregation level, these aggregations induce coincidental mate choice. However, if females over aggregate, they induce an indirect mate choice, as this response to male harassment increases male-male competition and the inequality of mating opportunities among males. Conversely, if females switch among groups and stay longer in less harassed groups, they express a direct mate choice. Indeed, they evaluated a criterion (the ability to provide protection against harassment), and they modified their behaviour (whether they leave) accordingly.

1.3 Living in groups

So far, we have seen that the distribution of resources can influence mating systems through its influence on female spatial distribution (1.1), that females may move among harems to sample males (1.2.2), and that females may benefit from living in groups to decrease harassment levels (1.2.3). Studying female mating tactics will therefore benefit from a better understanding of group living.

Living in groups has traditionally been studied from the cost/benefit approach (Krause and Ruxton 2002; Danchin et al. 2008). This approach searches for an optimal group size which optimizes the balance between costs and benefits of living in groups. Increasing group size is expected to benefit individuals by decreasing their predation risk (Hoare et al. 2004; White et al. 2012b; Marshall et al. 2012), increasing foraging efficiency (Pereira et al. 2013), or by decreasing sexual harassment levels (Carranza and Valencia 1999; Galimberti et al. 2000; Cappozzo et al. 2008). However, increasing group size is also costly for individuals as it increases foraging competition (Focardi and Pecchioli 2005; Marshall et al. 2012) and agonistic social interactions (Marshall et al. 2012). A broad conclusion of this approach is that observed group sizes are typically larger than the optimal group size, as it is often beneficial for a solitary individual to join a group at the optimal group size, but none of the individuals of this larger group would benefit from leaving it (Danchin et al. 2008).

Yet, summarizing the individual balance between costs and benefits is difficult due to the variability of individual requirements, sensitivity to predation risk, environmental variability (Danchin et al. 2008) or even personality (Bergvall et al. 2011). For instance, it has been shown that variation of the cost/benefit balance exists within a group. Individuals in central position experience higher foraging competition (Focardi and Pecchioli 2005) or increase negative social interactions (Hirsch 2011), but individuals in peripheral position are more sensitive to predation risk (Lung and Childress 2007; Morrell et al. 2010) or suffer higher sexual harassment levels (Carranza and Valencia 1999). In addition, the cost/benefit approach does not explain the large variability of group size, nor the frequent merging and splitting events groups may experience (but see Asensio et al. 2009).

The second, and more recent approach to study group living corresponds to the self-organization theory (Couzin and Krause 2003). Groups are seen, under this approach, as an inherent consequence of the random distribution of individuals in a landscape (Couzin and Krause 2003; Danchin et al. 2008). Randomly, the distribution of group sizes in a population corresponds to a Poisson distribution (Caughley 1964), i.e. many small groups and few large groups (more precisely described by the power-law function, Couzin and Krause 2003; Beauchamp 2011a), and the average group size is positively correlated to the population density (Beauchamp 2011b). Common interests of individuals for particular habitats increase the local population density and consequently increase the average group size.

Groups, by randomly bunching to each other, merge, and then split when group members are moving in different directions, which leads to high rates of fission and fusion events, as well as unstable group composition. Group cohesion, bird flock movement, fish school fleeing movement or the wave movement of the front of wildebeest herds during migration can be understood by simple rules, such as copying the neighbour behaviour

(Couzin and Krause 2003). In this context, the variations of group composition between days are difficult to interpret. Indeed, female movement among harems may either be due to active movement, such as those representing a mate sampling strategy, or be the result of the fission-fusion group dynamics.

In addition, the self-organization theory predicts that changes of the distribution of group size should be the result of individual behaviours which reflect the level of motivation to join or to leave groups (Couzin and Krause 2003; Beauchamp 2011a). The variations in opposite directions of these two components of the group dynamics, i.e. leaving and joining a group, can have the same effect on the resulting group size. For instance, an increase of individual motivation to join a group, as well as a decrease of their motivation to leave a group induces an increase in group size. In contrast to the cost/benefit approach, the self-organization approach can differentiate these mechanisms of the group dynamics. Identifying changes of one of these mechanisms is a first step to identify the males or females mating behaviours involve in group size variation.

1.4 Mating system and mating tactics in reindeer

Reindeer are the most northern ungulate species and one of the two permanently Artic species, the other being muskoxen *Ovibos moschatus*. Different subspecies are found in North America and Eurasia, in tundra and forest habitats (Røed 2005). The reindeer is a particular species: it is the only cervidae that has been domesticated (Røed et al. 2008), the only one in which both sexes carry antlers (Geist 1999; Melnycky et al. 2013), and is the species that exhibits one of the largest sexual dimorphisms among ungulates (Geist and Bayer 1988), with adult males attaining a mass twice that of females.

Reindeer mating systems vary widely across continents, especially in regards to the variation of the group dynamics of females. When males and females aggregate together in large migratory herds of hundreds or thousands (e.g. caribou *R.t. caribou* in Newfoundland, Lent 1965), males are not able to form harems, in which they would have exclusive access to females. Instead, they utilize tending tactics: they follow only one female in estrous at a time (Lent 1965; Henshaw 1970). When individuals are not migrating, and resources are clumped, female groups are spatially stable. These groups do not undergo fission-fusion dynamics, and males may either defend harems at a stable location, or they may defend the clumped resources (e.g. Svalbard reindeer *R.t. platyrhynchus* in the Spitsbergen Archipelago, Heatta 2009). Conversely, when resources are more evenly distributed, groups are not spatially stable and mating systems are harem-based (Espmark 1964; Kojola 1986; Skogland 1989; Hirotani 1994). However, these harems are highly volatile due to fission-fusion group dynamics (Hirotani 1989; Holand et al. 2006; L'Italien et al. 2012) and the instability of male hierarchies (Barboza et al. 2004; Holand et al. 2012). The semi-domestic herd of reindeer that I am studying displays fission-fusion group dynamics, and has a mating system based on temporary harems.

According to the strong sexual dimorphism in reindeer, male reproductive success is highly skewed in reindeer mating systems (Skogland 1989; Røed et al. 2002) but could be even more skewed as young males are still able to obtain some mates (Røed et al. 2005). Male body mass, age and dominance status are correlated with their reproductive success (Røed et al. 2002; Røed et al. 2005). Their mating tactics in temporary harems have been suggested to be particularly flexible (Clutton-Brock 1989). One reason for this flexibility may come from the fine adjustment of their mating behaviour to local conditions such as group size, number of competitors (Tennenhouse et al. 2011), or changes in hierarchies due to stochastic events (e.g. injuries; Holand et al. 2012). Another reason may be the variability of the relative success of male tactics according to their age (Tennenhouse et al. 2012), body condition and social rank (L'Italien et al. 2012). Finally, the flexibility of male mating behaviours could be time-dependent (Mysterud et al. 2004; Tennenhouse et al. 2012), as body condition (due to hypophagia) and the number of females in estrous vary throughout the rut period.

Females have a short estrous of 24-48h (Espmark 1964; Hirotani 1989; Ropstad 2000) which are also highly synchronous. Indeed, almost all females copulate during the peak rut period (Kojola and Nieminen 1988) which lasts about 10 days (Kojola 1986; Skogland 1989). Females that did not copulate during this period show an alternative estrous cycle (mean length 13-33 days, Ropstad 2000), which leads to a second peak rut 11-20 days after the first one (Hirotani 1989).

Inter-individual variations in fitness are mostly due to variations in female quality (Weladji et al. 2008), which correlates with female body mass and their social rank. Indeed, the body mass of female reindeer in the autumn is a strong determinant of their pregnancy rate (Reimers 1983; Ropstad 2000) and of the weight of fawns the next autumn (Holand et al. 2003). Female social rank is related to calving date (Holand et al. 2003; Holand et al. 2004), either through female-female mate competition or winter foraging competition. Females may compete for mates, thus securing early copulation may be important for them. Late copulations induce late calving which increases summer calf mortality and decreases autumn calf weight (Holand et al. 2003), therefore decreasing female reproductive success.

Female reindeer may choose their mates, as they regularly court males (Espmark 1964; Djaković 2012). Females also produce more male offspring when they copulate with a high quality male (Røed et al. 2007). Yet, the criteria females employ to choose their mate remain unknown, however it is apparent that they do not use relatedness (Holand et al. 2007) or compatibility of immune systems (Djaković 2012).

Fission-fusion group dynamics may influence female mating behaviour. Group composition is highly variable (10-20% of the group composition changes daily, Hirotani 1989) and it has been described as a result of random inter-harem movement (Hirotani 1989; Røed et al. 2002). Within the two weeks preceding their estrous, females can repeatedly visit one to seven dominant males (L'Italien 2010). Whether this encounter rate is a sampling behaviour or is due to the fission-fusion group dynamics remains unknown. In addition, females preferentially aggregate with close kin (mother-daughter) during the breeding season, especially in the presence of young males (Djaković et al. 2012). This behaviour may be a response to costly harassment by young males (Djaković 2012), as females lose an average of 2.5% of their body mass during the breeding season (Holand et al. 2006).

In conclusion, female reindeer show behaviours that could be indicative of mating tactics for direct or indirect benefits; through direct, indirect or coincidental mate choice. Previous studies on female mating behaviour have, however, been weakened by the unknown implications of fission-fusion group dynamics, and therefore these studies were limited to broad patterns during the whole breeding season. The explicit integration of group dynamics in this research will therefore improve our understanding of female mating behaviour in reindeer.

1.5 Objectives

The purpose of my research is to study female mating behaviours in reindeer. To achieve this goal, I must study and understand the fission-fusion group dynamics females are experiencing during the breeding season. Most of this research is based on remote sensing of the spatial movements and activities of individuals from an experimental herd of reindeer. Each individual of the herd was equipped with global positioning system (GPS) collars and activity sensors. Data resulting from similar experimental set up are rare, justifying an important part of this research work being on the development of adequate methods.

Objective 1: Developing remote sensing methods. Activity sensors are commonly included in GPS collars, but are not widely used in research due to the difficulty in linking the records to the actual behaviour of individuals. Chapter 2 develops and validates the recursive model, which is a new method of assessing the proportion of time individuals are active. This method is further extended to specific behaviours and applied in Chapter 4. GPS data are increasingly being used in ecology (Cagnacci et al. 2010), but examples of herds entirely equipped with GPS collars are rare. Thus, in Chapter 3, I detailed the method required to follow the fission-fusion group dynamics from GPS data. This methodological development allowed us to correct and validate a method of assessing the intensity of the group dynamics from direct observations of group size (Chapter 6).

Objective 2: Quantifying the fission-fusion group dynamics. I tested whether the self-organization theory applies to reindeer groups. This theory predicts that behaviours at the individual level explain changes in population patterns. Chapter 3 and Chapter 4 form a funneled approach. I first determined which components of the fission-fusion group dynamics (i.e. the rate of fusion or fission) induced changes in the average group size. Then, in Chapter 4, I linked these changes to specific behaviours. The functioning of the fission-fusion group dynamics is also described in detail in Chapter 3.

Objective 3: Identifying female mating behaviours. Building on conclusions drawn in Chapter 3, Chapter 4 tested three mating behaviours that may be responsible for the increase in group cohesiveness: male herding ability, female mate choice and female harassment avoidance. The study of female harassment avoidance is extended in Chapter 5, where I assessed the relative importance of the harassment dilution effect and the female foraging competition. There, I also contrasted the harassing contribution of adult and young dominant males and evaluated the harassing influence of satellite males.

Chapter 2 The recursive model as a new approach to validate and monitor activity sensors

The following chapter is based on the published manuscript: Body G, Weladji RB, Holand Ø (2012) The recursive model as a new approach to validate and monitor activity sensors. *Behavioral Ecology and Sociobiology* 66:1531-1541

2.1 Abstract

Activity sensors are increasingly being used to monitor animal activity but current methods, used to validate the relationship between the motion sensor information and the actual behaviour of animals, have weaknesses. This study aims to improve the methods used to estimate activity level from dual axis activity sensors and to validate the Tellus activity sensor for reindeer *Rangifer tarandus*. We developed a new approach, the recursive model (a recursive application of a logistic regression), to predict continuous values of activity without biased estimations or previous modifications of the dataset. We compared this new recursive model approach with two traditional approaches: the tree classification method and the standard model (based on simple logistic regression). Estimations from the tree classification and the standard model were dependent on the dataset used for validation, whereas the recursive model gave unbiased estimations. Estimations from standard and recursive models were also more accurate (lower average absolute errors) than those from the tree classification method and they had a slightly better discriminatory power (higher percentage of good classification). We successfully applied the recursive model for the first time and validated the Tellus activity sensor for reindeer. Any user can apply our methodology to obtain their own equations of the relationship between activity sensor values and the level of activity of the individual, and users monitoring reindeer activity with Tellus activity sensor can directly apply the provided equations under appropriate conditions.

2.2 Introduction

In recent history, position data from the global positioning system (GPS) have increasingly been used not only to remotely track the movements of animals but also to gather data to establish home ranges, migration routes, or habitat selection (Millspaugh and Marzluff 2001). Given the importance of studying animal behaviour and activity patterns in ecology and with the development of GPS technology, animal ecologists and managers have recently been using GPS coupled with activity sensors to record activity data automatically and continuously with high frequency and in unobservable time such as during night (Cagnacci et al. 2010). Large animals have received greater attention and a great number of activity sensor validations exist among ungulates (e.g., red deer *Cervus elaphus*: Löttker et al. 2009 and roe deer *Capreolus capreolus*: Gottardi et al. 2010, see Appendix 2.1). Generally, two types of validations have been performed: one distinguishes active from inactive behaviours (e.g. Moen et al. 1996; Gottardi et al. 2010) while the other differentiates the three major behaviours: eating, walking, and resting (Naylor and Kie 2004; Ungar et al. 2005; Moreau et al. 2009). There are concerns about the way these validations are established as there is no standardization of the method used to validate the relationship between the values recorded by the activity sensors and the real activities (Gottardi et al. 2010). Generally, authors find a threshold on the activity sensor value to classify activity on a binary scale, i.e., active versus inactive (see Adrados et al. 2003 for a method with three classes: active, inactive, and intermediate). With this binary approach, the authors must observe activities either in active or in inactive bits, but in reality, a mixture of active and inactive behaviours may be observed during the lapse of time in which the activity sensor is recording movement. Data in this format have been obtained using three simplification methods. Firstly, authors have used only "pure" data, i.e., with only one behaviour expressed during the recorded lapse of time, thereby disregarding mixed data (Naylor and Kie 2004; Coulombe et al. 2006; Bourgoin et al.

2008; Kozakai et al. 2008). Secondly, the authors have used a cut-off value of activity percentage below which the lapse of time is considered inactive and above which it is considered active (Moen et al. 1996; Gervasi et al. 2006; Gottardi et al. 2010). Thirdly, authors have used only the dominant behaviour of the lapse of time (Relyea et al. 1994; van Oort et al. 2004) or intentionally changed the definition of active and inactive behaviours (for example “standing” was considered an active behaviour by Coulombe et al. 2006).

The binary “active” versus “inactive” framework leads to some conceptual issues: first, by fitting a binary response, authors assumed that there is no inactive behaviour expressed during an active lapse of time. However, this low frequency of inactive behaviour could be of interest (e.g., studying the extent of the vigilance during eating bouts). Secondly, the choice of a cut-off for percentage of activity that differentiates active from inactive behaviours is subjective and will influence the results (see Gottardi et al. 2010). To overcome this problem, some authors have attempted to fit continuous equations of the proportion of activity recorded. While Ungar et al. (2005) were successfully able to set up equations for the time spent grazing, resting, or traveling in recorded times, Moen et al. (1996) were not able to find such a relationship.

When the activity sensor values of each behaviour are overlapping (see Relyea et al. 1994; Coulombe et al. 2006), we can predict that the threshold value used to distinguish active from inactive behaviours will vary with the relative proportion of these behaviours in the analyzed dataset. This variation is expected to be higher for intermediate values of the collar movement for which both active and inactive data can be either common or rare depending on the absolute frequencies of the behaviour. Accordingly, we expect the extent of the estimation bias, i.e., the difference between the mean of the predicted values and the mean of the observed values, to be correlated with the difference of the percentage of activity between the dataset used to establish the model (the training dataset) and the dataset used to apply it (the testing dataset, Cramer 1999; Oommen et al. 2011). This estimation bias is well-known in case–control studies which commonly estimate rare events based on an unbalanced dataset (Hosmer and Lemeshow 2000; He and Garcia 2009) and is suggested as the cause of "bias" of estimations by Pearce and Ferrier (2000). Even though statistical methods exist to correct such bias (King and Zeng 2001), none of the previous studies assessed the influence of their dataset on their validation procedure. In this paper, we present a method that gives a value to the level of activity, instead of a binary response, while avoiding the simplification of the training dataset, and we use the proposed method to validate the use of the Tellus activity sensor on reindeer. We first compared three different methods for estimating activity level from dual axis activity sensors, the common tree classification method, the standard logistic model, and a new method, a recursive application of a logistic regression (hereafter "recursive model"). Secondly, we explored the extent of the estimation bias due to the unbalanced nature of the datasets often used. Thirdly, we validated the estimation of activity level for the Tellus activity sensor on reindeer and assessed its accuracy as well as whether the results can be generalized to any individual.

2.3 Methods

2.3.1 Study area and GPS collars

The study was conducted at Kutuharju Field Reindeer Research Station in Kaamanen, Finland (69° N, 27° E). During the breeding season in 2010, 18 individuals of variable ages (13 males: one 6.5-year-old; one 5.5-year-old; four 4.5- year-olds; three 3.5-year-olds; three 2.5-year-olds; and one 1.5-year-old; five females: 10.5, 8.5, 6.5, 4.5, and 1.5 years old) were fitted with Tellus GPS collars from Followit (URL, http://wildlife.followit.se/). For a more intensive monitoring of the activities, two individuals (one 4.5-year-old male and one 6.5-year-old female, group A) were isolated in a 2-ha enclosure between the 11th and 15th of

October 2010, while the others (group B) were let free in a large enclosure (15 km^2) during the breeding season (from late September to late October).

2.3.2 Activity data

We observed behaviours from dawn to dusk during three consecutive days for group A and opportunistically during most of the breeding season for group B. We recorded the first 60 s of activity during each GPS recording. Thereafter, we only included in the dataset the behavioural observation in accordance with the time-to-fix (TTF, the time necessary for the GPS collar to fix its position) of each record and we deleted the data when the TTF was greater than 60 s. Behavioural observations were recorded per second (we noted the second at which there was a change in activity) and included the following: resting (lying down), standing (body and head are in upright position without locomotion), walking (locomotion more than two steps), eating (body upright, head down with less than two consecutive steps), fast locomotion (running), and other high level activities (e.g., bush trashing, sparing, and scratching). Time spent inactive was considered as the sum of the time spent resting and standing, while all other behaviours were classified as active. We standardized the time spent active by dividing these values by their respective TTF; thus, we obtained the proportion of time spent active.

2.3.3 Tellus activity sensors

Tellus GPS collars were equipped with dual axis activity sensors. The activity sensor records back–forth (x-axis) and left–right (y-axis) movements with an accelerometer (Tellus user manual revised on May 16, 2008). Each second during the TTF of the GPS localization attempt, the collar records whether the accelerometer goes over a threshold (Cecilia Thynell, Followit support manager, personal communication) which is adjustable by the sensitivity parameter (sensitivity = 5 in our study). Each activity sensor record corresponds to the number of seconds for which the accelerometer exceeded the threshold. As the TTF varies between attempts (minimum 30 s and programmed to stop the attempt at 90 s), we controlled for this variability by dividing the value recorded for each axis by the TTF of that particular attempt (hereafter called adjusted values: "X_{adj}" and "Y_{adj}"). GPS collars recorded their positions every 5 min for group A and every full hour for group B. The observers' watches were synchronized with the Greenwich Mean Time using handheld GPS time.

2.3.4 Activity validation method

2.3.4.1 Tree classification

Here, we classified our observations as inactive if the observed proportion of activity was ≤ 0.5 and as active if the proportion of activity was > 0.5. The tree classifications were built with the function "rpart" (package "mvpart") in R software (R Development Core Team 2011) and is hereafter called "Tree50." The difference in predictions induced by the use of different data simplification methods is investigated in the Appendix 2.2 by fitting a tree classification based on a cut-off at 0.3 ("Tree30") and by building the tree classification only on pure data ("Tree pure").

2.3.4.2 Standard model

Using the standard model method, we modeled the proportion of activity as a function of the value of each adjusted activity sensor's axis (X_{adj} and Y_{adj}) using a generalized linear model (GLM) with a quasibinomial distribution of errors ("quasi" to take into account overdispersion) and a logit link; this model is further referred as the "logistic regression" (Crawley 2007). We used the iteratively reweighted least squares algorithm to estimate parameters and weighted the data using the observations' respective TTF. We included a square term for each axis as well as the interaction between the two axes (Eq. 2.1) to increase the model flexibility by allowing the curve to decrease (Agresti 1996):

Eq. 2.1 $$Activity \sim f\left(X_{adj} + Y_{adj} + X_{adj}{:}Y_{adj} + {X_{adj}}^2 + {X_{adj}}^2\right)$$

where X_{adj} and Y_{adj} are the back–forth and left–right movement values of the collar adjusted by the TTF, $X_{adj}{:}Y_{adj}$ is the interaction term, Activity is the proportion of time spent active during the lapse of time, and *f(x)* represents the logistic function.

2.3.4.3 Recursive model

In order to avoid the bias of estimation due to the training set characteristics, a recursive model was developed. A recursive model requires two steps: we first (step 1) corrected our estimations from an unbalanced dataset as required if we knew both training and testing mean percentage of activity (MPA). Secondly (step 2), we integrated in our procedure the absence of an *a priori* knowledge of the MPA of the testing dataset ($\mathrm{MPA}_{\mathrm{testing}}$).

Step 1: "Prior correction" and "weighting" are two methods used by King and Zeng (2001) to correct the intercept of a fitted logistic equation for unbalanced datasets. However, these methods only work for binary type data, but not for proportions (see Appendix 2.3). Indeed, such a correction method has not yet been developed for continuous data (Maalouf 2011, Gary King, personal communication). We propose here a method to empirically assess the relationship between the intercept of equation 2.1 with the value of $\mathrm{MPA}_{\mathrm{training}}$ based on subsampling. We generated 250 random subsamples of 75 data points (a data point is composed of the observed proportion of activity, X_{adj}, and Y_{adj}) from our training dataset and we assigned to each data point the MPA value of its subsample. Then, we used all of these 18,750 new data points (each one composed of Activity, X_{adj}, Y_{adj}, and MPA values) to fit a logistic regression (see the "Standard model" for details) based on equation 2.2 below:

Eq. 2.2 $$Activity \sim f\left(X_{adj} + Y_{adj} + MPA + X_{adj}{:}Y_{adj} + {X_{adj}}^2 + {X_{adj}}^2\right)$$

where MPA is the mean percentage of activity of the subsample; see equation 2.1 for the other variables.

The number of subsamples (250) and their sizes (75 data points ~25 % of the dataset) were arbitrarily chosen, but they have been set up so as to increase the range of the MPA obtained from the various subsamples, thereby increasing the range of model reliability. We fitted equation 2.2 on the merged 250 subsamples rather than (1) fitting equation 2.1 on each sub- sample then (2) evaluating the relationship between the MPA of the subsamples and the value of the intercept of the different equations. The procedure using equation 2.2 is easier to run, more precise, and the strength of the relationship is directly understandable by the extent of the coefficient of the MPA variable in equation 2.2.

Step 2: We based our second step on the recursive application of equation 2.2; a method that has already been proposed for improving classification methods for unbalanced datasets (Hand and Vinciotti 2003; Maalouf and Trafalis 2011).We applied equation 2.2 with an *a priori* MPA = 0.5 (the *a priori* value does not affect the outcome of the recursive model) to obtain a first biased estimation of both the activity level of each data point and of the MPA of our testing dataset. Then, we again predicted our data points but using this first biased estimation of MPA as the value of MPA in equation 2.2. We ran this loop ten times (sufficient to reach a plateau) to obtain an unbiased estimation of both data points and $\text{MPA}_{\text{testing}}$. The R script of this method is provided in the Appendix 2.4a (step 1) and Appendix 2.4b (step 2).

2.3.5 *Statistical comparison of the methods*

2.3.5.1 Training and testing datasets

Using data from group A (each sex separately), we performed a repeated (200 times) random subsampling cross-validation procedure with 80 % of our data as the training dataset and 20%

as the testing dataset, used for all three methods. The group B datasets (separated by sex) were further used as additional testing datasets and were estimated from models fitted on the whole group A datasets. Table 2.1 is a summary of the training or testing datasets used, of the methods used to evaluate their quality, and of the corresponding results.

2.3.5.2 Comparison of the tree classification, the standard model, and the recursive model

We compared the three approaches based on their ability to estimate the MPA of testing datasets. We based our comparisons on the pattern of method's errors (error = estimated value − observed value) when estimating $MPA_{testing}$ (error > 0 represents an overestimation) and on their average precision (precision = |error|, i.e., the absolute error). This way of estimating the goodness of fit is applicable to both tree and logistic regression, contrary to the area under the receiver operating characteristic curve method for tree classification and to the estimation of bias and spread for logistic regression (see Pearce and Ferrier 2000 for both). We also

Table 2.1 Summary of the training and testing datasets, the equations and the methods used to evaluate the quality of the estimations for the two parts of the study: a) for the comparison of the tree classification, the standard model and the recursive model and b) for the evaluation of the quality of the prediction from the recursive model applied on the Tellus collar for reindeer. We also provide reference to tables and figures where results can be found

Training dataset	Testing dataset	Equation	Quality evaluation	Results section
a. Comparison of methods				
80% of Group A (200 replicates)	20% of Group A (200 replicates)	Tree: “rpart” function Standard model: Eq. 2.1 Recursive model: Eq. 2.2	Average MPA observed – average MPA estimated. For Tree: % of good classification Influence of the difference MPAtraining-MPA testing	Table 2.2 Figure 2.3
100% Group A	Group B	Tree: “rpart” function Standard model: Eq. 2.1 Recursive model: Eq. 2.2	Comparison MPA estimated versus Real MPA	Table 2.2
b. Application of the Recursive model to Tellus collar for reindeer				
80% of Group A (200 replicates)	20% of Group A (200 replicates)	Recursive model: Eq. 2.3	Bias and Spread evaluation method	Figure 2.3 Table 2.3
100% Group A	Group B	Recursive model: Eq. 2.3	Bias and Spread evaluation method	Fig 2.1, 2.2, 2.3 Table 2.3 Eq. 2.4

calculated the percentage of good classification of the models, when a threshold of 0.5 was applied to the observed and predicted values of percentage of activity to allow comparison of the discriminatory ability of the recursive and the standard model with the Tree50 classification method.

We assessed whether estimations of $\text{MPA}_{\text{testing}}$ were biased by modeling the relationship between the difference $\text{MPA}_{\text{training}}-\text{MPA}_{\text{testing}}$ and the errors of estimations by a linear model. If the difference of MPA does not induce biased estimations, we should observe a slope of 0. A bias, which is independent to the difference of MPA, is observed if the intercept of this relationship is different from 0. We compared average absolute errors among the recursive model, the standard model, and the Tree50 by paired t test in order to determine which method is the most precise. As an example, we compared the average estimates of MPA to the real value of MPA of group B dataset (separated by sex), using each of the three methods fitted on the whole group A datasets.

2.3.6 Application for evaluating activity level of reindeer from Tellus activity sensor

We used the recursive model to estimate the relationship between the Tellus activity sensor values and the level of activity of the male and female reindeers. However, the relationship fitted by the equation 2.2 for the male presents a default. The decrease of the prediction of the activity level at the highest values of X_{adj} and Y_{adj} (Fig 2.1a) is an artifact induced by the inclusion of the square terms in equation 2.2 and by the absence of behavioural observations corresponding to these values of activity sensor. Resting and standing which form our inactive class are not expected to produce high neck movements. Therefore, in order to fit better the particular pattern studied here (active versus inactive), we refined equation 2.2 to fit a monotonic pattern (i.e., the activity can only increase if X_{adj} and Y_{adj} increase) by the equation 2.3 below:

Eq. 2.3 $$Activity \sim f\left(X_{adj} + Y_{adj} + MPA\right)$$

We followed the procedure given by Pearce and Ferrier (2000) to evaluate the calibration (the agreement between predicted probability of activity and observed proportion of activity) of our model in terms of bias (consistent overestimate or underestimate of the probability of activity) and spread (a systematic departure of the regression line, fitted to the predicted and observed values, from a gradient of 45°; a positive spread (i.e., slope >1) indicates that predicted values greater than 0.5 are underestimating the percentage of activity and that predicted values less than 0.5 are overestimating the percentage of activity, Pearce and Ferrier 2000), by modeling the observed data as a function of the logit of the predicted values, using a logistic regression. We then used a series of likelihood ratio tests to evaluate bias and spread: first, we tested whether the intercept a = 0 (no bias) and the slope b = 1 (no spread), and if not, a second test evaluated the bias given the appropriate spread (b forced to be 1), and finally, we evaluated the spread given no bias. In order to evaluate the generalization of our model to individuals, other than the female and male of group A, we applied this calibration evaluation to the training dataset, the testing dataset, and the group B dataset to get internal, intra-individual, and inter-individual estimations. We pooled the observations and predictions of the 200 subsamples previously used for each sex before running the procedure for internal and intra-individual datasets. Group B dataset was only evaluated once from the equation obtained on the whole group A dataset (each sex separately). We calculated average precision (i.e., absolute errors) on the estimation of MPA as well as on the estimation of each data point on the training (i.e., the residual errors), the testing (the intra-individual errors), and the group B (the inter-individual errors) datasets for both sexes.

Finally, we called equation 2.3 the “final equation” because it was fitted from the whole group A dataset for each sex. The parameter estimates from these equations are provided and are the ones that should be applied in other studies. All analyses and comparisons were performed with R 14.0 (R Development Core Team 2011).

Except for figure 2.1, all figures are displayed only for the male dataset (see Appendix 2.5 for the female's figures). Graphical representations of equations are based on the equations fitted on the whole group A dataset, for each sex. Values of the mean errors and precision are presented with their ±standard deviation, i.e., the standard deviation of the means found in the 200 training and testing datasets, while slope and intercept values are reported with their ±standard error.

2.4 Results

2.4.1 Behavioural data and activity sensor value pattern

We obtained 306 data points for the male of group A (hereafter male A), 297 for the female of group A (female A), 44 data points from nine different males of group B (males B), and 49 data points from four different females of group B (females B). The observed mean percentage of activity was 48.87 % for male A and 58.63 % for female A. Adjusted values of the activity sensor did not exceed 0.57 on the x-axis and 0.65 on the y-axis for male A, with a statistically significant correlation between the values of the two axes (Pearson's correlation, $r = 0.60$, $p < 0.01$). Ninety-one ordered pairs $[X_{adj},Y_{adj}]$, or 29.7 % of the male A dataset, were around [0,0], i.e., less than 0.025 for the both axes. For female A, maximal values observed for X_{adj} and Y_{adj} were 0.67 and 0.47, respectively, with a statistically significant correlation between the axes ($r = 0.65$, $p < 0.01$). Sixty-six ordered pairs $[X_{adj},Y_{adj}]$, corresponding to 22.2% of the values, were around [0,0].

2.4.2 Graphical representation of the recursive model

Predictions of the percentage of activity for each ordered pair [X_{adj},Y_{adj}] as well as the distribution of the observed ordered pairs are displayed for MPA values corresponding to the real percentage of activity of male A and female A (MPA = 0.4887 and MPA = 0.5863, respectively, Fig 2.1). Graphical analysis showed that both axes of the activity sensor were important to describe variations of the level of activity for male A (Fig 2.1a), but the variations on the x-axis seemed to explain more variability than the y-axis for female A (Fig 2.1b).

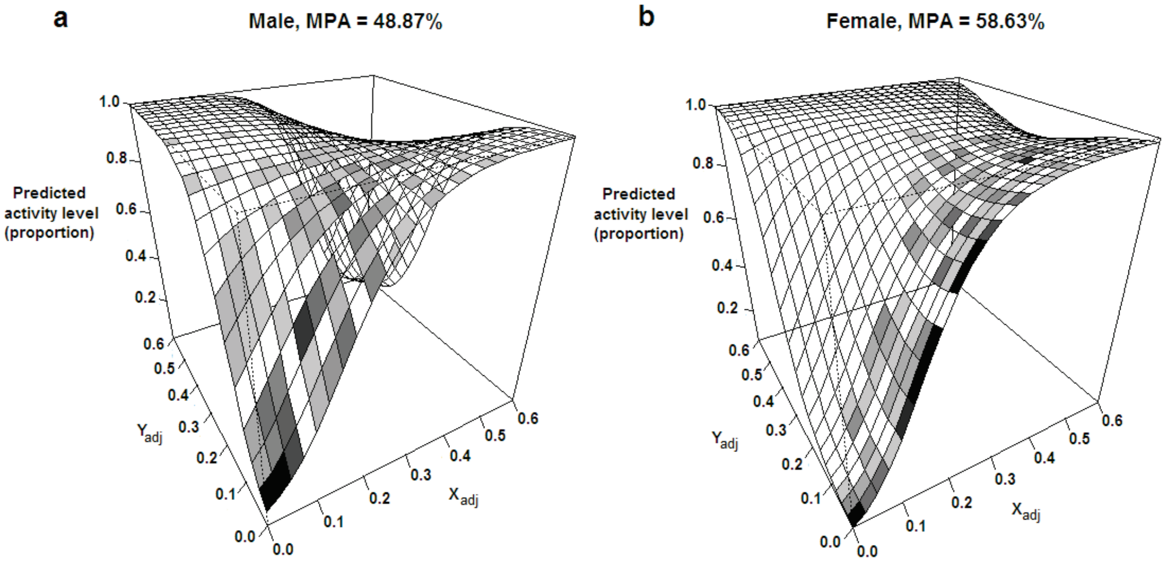


Figure 2.1 Predictions of the percentage of activity as a function of the back-forth (X_{adj}) and left-right (Y_{adj}) movements of the collar from the final equation for the (a) male and (b) female datasets. Values are for MPA equal to the observed MPA of Group A (MPA = 48.87% for the male and 58.63% for the female). The gradient of shading within squares represents the distribution of the observed ordered pairs [X_{adj};Y_{adj}] of Group A and range from 0 (white) to 13 or 8 (black) data for male and female respectively. Squares close to [0,0] have higher number (91 and 66 for male and female respectively) of data points and are darker. Total numbers of data points are 306 male and 297 female

2.4.3 Comparison of methods

2.4.3.1 Bias

We found that the difference of MPA between the training and the testing dataset significantly biased our estimations of $\text{MPA}_{\text{testing}}$ for the standard model and for the Tree50 on male A ($p < 0.001$ for both sexes; Table 2.2 and Fig 2.2), but only presented a trend for the female A Tree50 ($p = 0.07$). These two methods also had a significant bias independent of the difference of the MPA between the training and the testing dataset, overestimating the $\text{MPA}_{\text{training}}$ (all $p < 0.001$; Table 2.2 and Fig 2.2). On the contrary, estimations of the MPA from the recursive model were not biased in respect to the difference of MPA of the training and testing dataset ($p = 0.60$ and $p = 0.30$ for male A and female A, respectively; Table 2.2 and Fig 2.2) and were not biased independently of the difference of MPA for the female ($p = 0.32$). The bias independent to the difference of MPA was statistically significant for the male

Table 2.2 Comparison of the tree classification, the standard logistic model and the recursive models in term of bias dependent on the difference of mean percentage of activity (MPA) between the training and the testing datasets from the reindeer in Group A, of bias independent of this difference and on the precision of the estimation. The goodness of classification is also provided for trees. The provided goodness of classification of the recursive and the standard models are calculated for a threshold at 50%. Values are shown with their ± standard deviation (based on the estimation of the 200 training and testing datasets)

Group A	Method	Bias dependent on the difference of MPA	Bias independent of the difference of MPA	Mean of absolute MPA errors (% of activity)	Percentage of good classification (Tree only, %)
Male	**Recursive M.**	-0.02 ±0.05 NS	-0.006 ±0.003*[1]	3.16 ±2.55	(84.8 ±4.1)[2]
	Standard M.	0.19 ±0.03***	0.04 ±0.002***	4.67 ±2.73	(85.6 ±4.0)[2]
	Tree 50	0.20 ±0.07**	0.06 ±0.004***	7.58 ±4.84	83.1 ±4.0
Female	**Recursive M.**	-0.05 ±0.05 NS	-0.003 ±0.003 NS	3.64 ±2.50	(84.8 ±4.1)[2]
	Standard M.	0.18 ±0.04***	0.07 ±0.002***	6.65 ±3.25	(86.9 ±4.1)[2]
	Tree 50	0.15 ±0.08 NS	0.11 ±0.005***	11.50 ±6.13	84.1 ±4.1

Significance values are shown with stars: *(p<0.05), **(p<0.01), ***(p<0.001)

Non-significant values are noted as NS (p>0.05)

[1] This statistically significant bias has a small effect size and is disregarded in the discussion.

[2] These goodness of classification are provided for information, they are not the suitable measure of the model quality.

A (p = 0.04); however, the extent of the bias is biologically negligible (bias < 1 % of activity; Table 2.2 and Fig 2.2).

2.4.3.2 Predictive performance

Recursive models had significantly lower absolute errors on MPA than standard models and Tree50s, and the standard models had significantly lower absolute errors than Tree50s (all paired t test for both sexes, $p < 0.001$; Table 2.2; see Appendix 2.6 for a boxplot). Similarly, the percentages of good classification of the models were slightly better for a threshold at 0.5 than for the Tree50 (all paired t test for both sexes, $p < 0.05$, Table 2.2). Real MPA for males B was 82.28 % and was best estimated by the recursive models (79.56 %), followed by the standard models (85.68 %), and finally by the tree classification method (86.36 %). Similarly, the real MPA of females B (57.32 %) was much more precisely estimated by the recursive models (57.46 %) than by either the standard model (63.24 %) or the tree classification method (67.35 %).

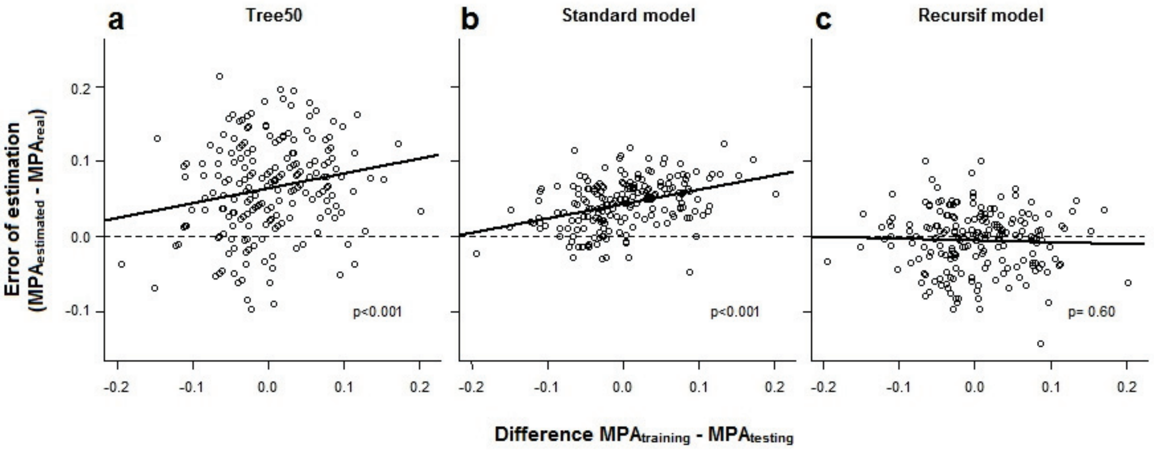


Figure 2.2 Comparison of the estimation of the mean percentage of activity of the Group A testing dataset (MPAtesting) of the different methods for the male dataset: (a) the Tree50 method, (b) the standard model, and (c) the recursive model. A bias dependent of the difference between the MPA of the training and testing datasets (positive values mean MPAtraining > MPAtested) is observed if the slope of the regression is different to 0 (the p value is provided on the figures). A bias independent of this difference is observed if the regression line (continuous) does not intersect the origin, i.e. at [0,0]

2.4.4 Application of Eq. 2.3 and evaluation of the model's calibration

Calibration did not show bias nor spread when tested on the 200 subsamples of the training dataset (overall test: $p = 0.37$ and $p = 0.75$ for male A and female A, respectively; Table 2.3 and Fig 2.3). We found a similar pattern for the female A testing datasets (overall test, $p = 0.09$), but the male A testing datasets had a slight spread with no bias (overall test, $p < 0.01$; test of bias, $p = 0.20$; test of spread, $p < 0.01$; slope = 0.95 ± 0.02; Table 2.3 and Fig 2.3). When applied to the group B datasets, we did not detect significant bias or spread for both sexes (overall test, $p = 0.07$ for both male B and female B; Table 2.3 and Fig 2.3). Equation 2.3 provided a more precise estimation of $MPA_{training}$ than $MPA_{testing}$ for both sexes (Table 2.3). However, the estimations of $MPA_{group\ B}$ had a comparable precision than the estimations of $MPA_{testing}$ (Table 2.3) for both sexes. The precisions obtained on each value were comparable among training, testing, and group B datasets for both sexes (Table 2.3). The final equation ran on the entire dataset of male (Eq. 2.4a) and female (Eq. 2.4b) A is

Table 2.3 Evaluation of the calibration of the recursive model from equation 2.3 for males and females datasets. The evaluation of the goodness of fit (the difference between the observed values of activity for each data and their estimation) is evaluated at an internal level (i.e. on the 200 training datasets from Group A), at an intra-individual level (i.e. on the 200 testing datasets from Group A) and at an inter-individual level using the final equation (from the whole Group A dataset) applied on the Group B dataset

Calibration	Bias	Spread	Precision on MPA	Precision on each data
Male				
Internal	0.006 ± 0.012^{NS}	0.989 ± 0.009^{NS}	0.19% ±0.15	18.3% ± 16.3
Intra-individual	0.014 ± 0.025^{NS}	$0.950 \pm 0.017^{**}$	3.24% ± 2.58	18.5% ± 16.6
Inter-individual	-2.478 ± 1.661^{NS}	1.733 ± 0.688^{NS}	5.8%	6.9% ± 14.7
Female				
Internal	0.009 ± 0.013^{NS}	1.003 ± 0.009^{NS}	0.21% ± 0.16	17.5% ± 17.3
Intra-individual	0.028 ± 0.026^{NS}	0.968 ± 0.0165^{NS}	3.69% ± 2.60	17.8% ± 17.6
Inter-individual	-0.581 ± 0.455^{NS}	0.646 ± 0.162^{NS}	5.2%	15.9% ± 24.0

Significance values are shown with stars: *(p<0.05), **(p<0.01), ***(p<0.001)
Non-significant values are noted as NS (p>0.05)

presented below:

Final equations:

Eq. 2.4a: Male $Activity \sim f\left(-3.44 + 11.29 \times X_{adj} + 11.12 \times Y_{adj} + 3.00 \times MPA\right)$

Eq. 2.4b: Female $Activity \sim f\left(-3.90 + 14.72 \times X_{adj} + 8.63 \times Y_{adj} + 2.50 \times MPA\right)$

where X_{adj} and Y_{adj} are the back–forth and left–right movement values of the collar adjusted by the TTF; Activity is the estimated proportion of time spent active of each data; MPA is the estimated mean percentage of activity of the dataset; and *f(x)* is the logistic function.

2.5 Discussion

In this study, we developed a new method of validation of activity sensors, the recursive model, which we compared with the common tree classification method and the standard logistic model. Our method successfully corrects for the disadvantages (the simplification of data and the biased estimation, discussed below) presented by the two other approaches and performs with improved accuracy of the estimations.

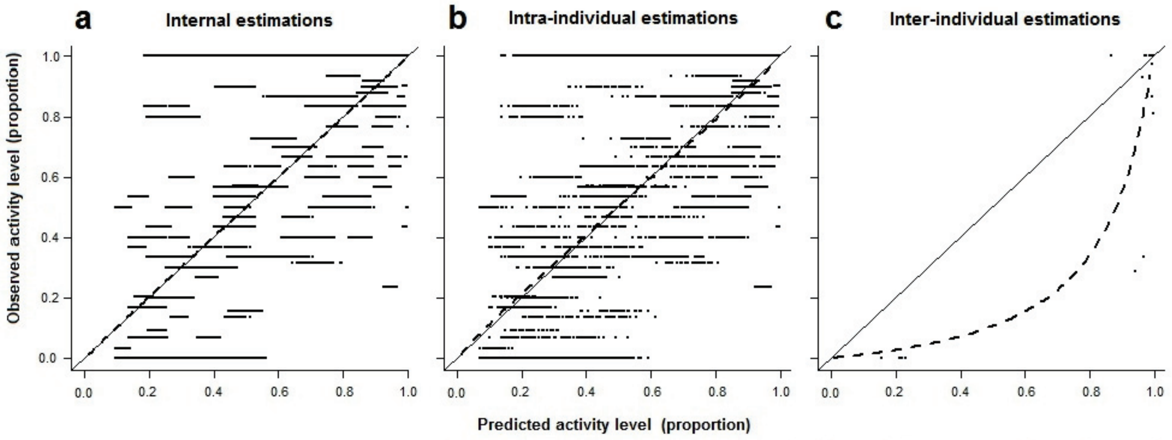


Figure 2.3 Calibration plots for (a) the internal estimations (on the Group A training datasets), (b) the intra individual estimations (on the Group A testing datasets) and (c) the inter-individual estimations (on the Group B dataset from the final equations) for the male dataset. We pooled together the 200 sub-datasets for panels (a) and (b). Continuous lines represent a perfect calibration with no bias or spread. Dotted lines represent the fitted relationship between the observed values and the logit of their estimation by a logistic regression (see Pearce and Ferrier 2000)

2.5.1 Data simplification

Differences among tree classification methods currently used come mostly from the pre-work simplification of data necessary to match the binary framework of the tree classification method. Deleting mixed data is clearly inducing uncontrolled changes on further estimations of values that are partially active. Changing the definition of active in order to include standing behaviour does not lead to wrong estimations; however, it limits the prospective information given by the activity sensor. Using a model which estimates percentage of activity rather than pure active or inactive time allows us to avoid finding a methodology to simplify the data and it better represents the authentic activity of the individual.

2.5.2 Biased estimations

Our results demonstrate that tree classification methods and the standard model provide biased estimations, overestimating low level of activity and underestimating high level of activity; the bias being higher with the difference in the MPA of the datasets. This bias is in accordance with statistical advancements made on case–control studies and is due to the well-known problem of unbalanced training dataset compared to the testing dataset (Cramer 1999; Oommen et al. 2011). This problem was previously discussed by Adrados et al. (2003), who advocated the use of a more “balanced” training dataset. However, this training dataset should be balanced in regard to its testing dataset, i.e., they should have the same MPA, which is impossible if we want to apply our validation to one or several unknown datasets. The consequence is that there are no training dataset characteristics that could improve the accuracy of the method for every testing datasets. Our recursive model eliminates such bias. In the absence of a mathematical method, such as the prior correction or the weighting method of King and Zeng (2001), to correct the bias due to the unbalanced dataset, we acknowledge that our method is sensitive to fluctuation due to sampling errors.

Biased estimations are not a problem within a study as far as it is constant. However, we show that the extent of the bias increases with the difference between the $MPA_{training}$ and the $MPA_{testing}$. The bias is toward the $MPA_{training}$ and will consequently decrease the difference between two samples that have different MPA. Therefore, the use of the recursive model will increase our ability to detect statistically significant differences between samples. However, the recursive model is mostly interesting when used to compare values from different studies that do not use the same training dataset, as the estimation will be independent of their respective training datasets. The observation of animal behaviour within enclosure may not necessarily be representative of the natural activity budget. This is not an issue here as the recursive model takes into account this limitation. Indeed, as long as the neck movement of the animals represents the natural neck movement of their behaviour in the wild, our model can be applied.

2.5.3 Predictive performances

The recursive model appeared to be the most accurate method compared to the standard model and to the tree classification method. Moreover, it is the only one which has no bias independent to the difference of MPA when estimating the MPA of the training as well as the group B datasets. The precision, or reliability, was better for the recursive model than for the Tree50 method, i.e., the predicted values of activity were closer to the actual values with the recursive model. Increasing the precision allows researchers to detect trends more easily. More importantly, estimating a continuous value of activity rather than a binary value allows researchers to investigate disturbance of a certain behaviour, for example the extent of vigilance during feeding bouts or the interruption of feeding bouts by the presence of young males during the rut (Holand et al. 2006). The use of a continuous value to represent the activity matches more closely the real proportion of activity when the recorded lapse of time is long (e.g., 5 min in Lotek GPS collars, Löttker et al. 2009) and it represents the proportion of data that are active or inactive when the lapse of time is really short (e.g., 3.5 s in ACTi-GPS, Brown et al. 2012). In order to be able to capture slight variation of the percentage of activity, a good calibration and precision of our equations are necessary.

The recursive model had a slightly higher discriminatory power than the Tree50, i.e., it better distinguished active versus inactive records. The recursive model had a goodness of classification (84.8 % including classification as active or inactive) similar to those found in the literature (e.g., percentage of good classification of 84 % for active and 97 % for inactive data, Gottardi et al. 2010; Appendix 2.1). Favoring the discriminatory power rather than the precision would be helpful to detect behaviour sequences but is not recommended if one wants to estimate activity budget or daily activity pattern.

2.5.4 Other advantages of the recursive model

When using a recursive model rather than a threshold method, first we can include square terms and interactions in our models, thereby increasing its level of flexibility in fitting any pattern of data. Therefore, the definition of the two classes of behaviours (here active versus inactive) can be changed to any association of behaviours, including grouping behaviours that are expected to have really different patterns of activity sensor values, the only restrictions being the two- category classification and the biological significance. For example, by including a unique behaviour to one category (e.g., fitting frequency of "eating" against all other behaviours), we can study the frequency of a specific behaviour without using a multi-category classification type of analysis that requires more complex validation procedures (see methods based on tree classification in Ungar et al. 2005, Löttker et al. 2009, and Moreau et al. 2009). That being said, our suggested approach can also be developed to allow use of multiple categories. Moreover, more flexibility can be added to this model by fitting a general additive model rather than a GLM as presented here. Second, it is easy to include information from other motion sensors, as neck direction or the up–down movement, in the recursive model, and more information should lead to better precision of the estimates. Third, the definition of what is active or inactive behaviour (i.e., the cut-off value) can be applied after the activity level estimation, when using our method. Applying this cut-off definition is easier with the recursive model as no further validation is necessary when the cut-off is changed and it allows using more than two classes. These cut-off values will also be determined based on a biological view point rather than a statistical view point (e.g., to increase the percentage of good classification). To this respect, the recursive model performs at least as well as a tree classification method (Table 2.2). Therefore, the recursive model is a better method to estimate the proportion of time spent in active behaviours and could be applied in parallel to methods that estimate the state of activity of individuals based on their trajectories (Van Moorter et al. 2010; Owen-smith et al. 2012). Using data from two different origins to assess the level or state of activity of the individuals could strengthen the confidence in the results.

2.5.5 Validation of the Tellus activity sensor on reindeer

The use of equation 2.3 in this application helped to remove the artifact originally created by the use of the more flexible equation 2.2. Estimations of the level of activity from Tellus activity sensor by the recursive model are accurate and do not show bias or spread for the female dataset and only present a small spread for the male dataset. Moreover, we found that those estimates are still accurate when applied to other male and female collars. Therefore, the male and female equation provided could be directly applied to other reindeer when using the same type of collars and with the same sensitivity. However, we acknowledge that our small sample size from group B datasets may prevent detection of bias or spread, and more attention should be given to inter-individual validation. Variability among males is expected to be higher, as there is more variation in male morphology and behaviour than in female's, especially during the breeding season. There is also a difference between the male and the female equations as an increase of the left–right movements of the collar means that the individual is more active for males, whereas it is unrelated to the level of activity for females. The variability among individuals of the relationship between the activity sensor records and the actual behaviour has been overlooked in previous validations. We show here that the variability could be important and therefore could decrease the accuracy of the previous validations, especially when behavioural differences in the sexes of highly dimorphic species have not been taken into account.

The application of the provided equations is however subject to limitations, particularly with respect to seasonality. In winter, accumulation of snow on the collar may disturb collar’s movements, and foraging behaviour is slightly different, as reindeer dig snow and brake ice to get access to grass. In summer, harassment by insect (Weladji et al. 2003) will influence many GPS collar records. During calving, social interaction and nursing may create new behaviour. All of these behaviours should be taken into account before using the final equations provided here during specific sea- sons. The sensitivity parameter (here fixed at 5) is also a key parameter and any modification prevents the use of our equations (for new studies, Tellus GPS sensitivity parameter could be slightly increased in order to obtain a larger range of adjusted values).

2.6 Conclusion

Gottardi et al. (2010) highlighted the lack of a standard method for activity sensor validation. The recursive model could be an important step in the standardization of this method. It is accurate, unbiased, and does not require simplification of the observed behavioural data. Any researcher or manager working with a dual axis activity sensor, or other sensors with higher dimensions, can apply the recursive model method to relate their own observed behavioural data with the values provided by the activity sensor (see Appendix 2.4a for the R script). Users with reindeer equipped with Tellus collars can directly apply the final equations under appropriate conditions. In order to facilitate the application of these equations, we provide the R script of the recursive model (Appendix 2.4b).

2.7 Appendices

Appendix 2.1 Table 2.4 summarizes previous validations of activity sensors for ungulates.

Appendix 2.2 Complementary analysis of the Tree30 and the Tree “pure”

Method

The Tree30 has been established on the same basis as the Tree50, except that a proportion of activity ≤0.3 was classified as inactive and a proportion of activity > 0.3 as active. We used the same training and testing dataset from Group A than for the other methods.

In order to establish the Tree based on “pure” data, we removed from each of the 200 training dataset the values which were neither “0” nor “1”. However, we kept all of the data from each testing dataset, as this difference cannot be performed *a priori*. Consequently we were not able to calculate a percentage of good classification for the Tree “pure” as we should have compared predictions that are in 2 classes to observations that are on a continuum scale.

Table 2.4 This table briefly describes the objectives of the validation (mainly Active/Inactive, Eat/Walk/Rest or a percentage of activity), the statistical method (mainly Tree method, a discriminant function or a regression model), and the type of data simplification performed before the analysis. It also provides other information such as the sample size, the number of axis on the collar and the dimension of the analysis (1D if only one variable is taken into account), as well as details about the percentage of good classification when provided (or its estimation when data were available). References are ordered by date. We use small letter case to emphasize studies with different goals or methods.

Reference	Species	GPS collar	Sensor axis	Dimension of the analysis	Number of individuals for calibration	Behaviours in focus	Inter-individual variability	Statistical method	Data simplification	% good classification
Relyea et al. 1994	Desert mule deer	-	One axis	1D	9 males 1 female	**a:** Eat/Walk/Rest **b:** Bedded/nonbedded **c:** Active/Inactive	Not estimated	Tree	Dominant behaviour	**a:** Failed **b:** 73-77% **c:** 74-81%
Moen et al. 1996	Moose	Lotek 1000	Dual axis	1D	1 female	**a:** Active/Inactive **b:** % activity	Not estimated	**a:** Tree **b:** Regression	Pure data Only mixt data	**a:** 91% (at 150) **b:** failed
Adrados et al. 2003	Red deer	Lotek 1000	Dual axis	1D	4 females	Active/Inactive	Individual-based method	Graphical method	Cut-off (3 classes)	66-81%
Naylor and Kie 2004	Rocky mountain elk	LORAN-C *Actiwatch*	Omni-directional	1D	8 (unspecified)	Eat/Walk/Rest	Not estimated	Discriminant function	Pure data	88%
van Oort et al. 2004	Reindeer	*Actiwatch*	Omni-directional	1D	5 males	Active/Inactive	Not estimated	Graphical method	Dominant behaviour	93-100%
Ungar et al. 2005	Cattle	Lotek 2000 Lotek 2200 *CW 1600 serie*	Dual axis	3D	5+9 (unspecified)	Eat/Walk/Rest	Included in the model	**a**: Regression **b**: Discrimination f. **c**: Tree	None or natural “pure” data	**a:** - **b:** 86% **c:** 88%
Coulombe et al. 2006	White tailed deer	LMRT3-VHF *STO-2a* Lotek 2200	One axis Dual axis	1D	4+4	Active/Inactive	Not estimated	Tree	Pure data Stand as Active	60-92% 83-92%
Gervasi et al. 2006	Brown bear	GPS-GSM *Vectronic Aerospace GmbH*	Dual axis	1D	1 male 1 female	Active/Inactive % activity	Individual-based method	Tree	Cut-off value	95%
Pépin et al. 2006	Red deer	*ETHOSYS*	One axis + head position	2D	3 females	Active/Inactive	Not estimated	Tree	Pure period of lying is inactive, the rest is active	77-99%
Loe et al. 2007	Reindeer	ARGOS	Head position	1D	-	Active/Inactive =head down/up	Not estimated	Tree	-	-
Bourgoin et al. 2008	Mouflon	Lotek 3300S	Dual axis + head position	3D	4 females	Active/Inactive	Standardization of sensors values among individuals	Discriminant function	Pure data Stand as active	85-87%
Kozakai et al. 2008	Japanese black bear	Lotek 3300S Lotek 4400S	Dual axis	1D	2 females 1 male	**a:** Active/Inactive Intensity of activity when active	Not estimated	**a:** Tree	**a:** Pure data	**a:** 98%

Umstätter et al. 2008	Sheep	BlueSkyTelemetry 200 series	One axis + head position	**a:** 4D **b:** 1D **c:** >4D	10 (not clear)	Active/Inactive	Not estimated	**a:** Discriminant f. **b:** Tree **c:** Graphical method	-	**a:** 93-97% **b:** 93-97% **c:** 91%
Löttker et al. 2009	Red deer	GPS-GSM Vectronic Series 600-2100-2300	Dual axis	1D	3 females 1 male	Eat/Run/Rest	Not estimated	Tree	Pure data	93%
Moreau et al. 2009	Goat	Vectronic *HOBO pendant G tri-axial*	Tri-axis	3D	3	Eat/Walk/Rest	Not estimated	Tree	Dominant behaviour	20-93%
Gottardi et al. 2010	Roe deer	Lotek 3300	Dual axis + head position	3D	2 females 1 male	Active/Inactive	Standardization of sensors values among individuals	Discriminant function	Cut-off value	84-97%
This study	Reindeer	Tellus	Dual axis	2D	1 male 1 female + 13 others	% activity	Sex separated + inter-individual variability evaluation	Recursive model	-	Comparable to 85-87% (at 0.5)

Results

We completed table 2 with the values of the Tree30 and Tree "pure" (Table 2.5). The three Trees displayed a significant bias independent to the difference $MPA_{training}$-$MPA_{testing}$ and only the Tree "pure" displayed a bias dependent to this difference for both sexes (Table 2.5, Fig 2.4). The Tree30 and the Tree "pure" had lower precision (mean absolute errors) on MPA than the Tree50, which had itself a lower precision than the standard and the recursive models. However, the Tree30 had a higher goodness of classification than the Tree50.

Table 2.5 Comparison of the tree classifications (Tree50, Tree30 and Tree "pure"), the standard logistic model and the recursive model in terms of bias dependent on the difference of mean percentage of activity (MPA) between the training and the testing datasets from the reindeer in Group A, of bias independent of this difference and on the precision of the estimation. The goodness of classification is also provided for trees. The provided goodness of classification of the recursive and the standard models are calculated for a threshold at 50%. Values are shown with their ± standard deviation (based on the estimation of the 200 training and testing datasets)

Group A	Method	Bias dependent on the difference of MPA	Bias independent of the difference of MPA	Mean of absolute MPA errors (% of activity)	Percentage of good classification (Tree only, %)
Male	Recursive M.	-0.02 ±0.05 ^{NS}	-0.006 ±0.003*[1]	3.16 ±2.55	(84.8 ±4.1)[2]
	Standard M.	0.19 ±0.03***	0.04 ±0.002***	4.67 ±2.73	(85.6 ±4.0)[2]
	Tree 50	0.20 ±0.07**	0.06 ±0.004***	7.58 ±4.84	83.1 ±4.0
	Tree 30	0.09 ±0.05 ^{NS}	0.15 ± 0.003***	14.94 ±4.25	89.9 ± 3.5
	Tree "pure"	0.16 ± 0.06**	0.14 ± 0.004***	13.76 ±5.50	
Female	Recursive M.	-0.05 ±0.05 ^{NS}	-0.003 ±0.003 ^{NS}	3.64 ±2.50	(84.8 ±4.1)[2]
	Standard M.	0.18 ±0.04***	0.07 ±0.002***	6.65 ±3.25	(86.9 ±4.1)[2]
	Tree 50	0.15 ±0.08 ^{NS}	0.11 ±0.005***	11.50 ±6.13	84.1 ±4.1
	Tree 30	0.12 ±0.06 ^{NS}	0.14 ±0.004***	14.31 ±5.22	90.7 ±3.4
	Tree "pure"	0.19 ±0.06**	0.12 ±0.004***	12.61 ±5.10	

Significance values are shown with stars: *($p<0.05$), **($p<0.01$), ***($p<0.001$)

Non-significant values are noted as NS ($p>0.05$)

[1] This statistically significant bias has a small effect size and is disregarded in the discussion.

[2] These goodness of classification are provided for information, they are not the suitable measure of the model quality

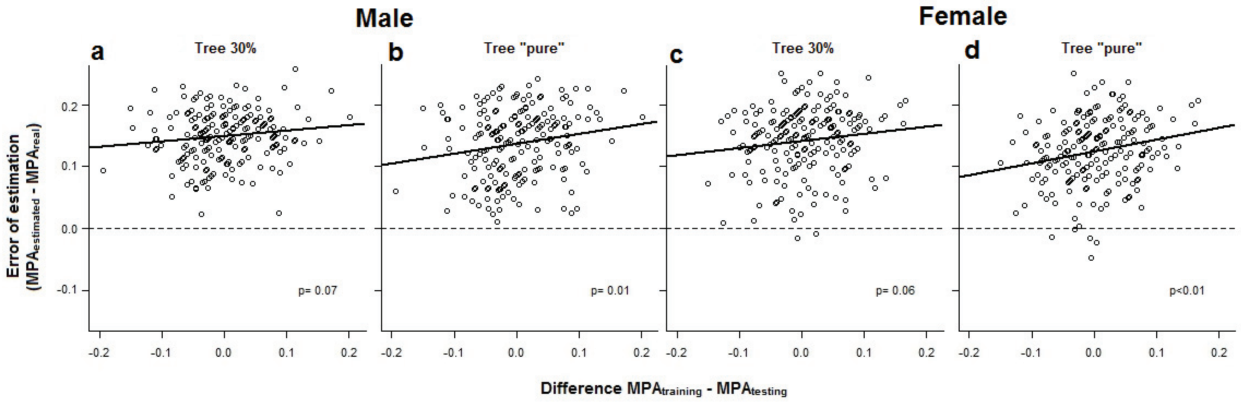


Figure 2.4 Comparison of the estimation of the mean percentage of activity of the Group A testing dataset ($MPA_{testing}$) from Tree30 (a, c) and Tree "pure" (b, d) for male (a, b) and female (c, d) datasets. Bias dependent of the difference between the MPA of the training and testing dataset (positive values mean $MPA_{training} > MPA_{tested}$) is observed if the slope of the regression is different to 0 (the p value is provided on the graphic) and bias independent of this difference is observed if the regression line (continuous) does not cross the 0 line (dashed line) at the x-axis equal to 0. The four Trees have a significant independent bias whereas only the Trees "pure" have a significant bias dependent of the difference between the MPA of the training and testing datasets

Discussion

Changing the cut-off value to differentiate active to inactive lapses of time, or only using the "pure" data to set up the threshold value, has an influence on the estimation of the MPA as well as on its precision. The tree the most precise (the Tree50) based on the absolute errors is the one that has a cut-off value (50%) closest to the actual mean percentage of activity of the Group A dataset (48.87% and 58.63% for male and female). The Tree30 and the Tree "pure" have a low precision on the MPA as their estimations have 12-14% of activity of error.

However, the percentage of good classification is higher for the Tree30 than the Tree50. This dichotomy goodness of classification versus goodness of fit highlights the difference between the two measures of the quality of the trees. The goodness of classification reflects the discriminatory ability of the tree, i.e. the ability to separate the "active" to the "inactive" lapses of time whereas the goodness of fit (the precision) reflects the ability of the method to estimate each value of activity with the smallest errors. Selecting a cut-off value in order to increase the goodness of classification (as performed by Moen et al. 1996 as well as by Gottardi et al. 2010) is therefore not a guarantee of an increased precision. If one wants to use a tree classification method rather than the proposed recursive model, we suggest using a more adapted tool to evaluate the tree quality, i.e. the Area Under the Curve (AUC), the curve referring to the Receiver Operating Characteristic (ROC) curve (Pearce and Ferrier 2000). This tool also presents the advantages of not being influenced by the issue of unbalanced datasets (Rakotomalala 2011).

Fitting the tree only on the "pure" data does not improve the precision of the estimation of $\mathrm{MPA}_{\mathrm{testing}}$ (Table 2.5) and increases the sensibility of the tree to the issue of unbalanced dataset. Moreover, if one estimates a percentage of good classification on both "pure" training and testing datasets, one will overestimate the quality of the tree. Indeed, the testing datasets will never be made of "pure" data and the data that are the most difficult to classify are the "combined" data which have intermediate values for the collar movements. We do not encourage this common practice (Appendix 2.1).

Appendix 2.3 Application of the "prior correction" and "weighting" method (King and Zeng 2001) to continuous data

In order to evaluate the theoretical influence of an unbalanced training dataset on the evaluation of the testing dataset, as well as the efficiency of the "prior correction" and the "weighting" method developed by King and Zeng (2001), we ran two sets of simulated data, one with binary data, and the other with continuous data.

Method

The dataset 1 is composed of 2 vectors, one of “x” values generated from a normal curve (mean = 3, sd = 2) and one of “y” binary values with a probability of being “1” depending on the associated “x” values given by the equation 2.5:

Eq. 2.5

$$probability\ of\ "1" = {}^{1}\!/_{(1+\ e^{-3+x})}$$

The dataset 2 is composed of the same 2 vectors, “x” values being generated as above and the “y” continuous values being generated from the mean of 20 draws from the previous equation 2.5. Datasets generated are presented in figure 2.5a (binary data) and figure 2.5e (continuous data).

For both datasets, we sub-sampled 1000 times 50 data points in order to create 1000 training datasets of 50 data points each, and we did the same for 1000 training datasets of 50 data points. We fitted the relationship between “x” and “y”, from the training dataset, by a logistic regression to establish the models. We predicted the “y” values of the testing dataset from this model for the uncorrected predictions. We applied the “prior correction” by adding the following term into the linear predictor of the model (King and Zeng 2001) in equation 2.6:

Eq. 2.6

$$logit(y_i) = a + b \times x_i + ln\left(\frac{MPA_{testing}}{1- MPA_{testing}}\right) - ln\left(\frac{MPA_{training}}{1- MPA_{training}}\right)$$

Where “$a + b \times x_i$” is the linear predictor of the model, and $MPA_{training}$ and $MPA_{testing}$ are the means of “y” values in the training and testing dataset respectively.

We applied the “weighting” method by fitting a new model including a weight parameters calculated as followed (Eq. 2.7, King and Zeng 2001):

Eq. 2.7 $$weight_i = \frac{MPA_{testing}}{MPA_{training}} \times y_i + \frac{1- MPA_{testing}}{1- MPA_{training}} \times (1 - y_i)$$

For each prediction, we calculated the MPA$_{estimate}$ (mean values of estimated "*y*") of the sample, the MPA$_{training}$ of the training dataset, the MPA$_{testing}$ which is the real value of the testing dataset. We compared the relationship between the errors of estimation (MPA$_{testing}$-MPA$_{estimate}$) and the difference of MPA between the two datasets (MPA$_{testing}$-MPA$_{training}$) by a linear model. Unbiased estimation should give a slope of zero.

Results

We found that the uncorrected predictions were biased for the binary dataset ($p<0.001$, Fig 2.5b) but the predictions were corrected when the "prior correction" or the "weighting" method has been applied ($p=0.79$ and $p=0.89$ respectively, Fig 2.5c-d).

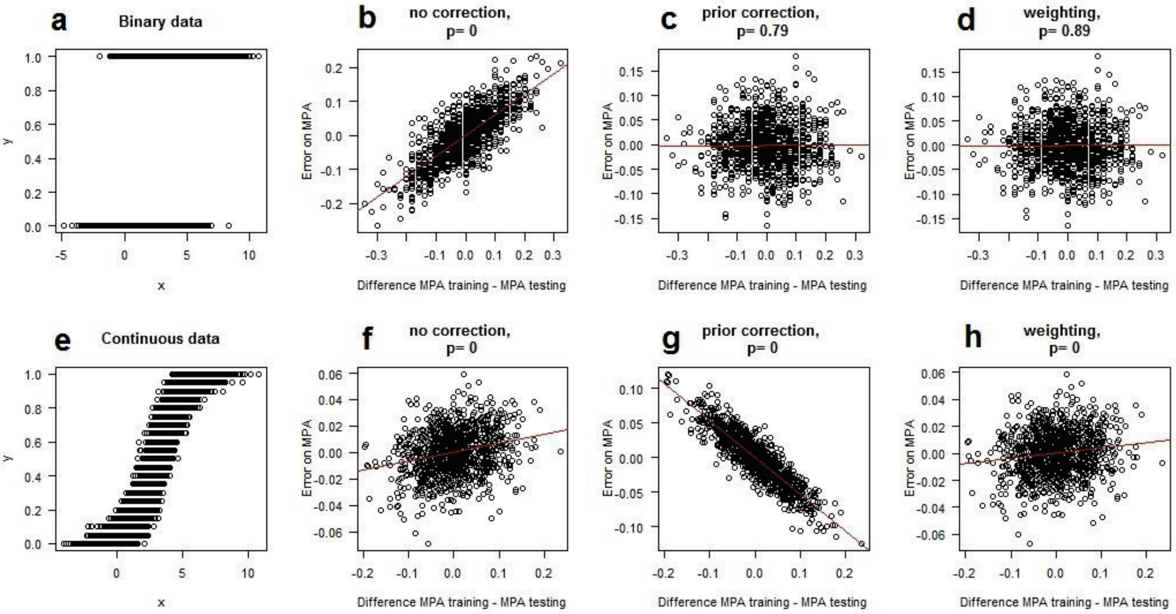


Figure 2.5 Generated binary and continuous dataset are presented for the relationship between "x" and "y" (a, e respectively). Predictions of MPA made from the training dataset to the testing dataset are presented for the binary dataset (b, c, d) and the continuous dataset (f, g, h) in relation to the difference of MPA of the training and testing datasets. MPA have been predicted without correction (b, f), with the prior correction (c, g) and with the weighting method (d, h) based on King and Zeng (2001). The p-values of the slope of the regression are presented below the method name

The uncorrected predictions of the continuous dataset were also biased, but in contrast to the binary dataset both corrections did not provide unbiased estimation of MPA (all $p<0.001$, Fig 2.5f-h).

Discussion

The prior correction or the weighting methods did not work with continuous data. This is the reason why we developed our methodology in order to predict the activity level from the values of the two axes of an activity sensor based on sub-sampling.

Appendix 2.4 R scripts

Appendix 2.4a: R script to set up the recursive model based on the equation 2.2

This script is made to set up your own final equations from equation 2.2. It neither runs cross-validation nor calibration procedures.

Red text indicates where the user can modify the script in order to apply it to its own dataset, to change parameters or definitions. Green text corresponds to comments and outputs.
Black text corresponds to the script itself.
Links between comments and script are presented by index #x# in blue.
Left panel corresponds to the script; right panel contains most of the comments and explanations.

Data used in this example comes from our female A dataset, given in a text file. (Female_activity.txt)

 # Building the recursive model equation

Script (Methods: Recursive model, Step 1)

```
############################
#     1) Downlowding data        #
############################
a<-read.table("Female_activity.txt",h=T)



### Parameters of the recursive model  ###
subsamplesize<-75                  #1# Size of subsamples
numbersubsamples<-250          # Number of subsamples


### Building the "Active" and "Inactive" class  ###
active<-a$E+a$W+a$F+a$BT
```

Outputs / Comments

Description:
Active behaviours: E=Eat, W=Walk, F=run (fast locomotion), BT=Bush Trash
Inactive behaviours: R=Rest, S=Stand
Sensor values: X= forward backward movement value, Y= left-right movement value
GPS information: TTF= Time to fix

We used a size of subsample approximately equals to 25% of the whole dataset #1#

The definition of what is “active” can easily be

```
inactive<-a$R+a$S
training<-
data.frame(active,inactive,X2=a$X/a$TTF,Y2=a$Y/a$TTF,TTF=a$TTF)
lengthtraining<-length(training[,1])
```

changed here

```
#####################
#   2) Subsampling      #
#####################

newMean<-c(rep(0, numbersubsamples))
newX=newY=newsuccess=newfailed=as.table(matrix(rep(0,
numbersubsamples * subsamplesize),nrow= subsamplesize))

for(i in 1: numbersubsamples) {
subsample<-training[sample(nrow(training), subsamplesize, replace=FALSE),
]
newX[,i]<- subsample $X2
newY[,i]<- subsample $Y2
newsuccess[,i]<- subsample $active
newfailed[,i]<- subsample $inactive
newMean[i]<-mean(subsample $active/ subsample $TTF) #2#
}

newtraining<-
data.frame(cbind(c(newX),c(newY)),c(newsuccess),c(newfailed),rep(newMean,each= subsamplesize)) #3#
colnames(newtraining)<-c("X","Y","Suc","Fail","MPA")
```

The subsampling procedure consists in sampling 250 times 75 data from the training dataset without replacement. For each subsample we record the mean percentage of activity (MPA, #2#)

“newtraining” #3# is the table containing the 250 subsamples of 75 data which are link to their subsample MPA. It has now 75*250 rows

e.g. of “newtraining”

X	Y	Suc	Fail	MPA
0.16	0.00	22	15	0.60
0.26	0.00	27	3	0.60
0.08	0.00	45	0	0.60
0.40	0.03	30	0	0.60
0.30	0.06	10	20	0.60
0.30	0.06	20	10	0.60

```
###################################
#     3) Fitting the logistic model        #
###################################
```

Here we apply equation 2.2 #4# to the “newdata” table

The “variation of the intercept of equation 1 with the variation of the $MPA_{training}$” is taken into account by the MPA coefficient.

```
newresponse<-cbind(newtraining$Suc,newtraining$Fail)
glmFinal<-
glm(newresponse~X*Y+MPA+I(X^2)+I(Y^2),data=newtraining,family=quasi
binomial(link=logit)) #4#
```

The equation can be modified if necessary, as we did to obtain equation 2.3. The function “func” (see below) will have to be changed too in such a case.

```
coeffFinal<-c(glmFinal$coefficients[1:7])   #5#
coeff<-coeffFinal
```

coefficients respectively for:
the intercept; X_{adj}; Y_{adj}; MPA; X_{adj}^2; Y_{adj}^2; $X_{adj} \times Y_{adj}$

```
write.table(coeff,file="coeff_Female_final_square.txt")  #5#
```

We save the coefficient of the Final equation for further use. #5#

```
######################################
#      5) Function for further application  #
######################################

func<-function(X2,Y2,MPA) 1/(1+exp(-(coeff[1]  + coeff[2]  *X2+ coeff[3]
*Y2 + coeff[4]  *MPA + coeff[5]*X2*X2 + coeff[6]*Y2*Y2 +
coeff[7]*X2*Y2 )))
```

Here is the Final equation as a function (see appendix 2.4b for an application)

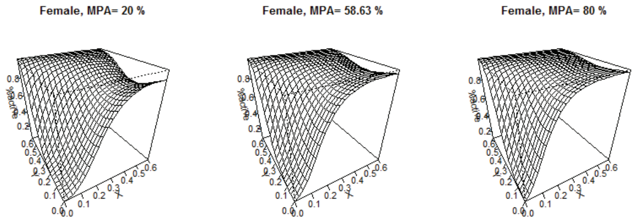
```
######################################
#      4) Graphical representation in 3D      #
######################################

# Enlarge horizontally the graphical windows to allow 3 panels
par(mfrow=c(1,3))
for(mpa in c(0.2,0.5863,0.8)) {

X3<-seq(0,0.6,len=25)
Y3<-seq(0,0.6,len=25)
newdata<-expand.grid(X2=X3,Y2=Y3)

fit.ii2<-matrix(func(X2=newdata$X2,Y2=newdata$Y2,MPA=mpa),25,25)
```

E.g of the representation, in miniature.



```
persp(X3,Y3,fit.ii2,theta=-
30,phi=30,ticktype='detailed',zlab="%active",xlab="X",ylab="Y",xlim=c(0,0.6
),ylim=c(0,0.6),main=paste("Female, MPA=",mpa*100, "%"))}
```

Appendix 2.4b: R script for application of the recursive model based on the final equations

This script is made to apply a preexisting equation 2.4 (Final equation) to a new dataset. Managers that wish to use our final equations (Eq. 2.4) can change the input file below and apply the following script. This script can also be used to apply the equation 2.2 obtained in the Appendix 2.4a.

Red text indicates where you can modify the script in order to apply it to your own dataset or to change parameters or definitions. Green text corresponds to comments and outputs.
Black text corresponds to the script itself.
Links between comments and script are presented by index #x# in blue.
Left panel corresponds to the script; right panel contains most of the comments and explanations.

Data used in this example are our female A dataset and is given in a text file. (Female_activity.txt).

Script (Methods: Recursive model, Step 2)

```
###############
#     INPUT     #
##############

data<-read.table("Female_activity.txt",h=T)

coeffM<-c(-3.44,11.29,11.12,3.00,0,0,0)     #6#
coeffF<-c(-3.90,14.72,8.63,2.50,0,0,0)      #7#

coeff<-coeffF       #8#
```

Outputs / Comments

”data” is your data file, presented as a dataframe with the following column
X Y TTF

Coefficients for the intercept, X, Y and MPA respectively, for males (coeffM #6#) and females (coeffF #7#) obtained from equation 2.3.
Notice that the last 3 coefficients are “0” as equation 2.3 does not take into account X^2, Y^2 and X:Y.

```
###############
#       RUN        #
################

testing<-as.data.frame(cbind(data$X/data$TTF,data$Y/data$TTF))
colnames(testing)<-list("X2", "Y2 ")

func<-function(X2,Y2,MPA) 1/(1+exp(-(coeff[1]  + coeff[2]  *X2+ coeff[3] *Y2
+ coeff[4]  *MPA + coeff[5]*X2*X2 + coeff[6]*Y2*Y2 + coeff[7]*X2*Y2)))

estimate<-c(0.5,rep(0,10))   #9#


for(i in 1:10){
estimateX<-func(MPA=(estimate[i]),X2=testing$X2,Y2=testing$Y2)
estimate[i+1]<-mean(estimateX)
}                       #10#

################
#     OUTPUT       #
################

estimate[11]           #11#
head(estimateX)       #12#
```

Change to coeff**M** for application on males from equation 2.3, or use the "coeff**Final**" from the previous script to apply coefficient from equation 2.2. #8#

Function of the equation.

"estimate" will save the successive estimation of the $MPA_{testing}$. One can display it to verify that a plateau is reached. Starting at MPA=0.5 is arbitrary and has no consequence. #9#

Successive estimations of each value of the sample considering the last estimated $MPA_{testing}$. #10#

"estimate[11]" is the final estimation of the Mean Percentage of Activity (MPA) of your dataset #11#

"estimateX" contain the final estimations of each datapoints #12#

Appendix 2.5 Figures for female datasets

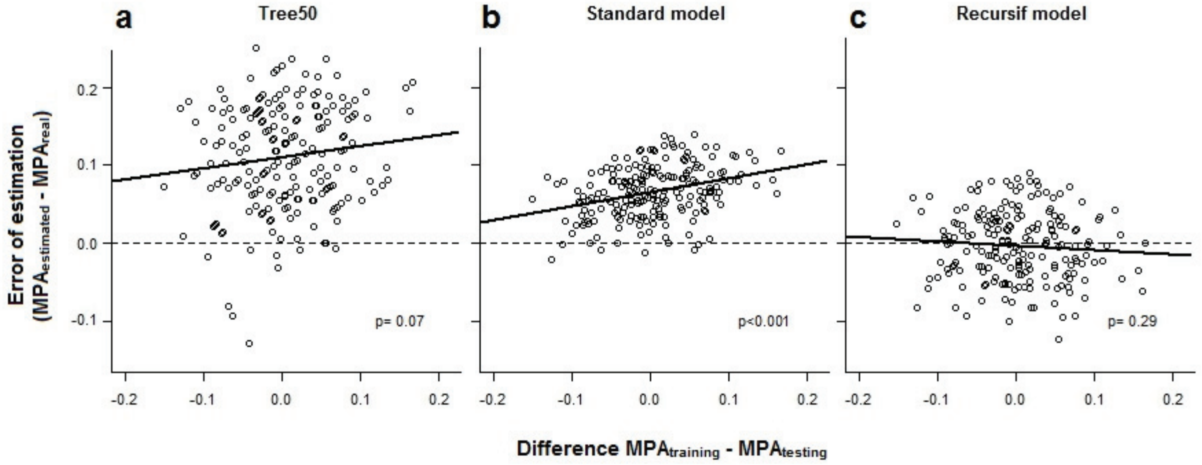


Figure 2.6 Comparison of the estimation of the mean percentage of activity of the Group A testing dataset ($MPA_{testing}$) from each method for the female dataset: (a) the Tree50 method, (b) the standard model, (c) the recursive model. A bias dependent of the difference between the MPA of the training and testing dataset (positive values mean $MPA_{training} > MPA_{tested}$) is observed if the slope of the regression is different to 0 (the p value is provided on the figures). A bias independent of this difference is observed if the regression line (continuous) does not intersect the origin, i.e. at [0,0]

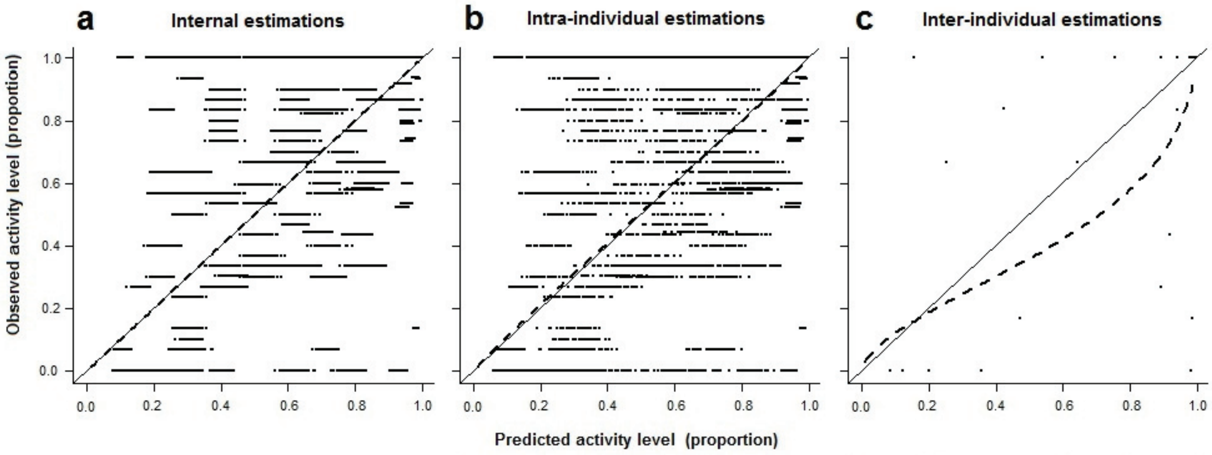


Figure 2.7 Calibration plots for (a) the internal estimations (on the Group A training datasets), (b) the intra individual estimations (on the Group A testing datasets) and (c) the inter-individual estimations (on the Group B dataset from the final equations) for the female dataset. We pooled together the 200 sub-datasets for the panels (a) and (b). Continuous lines represent a perfect calibration with no bias or spread. Dotted lines represent the fitted relationship between the observed values and the logit of their estimation by a logistic regression (see Pearce and Ferrier 2000)

Appendix 2.6 Boxplot of the precisions of the method's MPA estimations

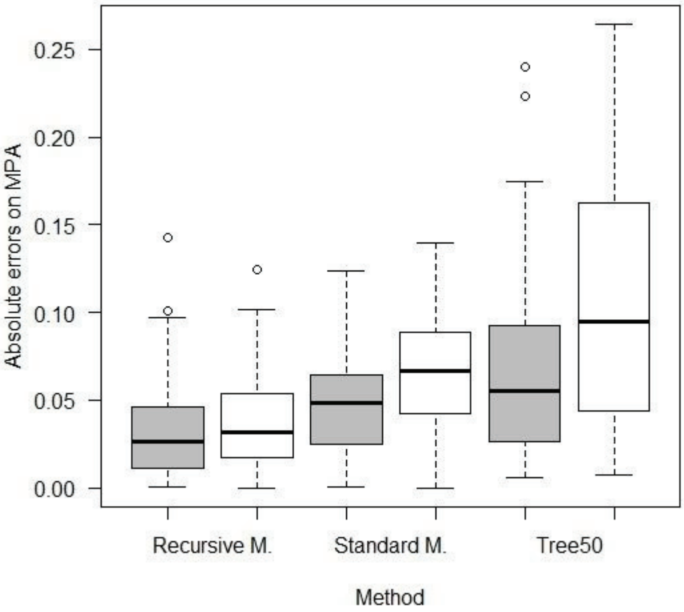


Figure 2.8 Precision (i.e. absolute error) of the Recursive model, the Standard model, and the Tree50 methods on the estimation of the $\text{MPA}_{\text{testing}}$ for male (grey) and female (white) datasets. Recursive models are statistically more precise than the two other methods, and the standard models are statistically more precise than the Tree50s (all paired t-test, $p < 0.001$), for the both sexes

.

Chapter 3 Fission-fusion group dynamics in reindeer reveal an increase of cohesiveness at the beginning of the peak rut

The following chapter is based on the manuscript: Body G, Weladji RB, Holand Ø, Nieminen M Fission-fusion group dynamics in reindeer reveal an increase of cohesiveness at the beginning of the peak rut. Resubmitted to *Acta Ethologica* on 10/01/2014.

3.1 Abstract

Knowledge of the processes of group formation is important to understand the evolution of animal societies. Ungulates typically aggregate during the breeding season. According to the self-organization theory, proximal processes such as mating tactics should explain variation in average group size through their influences on group dynamics. Here, we tested whether variation of the fusion rate or fission rate led to an increase of average group size at the beginning of the peak rut. We followed the movement of marked animals within an enclosed herd of reindeer *Rangifer tarandus* during two breeding seasons (2009, 2011). We used synchronized GPS collars that fixed the animals' position every hour (2009) and every 15min (2011). Group dynamics occurred in three steps: 1) a continuous aggregation of groups; 2) a continuous departure of single females from these groups, and 3) the aggregation of these solitary females to form new groups. We attributed the increase in average group size mainly to a decrease in number of groups, due to a decrease in the group and individual splitting propensities, rather than to an increase of their merging propensities. A decrease in the splitting propensity at the beginning of the peak rut may be due to males herding females, female mate choice or female harassment avoidance. Further research on fission-fusion group dynamics, should calculate merging and splitting propensities, by controlling for variables such as group size, group density, or habitat characteristics.

3.2 Introduction

Animal societies are generally organized in groups that are often unstable, varying in size and in complexity of social networks (Croft et al. 2008), with multiple level of organisation (Couzin 2006). The complexity of animal societies is captured by the three dimensions of the fission-fusion group dynamics: variation in spatial cohesion, variation of group size, and variation of group composition (Aureli et al. 2008). Knowledge of the functioning of animal societies is important to understand their evolution (Danchin et al. 2008). It also has practical applications, such as the management of disease spread (Proffitt et al. 2012), or the success of conservation programs (Haydon et al. 2008).

Recent studies have examined parameters influencing the variation of group size, such as population density (Beauchamp 2011b), food availability (Bercovitch and Berry 2010), and predation risk (White et al. 2012b). However, complex organisations at higher levels result from simple rules at the individual level (Couzin and Krause 2003), and variation in group size is an emergent property of local interactions (Gerard et al. 2002; Couzin and Krause 2003). Accordingly, it is important to quantify the parameters influencing group dynamics to understand the relation between local conditions and emergent group size.

Most social organisations can be produced by only two parameters influencing the group dynamics (Juanico 2009; Aureli et al. 2012): one representing behaviours related to mechanisms of group formation (fusion), and one representing behaviours related to mechanisms of group disintegration (fission) (Couzin 2006). These two parameters are, however, influenced by local conditions. For example, an individual’s perception of predation risk may enhance fusion processes (King et al. 2012), while its perception of foraging competition may enhance fission processes (Marshall et al. 2012). Variations in individual’s motivation to merge or split, i.e. the merging and splitting propensities, may influence the frequencies of fusion and fission events. External variables such as landscape structure

(Gerard et al. 2002; Pays et al. 2012) or population density would also influence the frequencies of fusion and fission events, through variations of the frequency of event's opportunities. A higher rate of fission compared to fusion will produce groups of small sizes, whereas a higher rate of fusion compared to fission will produce larger and more stable groups (Couzin and Laidre 2009). We can therefore expect the variation of the average group size of a population to reflect the variation of the relative rates of fusion and fission.

Changes in average group size during the mating season are common in ungulates, and those groups are generally described by fission-fusion group dynamics (Gower et al. 2009; Bercovitch and Berry 2012; White et al. 2012a). However, during the breeding season, the average group size typically increases (Aung et al. 2001; Jedrzejewski et al. 2006). Food distribution, social relationship, and information acquisition are the three main factors influencing fission-fusion group dynamics according to Sueur et al. (2011), and are not expected to change during the breeding season. Increasing the average group size has, however, a strong influence on the opportunity for sexual selection (Wade and Shuster 2004). Proximal processes influencing fission-fusion group dynamics during the breeding season are therefore most likely linked to mating behaviours. The theory of self-organisation (Couzin and Krause 2003) predicts that individual behaviours (e.g., mating tactics) should explain population patterns (e.g., average group size). Accordingly, any mating behaviour involved should be associated with a change in the mechanisms of group dynamics (the rate of fusion, or the rate of fission).

In ungulates, males and females draw advantages from forming larger groups for mating, and sex-specific tactics may interact (Bro-Jørgensen 2011). For non-territorial species, either males (interference competition), or females (indirect mate choice; Wiley and Poston 1996) can enhance male-male competition by increasing the fusion rate. Conversely, females may also increase the fusion rate to form larger groups and to decrease the harassment level incurred (dilution effect; Clutton-Brock and McAuliffe 2009). Mating behaviours associated with a decreased fission rate include male herding females, females staying with the chosen mate (direct mate choice; Wiley and Poston 1996), and harassment avoidance. For the latter, females may either stay under the protection of an adult male (Clutton-Brock and McAuliffe 2009), or avoid being solitary (Clutton-Brock et al. 1992).

We hypothesized that mating behaviours influence the pattern of variation in group size, through group dynamics. Because of the complex relations between mating tactics, group dynamics and group size, we used a simplified approach in this study by focusing on a broader scale. Accordingly, we tested whether 1) behaviours related to the fusion processes or 2) behaviours related to the fission processes could explain the variation of the average group size, regardless of the exact mating behaviour involved. An increase of the fusion rate or a decrease of the fission rate should explain the increase of the average group size. We tested these predictions using a reindeer *Rangifer tarandus* herd of about 50 individuals fully equipped with global positioning system (GPS) collars. We compared the average group size to the rates of fusion and fission, and we analyzed their variations throughout the mating season.

3.3 Methods

3.3.1 Study design and definitions

3.3.1.1 Area and study herd

We studied a semi-domestic herd of reindeer in Kutuharju Field Reindeer Research Station in Kaamanen, Finland (69°N, 27°E) during two breeding seasons (September through October in 2009 and 2011). In 2009, the herd was composed of 17 males (from 1.5 to 6.5 years old) and 42 females (from 1.5 to 7.5 years old), and in 2011 the herd was composed of 11 males (from 1.5 to 5.5 years old) and 34 females (from 1.5 to 10.5 years old). In 2009, the animals were released in the Sinioivi enclosure (13.4 km²) on September 16th and rounded up on November 3sd. In 2011, the animals were released on September 8th and rounded up on October 18th. Except for tameness, the behaviour of semi-domestic reindeer and wild reindeer are very similar (Reimers et al. 2012).

3.3.1.2 Study design: the full control experiment

We followed the movement of every individual by the mean of synchronized GPS collars, following the "reality mining" approach described in Krause et al. (2013). We used two sets of Tellus GPS collars (from Followit; URL, http://wildlife.followit.se/) the old and the new collars. The old Tellus collars had been previously used and had some battery issues, likely affecting their rate of positioning success (on average 56.2% ± 12.6 attempts were successful). The new Tellus collars with full batteries had a high rate of positioning success (on average 99.9% ± 0.1 attempts were successful). Position fix attempts were programmed to stop after 90 seconds.

Due to management practices, technology failures, or random events, we were not able to follow all of the individuals during the entire study period. In 2009, 40 out of 42 females and 12 out of 17 males were equipped with a GPS collars. Of those, 16 females wore an old Tellus collar, while 24 females and all 12 males wore a new Tellus collar. Unfortunately three old Tellus collars on females stopped working (September 18th, 23rd and 25th) and two male collars dropped off, probably during fights (September 29th and 30th). In 2011, 33 females out of 34 and 10 males out of 11 were equipped with a new Tellus collar and the last male was equipped in the field on October 1st. Among those, one male collar did not work and two female collars stopped working on October 1st and 16th, respectively. The data recorded from these collars were included in the analyses, even though they were not covering the whole season.

The GPS synchronously attempted to fix the collar positions every hour in 2009 and every 15 minutes in 2011. We removed the first and last 24h of the study to avoid the disturbance of the release and the round-up of the herd into the enclosure. Therefore, we studied 1142 recording times (47.6 days) in 2009 and 3800 recording times (39.6 days) in 2011. At each recording time *t* (time *t* hereafter), we obtained a map of the individuals' position. We estimated the GPS precision to be 42m (95th percentile error, average error = 13.8m ± 22.6; Appendix 3.1).

3.3.1.3 The spatial definition of groups

To identify aggregation, we used a chain rule based on the nearest neighbour distance (Carter et al. 2009; Aureli et al. 2012) stating that two individuals belong to the same group if their inter-individual distance is below the threshold r_{max}, i.e. the intra-group maximal distance estimated below.

To estimate r_{max}, we performed a spatial analysis on the point patterns obtained during our 2011 study period, only using times during which every GPS collar recorded their positions (i.e. with 42 data points). For each time *t*, we calculated the cumulative distribution function of the nearest-neighbour distance (the G function; Ripley 1988) with edge corrected by the Kaplan-Meier method (referred below as the $G_{km}(r)$ function, where r is the distance to the nearest-neighbour in meters; Baddeley and Gill 1997). To do this, we used the "Gest" function from the Spatstat package in R (Baddeley 2010). We calculated for each curve the distance "r*" at which the observed curve was the most different to the upper limit of the 95% confidence envelope (one sided) of the theoretical $G_{km}(r)$ function. This envelope was generated by a Monte-Carlo procedure. We generated 1000 times 42 random points in our enclosure and we calculated their G_{km} functions. We defined the one sided confident interval as follow: the lower limit was equals to 0 for any r; and for each r value, the upper limit was defined by the 50th (5%) highest value of the $G_{km}(r)$ function. We selected "r_{max}" as the 95th upper percentile of the r*. Therefore, r_{max} represents the distance to the nearest neighbour from which 95% of the most aggregated pattern can be estimated.

The data set from 2011 included 2863 recording times (75.3% of the whole dataset). The observed $G_{km}(r)$ functions showed a significant aggregated pattern below 463m (based on the intersection between the average of the Gest function and the upper limit of the 95% confidence intervals of the simulated $G_{km}(r)$ function (figure in Appendix 3.2) and our r_{max} was estimated to be 89m.

3.3.1.4 The definition of groups from their composition

Group dynamics can be sex-dependent (Johnson 1983; Jedrzejewski et al. 2006). Therefore, we separated spatial aggregations into "*male groups*" (0 female, ≥ 1 male) and "*parties*" (≥ 1 female, ≥ 0 male). We only studied the dynamics of parties and only used the female component of parties to define group dynamic events (see group dynamics definition below), because female distribution should match resource distribution, while male distribution should match female distribution (Emlen and Oring 1977). During the mating season, individual behaviour of females is likely to differ greatly from group behaviour as females may individually perform mate sampling tactics (Byers et al. 1994), or be subject to intense male harassment when solitary (Clutton-Brock et al. 1992). Therefore, "*parties*" were classified as "*solitary females*" (1 female, ≥ 0 male) or "*groups*" (≥ 2 females, ≥ 0 male). This way, the departure of a single female from a group (an individual behaviour) was not considered fission of the group (a group event).

3.3.1.5 The definition of group dynamics

The separation of parties in groups and solitary females has consequences on the definition of merging and splitting events which constitute the elements of the group dynamics. The group dynamics as recorded here is spatially implicit, and we did not run spatially explicit models such as the correlated random walk (see an example of this method in Haydon et al. 2008).

We compared the female composition of each group at time *t* with that at time *t-1* to establish the group dynamics. Because groups composition may be similar from one recording time to the next, we described the fission-fusion group dynamics using seven events: "*identical*" (the group was similar to the previous recording time, and its identity remained the same), "*fusion*" (≥ 2 groups merging into one group), "*fission*" (one group splitting in ≥ 2 groups), group "*appearance*" (the aggregation of previously separated females), group "*disappearance*" (the separation of all of the females present in one group), individual "*junction*" (a female joining a group), and individual "*departure*" (a female leaving a group). Junction and departure events can happen in addition to identical, fusion and fission events (see detailed definitions used for computation in Appendix 3.3). Short behaviours may be important for mating; however, they may also be GPS errors (of 100m or more), or may be unrelated to mating. Typically, a female may eat or rest in one place while her group is moving, and she rejoins them later. To eliminate these problems, we decided to disregard the departure events when the female came back to the same group within one hour and the junction events when the female had left the same group in the previous hour. By doing so, the reported frequencies were conservative, as we recorded a lower number of events to decrease the number of errors.

"*Merging*" events were represented by fusion, appearance and junction events, while "*splitting*" events were represented by fission, disappearance and departure events. These definitions imply that, mathematically, the variation of the number of groups is due to the variation of the merging and splitting tendencies at the "*group level*" (fusion vs fission) or at the "*interface level*" (appearance vs disappearance). Similarly, the variations of the number of solitary females are mathematically linked to variations of the merging and splitting tendencies at the "*individual level*" (junction vs departure) or at the interface level.

Appearance and disappearance events were named “interface level” due to their effect on the variation of the number of groups and of the number of solitary females.

3.3.1.6 The validation of the group dynamics record

Missing data may decrease the quality of the group dynamics recorded, for instance by transforming a fission event into a departure or a disappearance event, or by missing some departure or junction events. To validate the quality of the group dynamics data, we compared the observed number of groups to the number of groups expected through the recorded group dynamics (see Appendix 3.4). A similar comparison was performed for solitary females (see Appendix 3.4).

The recorded group dynamics accurately described the variation of the number of groups in 2009 and 2011, and of the number of solitary females in 2011 but not in 2009 (Appendix 3.4). Therefore, we did not analyze the variations of the departure and junction rates, as well as the grouping and splitting propensities for 2009.

3.3.2 Statistical analyses

3.3.2.1 Mean group size, number of parties and rut period

We investigated the change in group size by calculating the mean group size (in number of females from groups with ≥ 2 females) at each time *t*. Then, we evaluated whether these variations were either due to a decrease in the number of groups or a decrease in the number of solitary females. These variations were assessed in respect to the early, peak and late rut periods (Tennenhouse et al. 2012). The peak rut period lasts one week and was estimated by a backdating procedure from birth dates. We subtracted to the median birth date the mean gestation length (221 days, Mysterud et al. 2009) to obtain the peak rut day, and we obtained the peak rut week by adding and subtracting three days (similar to Tennenhouse et al. 2011 and L’Italien et al. 2012).

We recorded 38 births from May 5th to May 19th in 2010 with May 10th as a median birth date. Accordingly, we defined the peak rut in 2009 between September 29th and October 5th. Nine copulations were observed between September 27th and October 7th in 2009. In 2012, we recorded 26 new born from May 5th to May 20th, with May 11th as a median birth date. Accordingly peak rut in 2011 started September 30th and ended October 6th. In 2011, seven successful copulations were observed between September 30th and October 8th.

3.3.2.2 The temporal variation of group dynamics

We described the temporal variation of the mechanisms of the group dynamics by analysing three aspects of the group dynamics: the frequency of each event, the relative importance of splitting and merging events, and the propensities of splitting and merging events. Because of the difference in recording schedule, we analyzed each year separately by fitting general additive models (GAM, Crawley 2007). To obtain curves smoothed in a comparable way for the two years, the smoothing parameters of the GAM were set to be $k = 10$ for 2009 and $k = 8$ for 2011, in agreement with the relative length (in days) of the studied periods.

The number of each event occurring at a given time t was used as the frequencies of fusion, fission, appearance, disappearance, junction and departure, per hour for 2009 and per 15 min for 2011. For graphical purpose, we used the daily averages of these frequencies.

We described the relative importance of the merging and splitting frequencies at each level by plotting the cumulative sum of the variation in the number of parties, delta (Δ), corresponding to difference in number of parties before and after the event (group level: Δ*fission*-Δ*fusion*; interface level: Δ*disappearance*-Δ*appearance*; individual level: Δ*departure*-Δ*junction*). For instance, fission of one group into three groups has a Δ*fission* of two, while appearance of one group from two solitary females has a Δ*appearance* of one.

We assessed the propensity of merging as a unique process. We calculated the proportion of parties which could engage into a merging event (i.e. the merging potential =

the parties present at time *t-1* or created by a splitting event at time *t*), and that actually engaged into it (Eq. 3.1). We included in the model the merging potential as a covariate to control for the increasing probability of merging when the parties' density increases.

Eq. 3.1

$$P(herding)_t = \frac{fusion(Gp)_t + appearance(F)_t + junction_t}{Groups_{t-1} + Solitary\,females_{t-1} + \ departure_t + \Delta fission_t + \Delta disappearance_t}$$

with *(Gp)* and *(F)* representing the number of groups and females respectively engaged into the corresponding event; and Δ the net variation of parties induced by the corresponding event.

Contrary to the merging propensity of individuals, the splitting propensity cannot be understood as a unique process, and we therefore separated the leaving propensity, an individual decision, from the group splitting propensity representing group behaviour. We calculated the leaving propensity (probability that a female would leave a group) by dividing the number of departures by the number of females that could leave (i.e. the leaving potential = every female in groups except the two first of each group, Eq. 3.2). We included the typical group size (i.e. the group size in which an average female is, Jarman 1974, Barrette 1991) as a covariate to account for the fact that individuals are less likely to leave larger than smaller groups.

Eq. 3.2

$$P(leaving)_t = \frac{departure_t}{\sum_i (Group\,size_i - 2)_{t-1}}$$

with $Group\,size_i$ the number of females present in the group *i* at time *t-1*.

We calculated the group splitting propensity (the probability that a group would split) by dividing the number of groups that split (by fission or disappearance events) by the number of groups (i.e. the splitting potential, Eq. 3.3). We included the mean group size as a covariate because larger groups are more likely to split apart than smaller ones. The GAMs corresponding to merging, leaving and splitting potentials had a binomial link function and data were weighted by their merging, leaving and splitting potential values, respectively.

Eq. 3.3

$$P(splitting)_t = \frac{fission(Gp)_t + disappearance(Gp)_t}{Groups_{t-1}}$$

with *(Gp)* representing the number of groups and females engaged into the corresponding event

3.4 Results

3.4.1 The average group size and parties number

In 2009 and 2011 we observed an increase of the average group size before the peak rut and a decrease after the peak rut (Fig 3.1a). The increase of the average group size occurred within a relatively short period of time (about 10 days in 2009 and about seven days in 2011) right before the beginning of the peak rut and it decreased quickly after the peak rut.

Both the number of groups (Fig 3.1b) and the number of solitary females (Fig 3.1c) decreased before the peak rut. However, for both years, the decrease of the number of solitary females seemed to happen before the decrease of the number of groups. The number of groups increased more after the peak rut than the numbers of solitary females, and showed a strong negative correlation with the average group size in both years (Pearson correlation: r = -0.87 and r = -0.86 in 2009 and 2011 respectively) whereas the negative correlation was weaker with the number of solitary females (Pearson correlation: r = -0.41 and r = -0.40 in 2009 and 2011 respectively). The number, the dynamics and the spatial movements of parties are displayed for 2009 (Appendix 3.5) and 2011 (Appendix 3.6).

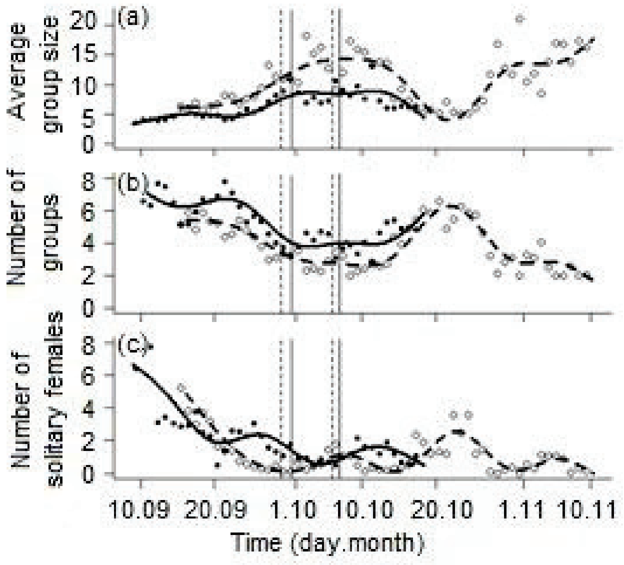


Figure 3.1 Temporal variations of (a) the average group size (in number of females), (b) the number of groups and (c) the number of solitary females for 2009 (open circles, dashed lines) and 2011 (solid circles, continuous lines). Points represent the observed daily averaged values; the lines represent the predictions from the GAM (k = 10 for 2009, k = 8 for 2011) fitted from the original data. Peak rut is represented by the period within the vertical lines (dashed lines for 2009, continuous lines for 2011)

3.4.2 The temporal variation of group dynamics

For both years, the temporal variations of the frequency of fusion matched those of the fission (Fig 3.2a,b), as both the frequencies of appearance and disappearance events (Fig 3.2c,d), as well as the frequencies of junction and departure (Fig 3.2e,f), matched each other. The frequencies of the fusion and fission events decreased until the beginning of the peak rut, when it started to increase for both years (Fig 3.2a,b), before decreasing again after October 20th in 2009 (Fig 3.2b) and after October 14th in 2011 (Fig 3.2a). In 2009, the temporal variation of the frequencies of the appearance, disappearance (Fig 3.2d), departure and junction events (Fig 3.2f) were close to the temporal variation of the fusion and fission events (Fig 3.2b). In 2011, the temporal patterns of the frequencies of the appearance, disappearance (Fig 3.2c), junction and departure (Fig 3.2e) were continuously decreasing until the beginning of the peak rut but were only slightly increasing after the peak rut.

For both years, parties were aggregating as a result of both fission-fusion processes (n = 334 fusions versus n = 322 fissions in 2009, and n = 484 fusions versus n = 462 fissions in 2011, Fig 3.3a,b) and appearance-disappearance processes (n = 118 appearances versus n = 107 disappearances in 2009, and n = 121 appearance versus n = 107 disappearance in 2011,

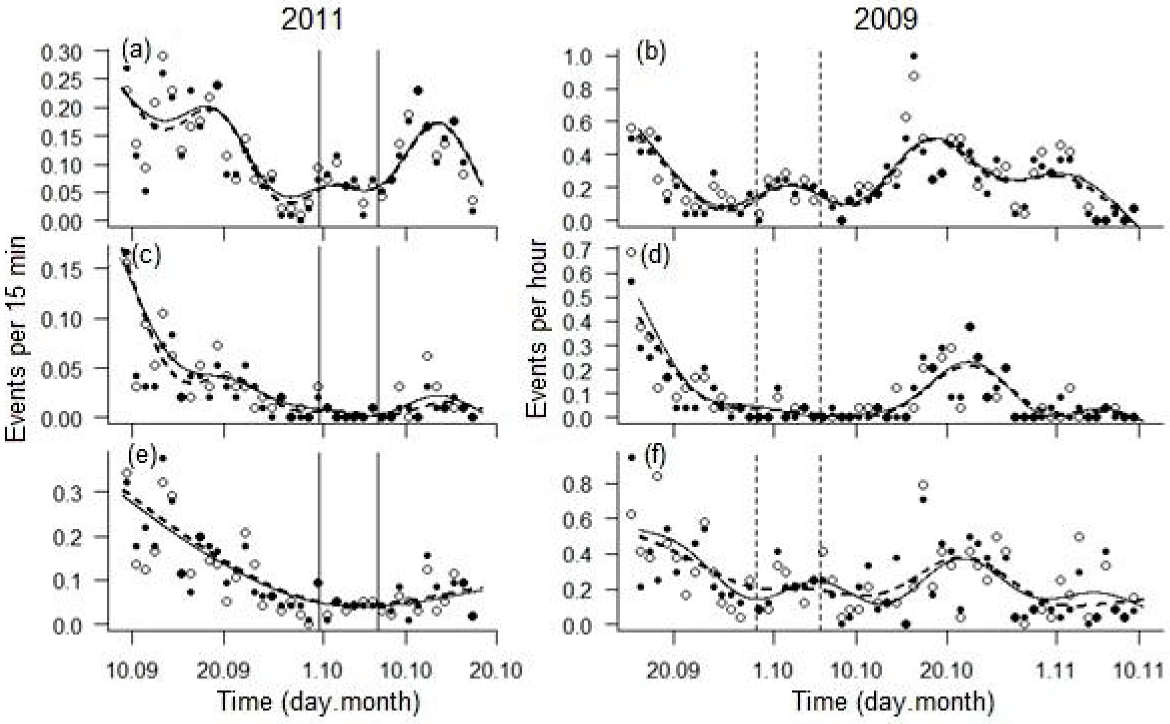


Figure 3.2 Temporal variation of (a, b) the frequencies of fusion (solid circles, continuous lines) and fission (open circles, dashed lines), (c, d) the frequencies of spontaneous group appearance (solid circles, continuous lines) and group disappearance (open circles, dashed lines), and (e, f) the frequencies of junction (solid circles, continuous lines) and departure (open circles, dashed lines) for 2009 (right panels: b, d, f) and 2011 (left panels: a, c, e). Circles represent the observed daily average of the frequency of the event per 15 min (2011) or per hour (2009); the lines represent the predictions from the GAM (k = 10 for 2009, k = 8 for 2011) fitted from the original data. Peak rut is represented by the period within the vertical lines

Fig 3.3a,b). In contrast, from figure 3.3a, it appeared that in 2011 females left groups more often than they joined groups (n = 342 junctions versus n = 358 departures, Fig 3.3a).

In 2011, the temporal patterns obtained for the group splitting propensity (Fig 3.4a) matched the temporal variation of the number of groups (Fig 3.1b) and the leaving propensity temporal pattern (Fig 3.4b) matched the temporal variation of the number of solitary females (Fig 3.1c). The group splitting propensity decreased until the beginning of the peak rut then increased again, while the leaving propensity decreased continuously through time until October 1st. Conversely, there was no recognizable temporal pattern for the merging propensity of individuals (Fig 3.4c).

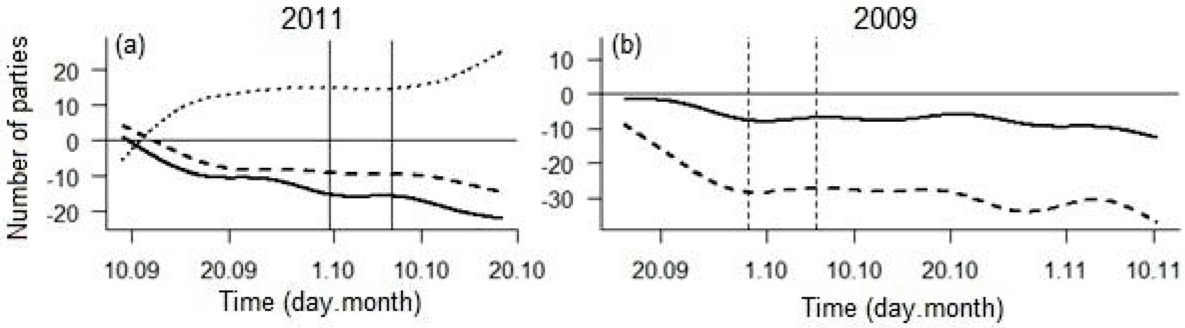


Figure 3.3 Predicted variations of the number of parties due to difference between the fusion and fission events (continuous line), between the spontaneous group disappearance and the spontaneous group appearance events (dashed line) and between the departure and junction events (dotted line), for (a) 2011 and (b) 2009. The difference between junction-departure is not plotted for 2009 for reason explained in the text. In both year we are losing parties by the fission-fusion and the group appearance-disappearance processes (i.e. groups and solitary females are merging together), whereas we are gaining parties as there is more individual females leaving groups than joining groups. Peak rut is represented by the period within the vertical lines

In addition to the temporal variations, the mean group size increased the probability of group to split (slope ± SE = 0.16 ± 0.02, $p < 0.001$) whereas the typical group size decreased the probability of females to leave groups (slope = -0.22 ± 0.09, $p = 0.02$). The current number of parties (i.e. the merging potential) increased the probability that a party engaged into a merging event (slope = 0.14 ± 0.01, $p < 0.001$).

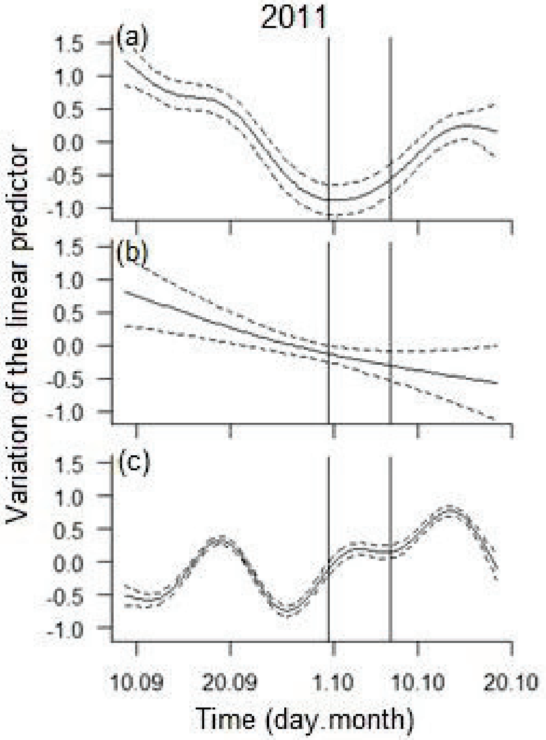


Figure 3.4 Temporal variation of (a) the group splitting tendency, (b) the individual leaving tendency and (c) the herding tendency for 2011. We only display the temporal variation of the linear predictor from the fitted GAM (k = 8) and not the effect of covariates. Dashed lines represent the lower and upper limits of the 95% confident intervals. Peak rut is represented by the period within the vertical lines

3.5 Discussion

3.5.1 The fission-fusion group dynamics

This study revealed a high instability of reindeer groups during the rut as shown by the observed high number of merging and splitting events. Given that females bonds are weak, at least as far as kinship is concerned (Djaković et al. 2012), we classify this reindeer herd as experiencing high fission-fusion group dynamics, corresponding to group C as per Aureli et al. (2008)'s classification. Our assessment point to the group dynamics being described in three simultaneous steps (Fig 3.5): first groups merge, then females leave groups, and finally solitary females form new groups by joining together. This pattern appeared to be consistent throughout the study period, and being independent of the frequencies of events. However, evaluating the relative frequencies of merging and splitting events has proven difficult, as their relative frequencies are different by less than 10% in our study. The relative frequencies of departure and junction is even more difficult to measure, and has only been reported by one study, to our knowledge (Haydon et al. 2008). Indeed, it is not possible to differentiate group fission from individual departure when following a sample of individuals (e.g. Fortin et al. 2009). Yet, it is an important component of the observed group dynamics. Indeed, our

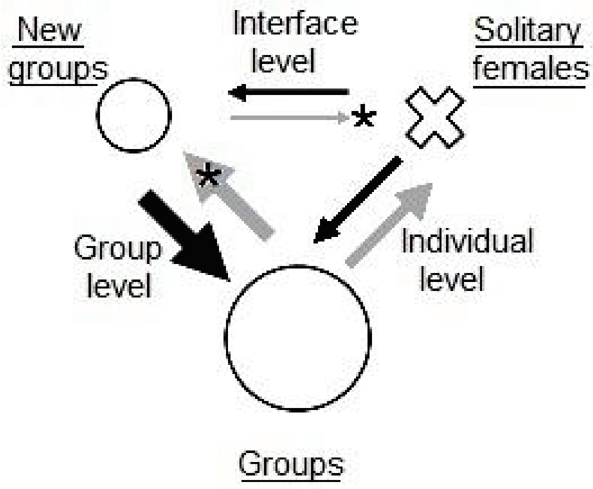


Figure 3.5 Schematic representation of the group dynamics. The cross represents solitary females whereas circles represent groups, usually small when they appear or disappear and larger after merging together. Merging events (i.e. fusion, appearance and junction) are represented by black arrows whereas splitting events (i.e. fission, disappearance and departure) are represented by grey arrows. Arrow sizes are correlated to their frequencies (based on 2011 dataset). Note that counter clockwise arrows (external) are always larger than the clockwise arrows (internal), therefore giving a net direction to the group dynamics (i.e. counter clockwise). Yet, it is the decrease of the propensity of group splitting (arrows with a star) that induced change in the variation of the number of groups

finding of the departure rate being larger than the junction rate happened to be the only process compensating the tendency for groups and individuals to merge together.

3.5.2 The change in group dynamics

We described in this study two important patterns of the group dynamics during the breeding season. As expected for an ungulate species, we reported an increase of the average group size during the peak rut. We also reported a decrease in the frequency of the various events of the fission-fusion group dynamics until the beginning of the peak rut. The increase of the average group size was due primarily to the decrease of the number of groups, but also to the decrease of the number of solitary females as previously observed in reindeer (Kojola 1986) and in red deer *Cervus elaphus* (Clutton-Brock et al. 1982). From our data, it appeared for both years that the decrease of the number of groups was dramatic, leading to only one large group at the beginning of the peak rut (Appendix 3.5: group dynamics video for October 3rd 02:00 and Appendix 3.6: group dynamics video for October 1st 16:00, local time). The decrease of the number of parties was due to the decrease of the splitting propensity, rather than to an increase of the merging propensity of groups or individuals. In fact, the merging propensity was decreasing right before the peak rut in 2011, which is opposite to the expected trend; and therefore variation of the merging propensity could not explain the variation of the number of groups.

The decrease of the number of solitary females was related to the decrease of the number of groups. The splitting propensity of females decreased, therefore reducing the number of new solitary females. As solitary females were still merging, the appearance of new groups compensated the decrease of fission rate, resulting in a stable number of groups, and a smaller number of solitary females. This scarcity of solitary females' resulted in a reduction of the frequency of group appearance, and the total number of groups therefore decreased to reach its minimum at the beginning of the peak rut. The decrease of the splitting propensity was also responsible of the stabilization of the group dynamics, i.e. a reduced frequency of all events. Indeed, by decreasing the splitting frequency, it decreased the opportunity for reversion events, an important component of the group dynamics (Pays et al. 2007). To follow up on these findings, it is crucial to discriminate which mating behaviour actually decreased the fission propensity.

3.5.3 Theoretical and empirical implications

We identified two aspects of our results that may be generalized to other systems, and that should be tested by further studies. First, we showed that reindeer tended to merge more than split at all but the individual level. Other studies on ungulates reported the same trend at the group level (Pays et al. 2007; Pépin and Gerard 2008). This pattern could be driven by either foraging tactics (e.g. copying behaviours) or anti-predation tactics (e.g. dilution effect), and formal tests that take into account basic patterns of the fission-fusion group dynamics should be performed to disentangle the two hypothesis. Second, we demonstrated that, in our system, the decrease of the splitting propensity had a major influence on group dynamics. This result may reflect a more general rule: variation of the splitting propensity would have a more important effect on group dynamics than the equivalent variation of the merging propensity. Indeed, splitting events can happen within any group at any time, while merging events require spatial proximity of groups.

As predicted, in theory, by Couzin and Krause (2003), we reported in fission-fusion group dynamics a difference between frequency and propensity. Knowledge of such differences can be crucial: as demonstrated for mating behaviour by de Jong et al. (2012), the frequency of an event is the result of the propensity of individuals to perform the event and the opportunity they have to perform it. Accordingly, we are proposing to classify most variables influencing fission-fusion group dynamics in these two categories: internal variables, influencing the individuals' grouping propensity; and external variables, which only influence events' probability, i.e. the opportunity for the event. Grouping propensity can change according to predation risk (White et al. 2012b), mating tactics (this study), or social factors (Sueur et al. 2011b; Fishlock and Lee 2013). External variables include the effect of population density (Caughley 1964; Beauchamp 2011b), food distribution (Borkowski and Furubayashi 1998; Bercovitch and Berry 2010), habitat structure (Gerard et al. 2002; Fortin et al. 2009; Pays et al. 2012), group size (Focardi and Pecchioli 2005; Pépin and Gerard 2008), and parties' density (i.e. the number of parties per unit area; Pépin and Gerard 2008).

In our study, short term variations of the average group size, and consequently of the group dynamics, might be explained by external variables. We showed that the fission probability of groups increased with group size, expressing the likelihood that individuals in the group moved in divergent directions. In contrast, the individual departure probability decreased with group size, reflecting either the individual interest to remain in a large group or the individual's increased facility to keep contact with group members in a large group. The fusion probability increased with the parties' density, in accordance with the higher encounter rate when the parties' density is higher. These relations accord with previous studies (Pépin and Gerard 2008; Haydon et al. 2008). We could not take into account some of the external variables (e.g. landscape structure and food distribution), but these should be investigated.

3.6 Appendices

Appendix 3.1 GPS precision

We assessed the precision of our GPS in March 2011. We placed 34 new Tellus GPS collars in one location, and they recorded positions every 10 minutes for 57 hours, for a total of 12966 recordings. We then calculated the average position (average latitude and longitude, using UTM 35N projection in meters) using the records from every collars. The position error was calculated from the Euclidian distance of each record to this average position (Fig 3.6). Errors greater than 1000m appeared to be easily detected, being outside the enclosure, and the corresponding records were removed from the data used in this paper. We derived two error measures corresponding to the distribution of these errors (Fig 3.7): the 95th percentile error and the average error. We found a 95th percentile error of 42m and an average error of 13.8m ± 22.6m. The reported error could be underestimated as the record frequency increase the quality of the GPS position fix, but could also be overestimated as the location used for this experiment had a lower range of visible sky (presence of building next to the location) than the Sinioivi enclosure.

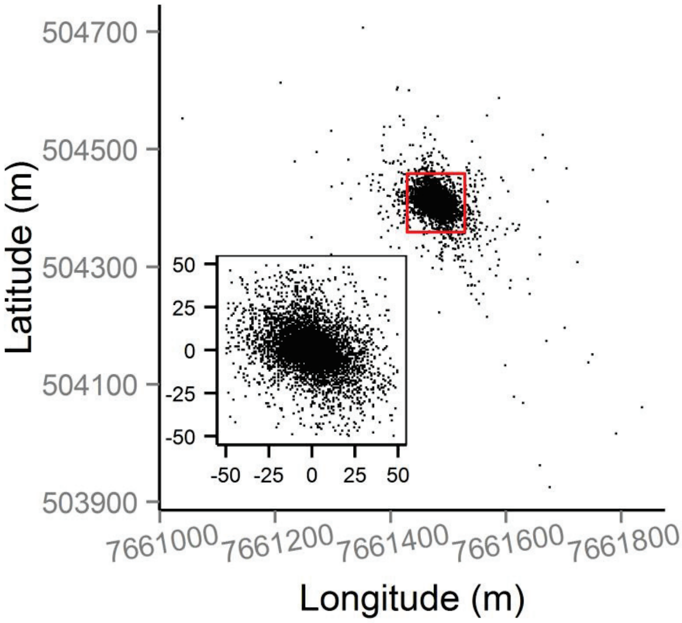


Figure 3.6 Position records according to the UTM longitude and latitude. Bottom left panel is a zoom of the red square and display the coordinates (in meters) of relocations centered to the average relocation

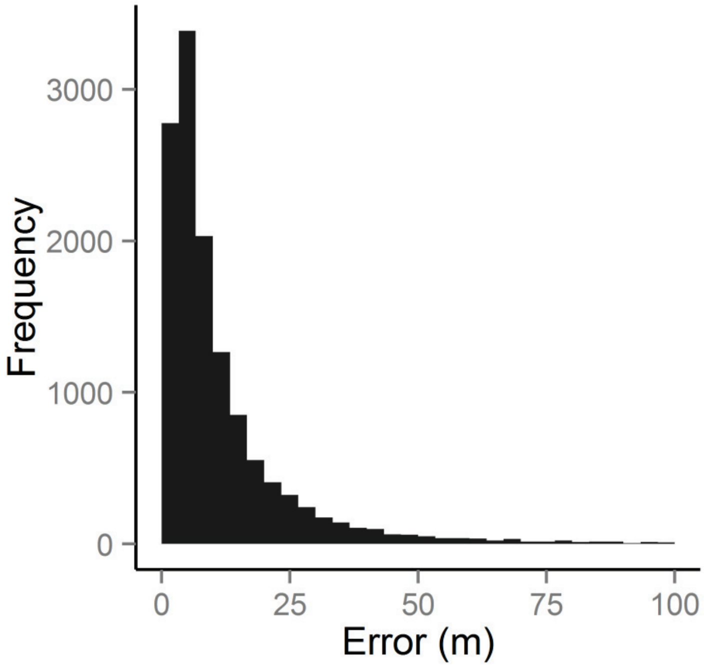


Figure 3.7 Histogram of errors. Errors greater than 100m are not displayed

Appendix 3.2 Simulated and observed nearest-neighbour distance (Fig 3.8)

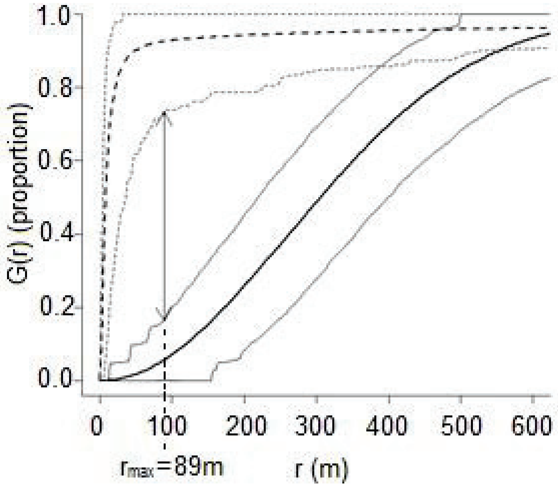


Figure 3.8 Observed (bold dashed line) and simulated (bold continuous line) cumulative distribution function of the nearest-neighbour distance (G function) with their 90% confidence intervals. G(r) represents the proportion of the individuals in the population (y-axis) that has their nearest-neighbour within the distance r (x-axis). We display the difference between the two confident intervals at the estimated intra-group maximal distance (r_{max} = 89m)

Appendix 3.3 Definitions of fission-fusion group dynamics events used for computation

We compared the female composition of each group at time *t* with those at time *t-1* to establish the group dynamics. Any groups at time *t* and *t-1* that had in common at least two females were linked by an event which could be "*identical*", group "*fusion*" or group "*fission*" and events were associated to groups at time *t*. A fusion event (≥ 2 groups merging into one group; "*fusion*" hereafter) occurred when a single group at time *t* was linked with two or more groups at time *t-1*. A fission event (one group splitting in ≥ 2 groups; "*fission*" hereafter) occurred when two or more groups at time *t* had a link with a common group at time *t-1*. A group at time *t* can be linked by several fusion and fission events. When a group at time *t* did not come from any fusion or fission event, the link was "*identical*" and the group identity remained the same. A group at time *t* with no link to any group at time *t-1* was called an appearing group (hereafter "*appearance*" i.e. the aggregation of previously separated females). A group at time *t* with no link with to any group at time *t+1* was called a disappearing group (hereafter "*disappearance*" i.e. the separation of all of the females present in one group).

We described two more events that were associated with solitary females. A "*departure*" event happened when a female in a group A at time *t* was not found in the same group or in any group linked with group A at time *t+1*. She left group A at time *t* with no other female to be solitary, to later join another group, or to form a new group. The reverse situation, a "*junction*" event happened when a female present in a group B at time *t* was not in the same group at time *t-1* or in any group at time *t-1* linked with group B. Three different situations can lead to a female joining group B at time *t*: she was solitary, she came from a disappearing group or she had left another group at time *t-1*. Junction and departure events can happen in addition to fusion and fission events.

Appendix 3.4 The validation of the group dynamics record

Missing data may decrease the quality of the group dynamics recorded, for instance by transforming a fission event into a departure or a disappearance event, or by missing some departure or junction events. To validate the quality of the group dynamics data, we compared the observed number of groups to the number of groups expected through the recorded group dynamics. A similar comparison was performed for solitary females. We followed the expected variation of the number of groups ($\Delta groups_t$) by equation 3.4 and the expected variation of the number of solitary females ($\Delta solitary\ females_t$) by equation 3.5. We compared the cumulative sum of $\Delta groups_t$ to the observed number of groups and the cumulative sum of $\Delta solitary\ females_t$ to the observed number of solitary females by graphical inspection of the accordance of the respective curves and by assessing the strength of the relationship using Pearson correlation. We only performed analyses when the correlation between the two curves was high.

Eq. 3.4

$$\Delta groups_t = \Delta fission_t - \Delta fusion_t + appearance(Gp)_t - disappearance(Gp)_t$$

Eq. 3.5

$$\Delta solitary\ females_t = disappearance(F)_t - appearance(F)_t + departure_t - junction_t$$

Where:

$\Delta groups_t$ and $\Delta solitary\ females_t$ are the variation of the number of groups and the number of solitary females, respectively, expected by the recorded group dynamic;

$\Delta fusion_t$ and $\Delta fission_t$ are the net change of the number of parties between time *t-1* and time *t* due to fusion and fission, respectively;

$appearance(Gp)_t$ and $disappearance(Gp)_t$ are the number of groups engaged in appearance and disappearance between time *t-1* and time *t*, respectively;

$disappearance(F)_t$; $appearance(F)_t$; $departure_t$ and $junction_t$ are the number of females engaged in disappearance, appearance, departure and junction between time *t-1* and time *t*, respectively

Results

The variations of the expected number of groups by the recorded group dynamics matched the observed variations of number of groups for both years (Pearson correlation: r = 0.87 and r = 0.95, for 2009 and 2011 respectively). Visual inspection of the variation of the number of solitary females and its expectation by the recorded group dynamics showed that the two curves matched each other for 2011 (Pearson correlation: r = 0.78), but not for 2009 (Pearson correlation: r = 0.55). Therefore, we did not analyze the variations of the departure and junction rates, as well as the grouping and splitting propensities for 2009.

Appendix 3.5 Video 2009

Appendix 3.6 Video 2011

Chapter 4 High quality reindeer males control female behaviour during the rut

The following chapter is based on the manuscript: Body G, Weladji RB, Holand Ø, Nieminen M High quality reindeer males control female behaviour during the rut. Submitted to *PLoS One* on 14/02/2014.

4.1 Abstract

During the rut, female ungulates move among harems or territories, either to sample mates or to avoid harassment. Females may be herded by a male, may stay with a preferred male, or aggregate near a dominant male to avoid harassment from other males. In fission-fusion group dynamics, female movement is best described by the group's fission probability, instead of inter-harem movement. In this study, we tested whether male herding ability, female mate choice or harassment avoidance influence fission probability. We recorded group dynamics in a herd of reindeer *Rangifer tarandus* equipped with GPS collars with activity sensors. We found no evidence that the harassment level in the group affected fission probability, or that females sought high quality (i.e. high rank and hence successful) males. However, the behaviour of high quality males decreased fission probability. Male herding activity was synchronous with the decrease of fission probability observed during the rut. We concluded that male herding behaviour stabilized groups, thereby increasing average group size and consequently the opportunity for sexual selection.

4.2 Introduction

For reasons that remain unclear (e.g. Stopher et al. 2011), females of polygynous species commonly move among mating groups or territories (Byers et al. 1994; Maher 1997; Naulty et al. 2013). Female ungulates alter their movement patterns during the breeding season (Ozoga and Verme 1975; Richard et al. 2008), often aggregating around the same male or the same place. These changes might be associated with either male or female mating behaviour, likely resulting in increased group size (Clutton-Brock et al. 1982; Aung et al. 2001; Jedrzejewski et al. 2006), which ultimately increases the intensity of sexual selection (Wade and Shuster 2004). Understanding factors influencing female movement is therefore important to identify which mating behaviours drive sexual selection (Bro-Jørgensen 2011).

Males may increase their mating opportunities by stabilizing their harems (Clutton-Brock et al. 1982; Bro-Jørgensen 2011), whereas females may gain indirect benefits by leaving harems to sample mates (Byers et al. 1994; Byers et al. 2005; Naulty et al. 2013). Females may also move to optimize direct benefits by selecting resource-rich territories (Carranza 1995) or by avoiding harassment (Clutton-Brock and McAuliffe 2009; Bro-Jørgensen 2011). Male coercion, female mate choice and harassment avoidance, can individually or concurrently, constrain female movement. The relative importance of these behaviours on female movement has rarely been estimated, despite their potential for enhancing our understanding of the drivers of sexual selection.

In groups with fission-fusion dynamics (Aureli et al. 2008; Sueur et al. 2011b), group sizes are influenced by the relative rates of group splitting and merging events (Couzin and Laidre 2009). Accordingly, males may benefit more from increasing group stability than preventing single females from leaving the group, which is not easy to achieve (Bro-Jørgensen 2011). Avoiding harassment may also increase group stability. Indeed, females may either aggregate to dilute harassment (Carranza and Valencia 1999) or stay under the protection of the harem holder, i.e. the dominant male (Bowyer et al. 2011). Because females often copy or follow each other's movement, a female leaving a group to sample mates may induce fission of the group. Once females have chosen a mate, they would stop sampling, and remain with his group which is less likely to split. Therefore, the influence of male or female mating behaviour on female movement may best be represented in fission-fusion group dynamics by an index of group stability, which should be negatively correlated with the group's fission probability.

Coercion and deception are used by males to prevent individual females from leaving their harems (Bro-Jørgensen 2011). Herding of females, a common behaviour in ungulates, is likely more efficient to decrease the fission probability, and increase male reproductive success, than identifying and following individual females. Although males do not specifically herd females in estrous (Hirotani 1989), male reproductive success has been shown to strongly correlate with their social rank (Røed et al. 2002; Mainguy et al. 2008; Clutton-Brock and McAuliffe 2009), which is positively correlated with the stability of their groups (L'Italien et al. 2012).

Female ungulates are as likely as female birds to choose their mates (Clutton-Brock and McAuliffe 2009), but the way they evaluate phenotypic quality is unclear. A number of criteria has been suggested, including vocalization (Charlton et al. 2007), antler size (Vanpé et al. 2007a), horns size (Ezenwa and Jolles 2008), body size (Røed et al. 2007) or male social rank (Clutton-Brock and McAuliffe 2009). Male social rank is an integrative measure of phenotypic quality and may correlate with the characteristics females evaluate when sampling males (Clutton-Brock and McAuliffe 2009; Bro-Jørgensen 2011). Two strategies, threshold sampling and Bayesian sampling, predict a lower probability to leave a male of higher phenotypic quality (Wiegmann et al. 2013), and consequently, a lower fission probability.

Harassment avoidance is expected to influence the behaviour of female ungulates during the breeding season (Clutton-Brock and McAuliffe 2009; Bro-Jørgensen 2011). Harassment level can be diluted by increasing group size, and by joining a harem controlled by a high quality (i.e. high ranked) male (Carranza and Valencia 1999; Sánchez-Prieto et al. 2004). Solitary females are particularly exposed to harassment (Clutton-Brock et al. 1992), so that females prefer to remain in a group, decreasing the fission probability. Females may also seek the protection of a dominant male (Holand et al. 2006; Bowyer et al. 2011) who will chase satellite males away, thereby keeping harassment to a minimum. Satellite males, usually young and low ranked males, are indeed responsible for harassing females, which may occasionally lead to extreme consequences such as death (Réale et al. 1996).

Reindeer *Rangifer tarandus* are highly sexually dimorphic (Geist 1999; Melnycky et al. 2013) and exhibit fission-fusion group dynamics (Hirotani 1989; L'Italien et al. 2012). According to sexual selection theory (Darwin 1871; Danchin et al. 2008), male herding ability (P_1), female mate choice (P_2) or harassment avoidance (P_3) would decrease the group's fission probability. If males successfully herd females (P_1) fission probability should decrease with the time dominant males spend herding or in herding-like activities (Table 4.1). If females choose their mates (P_2), fission probability should decrease when the group is led by

Table 4.1 Predicted relationship trend between the group's fission probability and the dominant male activity level, its quality, the number of satellite males and the activity of satellite males and females in the group. "+" and "-" signs represent the predicted direction of the relation between a variable and the group's fission probability. Among male quality categories, they represent the relative influence of the variable on the group's fission probability

Group's fission probability	Dominant male mating activity level		Dominant male presence			Satellite males		Females
	High quality	Low quality	High quality	No male	Low quality	Number	Mating activity level	Eating frequency
Herding ability (P_1)	- -	-						
Female mate choice (P_2)			-	0	+			
Harassment avoidance (P_3)						-	-	+

*This prediction was also assessed using the temporal synchrony between herding and the group's fission probability a high quality male as compared to female only groups (i.e. groups without males), and should increase if the group is led by a male of low quality (Table 4.1). Finally, we predict that (P_3) the fission probability should decrease with increasing number of satellite males, and with the level of their involvement in mating-related activities. It should also correlate positively with the time female spend feeding, considered to be the time when they are undisturbed, as a decrease in time feeding may result from harassment (Table 4.1).

4.3 Methods

4.3.1 Area and study herd

We studied a semi-domestic herd of reindeer in Kutuharju Field Reindeer Research Station in Kaamanen, Finland (69°N, 27°E) during the 2011 breeding season (September 8th -October 18th). The herd, composed of 11 males (from 1.5 to 5.5 years old) and 34 females (from 1.5 to 10.5 years old), was released into the Sinioivi enclosure (13.4 km²). We removed from the analysis the first and the last 24 h to avoid the influence of the herd release and roundup. Ten males and 33 females were originally equipped with a Global Positioning System (GPS) Tellus medium collar and the last male was equipped with a GPS collar in the field on October 1st. During the season, one male collar (ranked 4) did not work and two female collars stopped working on October 1st and October 16th, respectively. All GPS collars synchronously recorded their position every 15 minutes, for a total of 3800 recordings. At each recording time *t*, we generated a map of individual positions.

4.3.2 Ethics statement

Handling of animals and data collection was done in agreement with the Animal Ethics and Care certificate provided by Concordia University (AREC-2010-WELA and AREC-2011-WELA) and by the Finnish National Advisory Board on Research Ethics.

4.3.3 Group definition

We defined groups from the spatial aggregation of individuals. We used a chain rule based on the nearest neighbour distance (Carter et al. 2009; Aureli et al. 2012) stating that two neighbours belong to the same group if their inter-individual distance was below 89 m (Chapter 3). Then, we followed each group (≥ 2 females and ≥ 0 male) until it disappeared. A group could disappear if it split (fission) or merged with another one (fusion). Male and individual female departures and junctions from the group or to the group did not influence the group identity. To prevent registering excessive splitting events due to GPS errors or GPS location failures (i.e. missing data), we applied a smoothing procedure to the group identity. Any reversion, i.e. a group splitting followed by the sub-groups merging together (Pays et al. 2007), which lasted less than 30 min was disregarded and the same group identity was subsequently used. Because small groups appeared particularly sensitive to GPS errors, we increased this time up to 60 min for groups containing only two females. For descriptive purpose, we also assessed the number of groups present in the enclosure every 25 hours (to insure data independence) as well as their individual duration. We report the average number of group and their half-life (i.e. the median group duration) according to the period of rut and the quality of the dominant male (see definitions below).

4.3.4 Survival analysis

4.3.4.1 Model

We ran a non-parametric survival analysis model (a Cox model with the coxph function using the package “survival” in R, Crawley 2007) with the duration of the group as index of survivorship (for similar analysis, see Fortin et al. 2009). As we were interested in the group’s fission probability, we recorded splitting events as death events, whereas merging events were recorded as censoring events. Indeed, the group had not split when the fusion happened, but it cannot be followed further as its composition dramatically changed.

4.3.4.2 Explanatory variables

We included the following variables in the full model according to our predictions (P_1, P_2, P_3): the quality (see below) of the dominant male (*Male*; P_1, P_2) and the proportion of time it spent in mating-related activities (*DomAct*; P_1); the number of satellite males in the group (*NbSat*; P_3), and the proportion of time they spent in mating-related activities (*SatAct*; P_3); and the percentage of time females spent feeding (*FemEat*; P_3). We also included two covariables: the group size (*GpSize*), as larger groups are expected to split more easily (Pays et al. 2007), and the period of the rut (*Period*) (see below) as preliminary analyses revealed temporal variability of group dynamics. We had, unfortunately, no data to control for the possible influence of habitat structure (Gerard et al. 2002; Fortin et al. 2009; Pays et al. 2012). However, it is unlikely that habitat selection varied during the breeding season in a way that would influence the reported results.

We classified males (*Male*) as "high quality males" or "low quality males" based on their social rank (used as indicator character *sensu* Wiegmann et al. 2013). We established a linear hierarchy among males from field observations of agonistic behaviours. Because male ranked 4 was not followed by GPS, the top three males were classified as "high quality" (i.e. high rank) and the remaining eight males as "low quality" (i.e. low rank). This threshold is based on field observations as the three top ranked males were most often seen holding a harem. Moreover, this classification enhanced statistical power (as some "low quality" males were still able to lead medium size groups) and was related to the body mass and antler size: high quality males weighted more than 125 kg and their antlers measured more than 85 cm, while low quality males were lighter than 115 kg and their antlers were smaller than 85cm. The variable *Male* included a third class ("without male") for female-only groups.

The breeding season was divided in two periods (*Period*). The rutting period was defined as the peak rut week and the early peak rut week (Tennenhouse et al. 2011) for a total period of two weeks (September 23rd to October 6th), when mating behaviours were more frequent. The time before and the time after the rutting period, were considered as "outside rut". Groups were ascribed to a given period based on the average date of the group (Eq. 4.1).

Eq. 4.1 $$Date_{average} = (Date_{group\ ends} - Date_{group\ starts})/2$$

We determined the median group composition from GPS records. *GpSize* was consequently the median number of females in the group and *NbSat* the median number of males in the group, excluding the dominant male. We assumed that the male with the highest social rank in the group was the dominant male. When the dominant male changed during the duration of the group, we removed the group from analysis.

The percentage of time males or females spent in a given activity was estimated from the activity sensor records using the recursive model (Body et al. 2012) at each recording *t* (see Fig 4.4 in Appendix 4.1 for details). Once resting periods were detected (Fig 4.5, Fig 4.6 in Appendix 4.1), we estimated during the active periods the percentage of time males spent in mating-related activities (Fig 4.4 in Appendix 4.1), i.e. standing, walking or running which represent short behaviours such as chasing males, herding females, threatening, grunting, courting, seeking copulation and being vigilant toward other males (Tennenhouse et al. 2012). The average percentage of time the dominant male spent in mating-related activities in the group formed the *DomAct* variable. We used the average of the cumulative percentage of time satellite males spent in mating-related activities to form the *SatAct* variable. Similarly, we estimated the average percentage of time active females spent feeding in the group at each recording *t* (Fig 4.4 in Appendix 4.1), and we averaged these values throughout the duration of the group to form the variable *FemEat*.

4.3.4.3 Model selection

The most complex model fitted to explain the group's fission probability included *Period*, *Male*, *GpSize*, *NbSat*, *DomAct*, *SatAct* and *FemEat*, and a number of interactions among those variables. We included a two-way interaction *Male:Period* to take into account the fact male quality influences the timing of their mating behaviour (Tennenhouse et al. 2012). Within these different periods, the quality of the male can also influence the efficiency of his mating behaviours or his ability to manage a larger group. Consequently, we included the three-way interactions *Male:Period:DomAct*, and *Male:Period:GpSize*. We performed all possible subsets of models (Symonds and Moussalli 2010) and extracted the Akaike Information Criteria (AIC) from each (Crawley 2007). The number of different possible models, 488, was lower than the sample size (see results) as preconized (Burnham et al. 2011). We calculated AIC weights for each variable from all subsets (Symonds and Moussalli 2010), but we only displayed models with a ΔAIC ≤ 2. Then, we selected, among these models, the one including the variables with the highest relative importance (obtained by summing AIC weights, Burnham and Anderson 2002) for both graphical purpose and effect sizes, which dealt with model uncertainty (Symonds and Moussalli 2010).

4.3.5 Temporal Synchrony

4.3.5.1 Temporal variation in herding frequency

To obtain a more precise measure of herding, we used a long-term dataset (15 years, from 1996 to 2011, without 1998) of direct observations of dominant male behaviour during the rut season to assess the synchrony between herding and group's fission probability. Behavioural records were collected using a 15 min focal observation method (Tennenhouse et al. 2012). As herding a female regularly switched to a chase (Espmark 1964), we summed behaviours classified in the field as either "herd" or "chase female" to assess the frequency of the herding behaviour. We modelled the proportion of time spent herding as a function of the number of days to the beginning of the peak rut using a generalized additive model (GAM), with a smoothing parameter k of 8. The beginning of the peak rut was calculated for each year by the back-dating procedure (as described above for the variable *Period*) and all years were then pooled together.

4.3.5.2 Temporal variation of the group's fission probability

Using the above GPS dataset recorded in 2011, we calculated the group's fission probability at each recording time t as the proportion of group at time t_{-1} that split at time t (Chapter 3). We analyzed the temporal variability of the fission probability using a GAM with the time as explanatory variable, with a smoothing parameter k of 8. We included the mean group size as covariate. The GAM had a binomial link and data were weighted by the number of groups at time t_{-1}.

4.4 Results

4.4.1 Number of groups and group half-life

Outside the rut period, there were on average (± SE) 1.5 ± 0.3 groups without males, 1.0 ± 0.3 groups with a low quality male, and 1.0 ± 0.2 groups with a high quality male (Fig 4.1a). The median duration of these groups were respectively 15.5 ± 4.3 hours, 7.4 ± 1.7 h and 9.7 ± 3.1 h (Fig 4.1b). At any time during the rut, there were 0.5 ± 0.1 groups without males, 0.9 ± 0.2 with a low quality male and 1.7 ± 0.2 with a high quality male (Fig 4.1a). These groups lasted on average 47.6 ± 12.2 h, 27.0 ± 5.9 h, and 33.4 ± 8.7 h, respectively (Fig 4.1b).

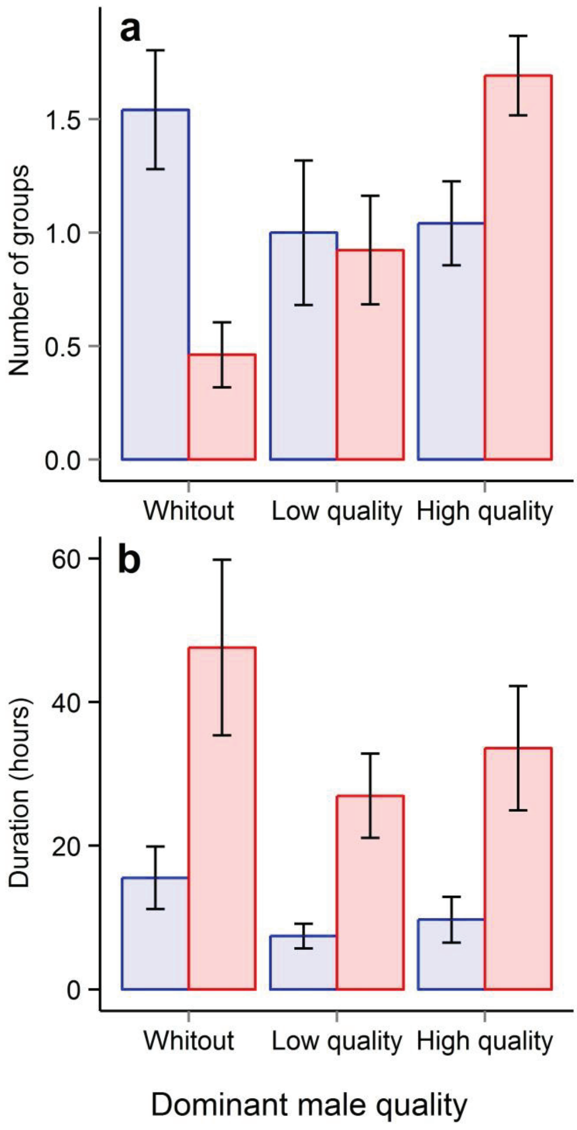


Figure 4.1 Number of groups (a) and their half-life (b) according to the quality of the dominant male and the period of the rut. Averages are represented in each category with their standard errors. Left-blue bars and right-red bars correspond to the outside rut and during rut periods, respectively.

4.4.2 Survival analysis

We analyzed 1075 groups which included 335 splitting events. Among these groups, 879 were recorded outside the rut period (N = 300, 276 and 303 without males, with low quality and high quality males, respectively), whereas 196 were recorded during the rut period (N = 42, 60 and 94 without males, with low quality and high quality male, respectively). Model certainty to explain the group's fission probability was low, as it took 166 models to reach 0.95 of the AIC weights. Twelve models had ΔAIC ≤ 2 (Table 4.2) and they represented together 0.31 of the AIC weights. Confidence in variable selection was high (Table 4.2), as

Table 4.2 Selection of the model explaining variations of the group's fission probability. We represented the variables included in the 12 best models that have ΔAIC ≤ 2, with their respective AIC values, ΔAIC, their AIC weights (AIC_w) and the cumulative sum of the AIC weights (acc AIC_w). We also present the cumulative sum of the AIC_w in which each variable is presented, giving the variable's AIC weight (w_i ; in line)

Candidate models								Interaction between *Male* and:			Interaction between *GpSize* and:						
variable	Period	GpSize	Male	DomAct	NbSat	SatAct	FemEat	*:Period*	*:DomAct*	*:both*[1]	*:Male*	*:Period*	*:both*[2]	**AIC**	**ΔAIC**	**AIC_w**	**acc AIC_w**
w_i	0.99	~1.00	~1.00	0.96	0.34	0.45	0.33	0.65	0.84	0.07	0.81	0.46	0.15				
1	×	×	×	×				×	×		×			3930.9	0	0.044	0.044
2	×	×	×	×		×		×	×		×			3931.3	+0.41	0.036	0.081
3	×	×	×	×				×	×		×	×		3931.4	+0.54	0.034	0.115
4	×	×	×	×					×		×			3931.4	+0.56	0.034	0.149
5	×	×	×	×		×			×		×			3932.0	+1.12	0.025	0.174
6	×	×	×	×		×		×	×		×	×		3932.1	+1.20	0.024	0.199
7	×	×	×	×	×	×		×	×		×			3932.3	+1.37	0.022	0.221
8	×	×	×	×			×	×	×		×			3932.3	+1.42	0.022	0.243
9	×	×	×	×	×			×	×		×			3932.6	+1.73	0.019	0.262
10	×	×	×	×		×	×	×	×					3932.7	+1.83	0.018	0.280
11	×	×	×	×				×	×					3932.8	+1.93	0.017	0.297
12	×	×	×	×	×	×		×	×		×	×		3932.9	+1.97	0.017	0.313

[1] interaction: *Male:Period:DomAct*

[2] interaction: *Male:Period:Gpsize*

the variables *Period*, *Male*, *DomAct*, and *GpSize* had AIC weights over 0.95, while *SatAct*, *NbSat*, and *FemEat* had AIC weights ≤ 0.45. The three interactions formed with the variable *Male* (*Male:Period*; *Male:DomAct*; *Male:GpSize*) had high AIC weights (≥ 0.64, Table 4.2), while the other interactions had AIC weights ≤ 0.46 (Table 4.2). The model 1 (i.e. with the lowest AIC) in Table 4.2 was the combination of the two most parsimonious models (models 4 and 11, Table 4.2), and included all the variables with high AIC weights, in contrast to models 4 and 11. Therefore, model 1 was the best model to represent AIC weights of the different variables, and it was used for interpretation hereafter.

The variables related to harassment avoidance, i.e. *NbSat*, *SatAct* and *FemEat*, did not influence the group's fission probability, which was independent of group size when males were absent (Table 4.3, Fig 4.2a), and increased with group size when the dominant male was of low (Table 4.3, Fig 2b) or high (Table 4.3, Fig 4.2c) quality. The fission probability was lower in absence of males (Fig 4.2a), than in their presence, regardless of their quality (Fig 4.2b,c). The mating-related activities of low quality dominant males did not influence the fission probability (Table 4.3, Fig 4.2d). Conversely, the proportion of time high quality dominant males spent in mating-related activities decreased the fission probability (Table 4.3, Fig 4.2e). As expected, the fission probability decreased during the rut period, especially for

Table 4.3 Parameter estimates and corresponding standard error (SE) of the final model explaining the splitting probability of (a) groups without male, (b) groups controlled by a low quality dominant male, and (c) groups controlled by a high quality dominant male

	(a) No male		(b) Low quality male		(c) High quality male	
	Estimates ± SE	P value	Estimates ± SE	P value	Estimates ± SE	P value
Intercept	0		0.03±0.39	p = 0.930	0.55±0.36	p = 0.130
Period	0.26±0.36	p = 0.48	-0.30±0.27	p = 0.270	-0.67±0.21	p = 0.002
Group size	-0.03±0.07	p = 0.69	0.11±0.02	p < 0.001	0.06±0.01	p < 0.001
Dominant male sexual activity			-0.10±0.35	p = 0.780	-1.46±0.50	p = 0.003

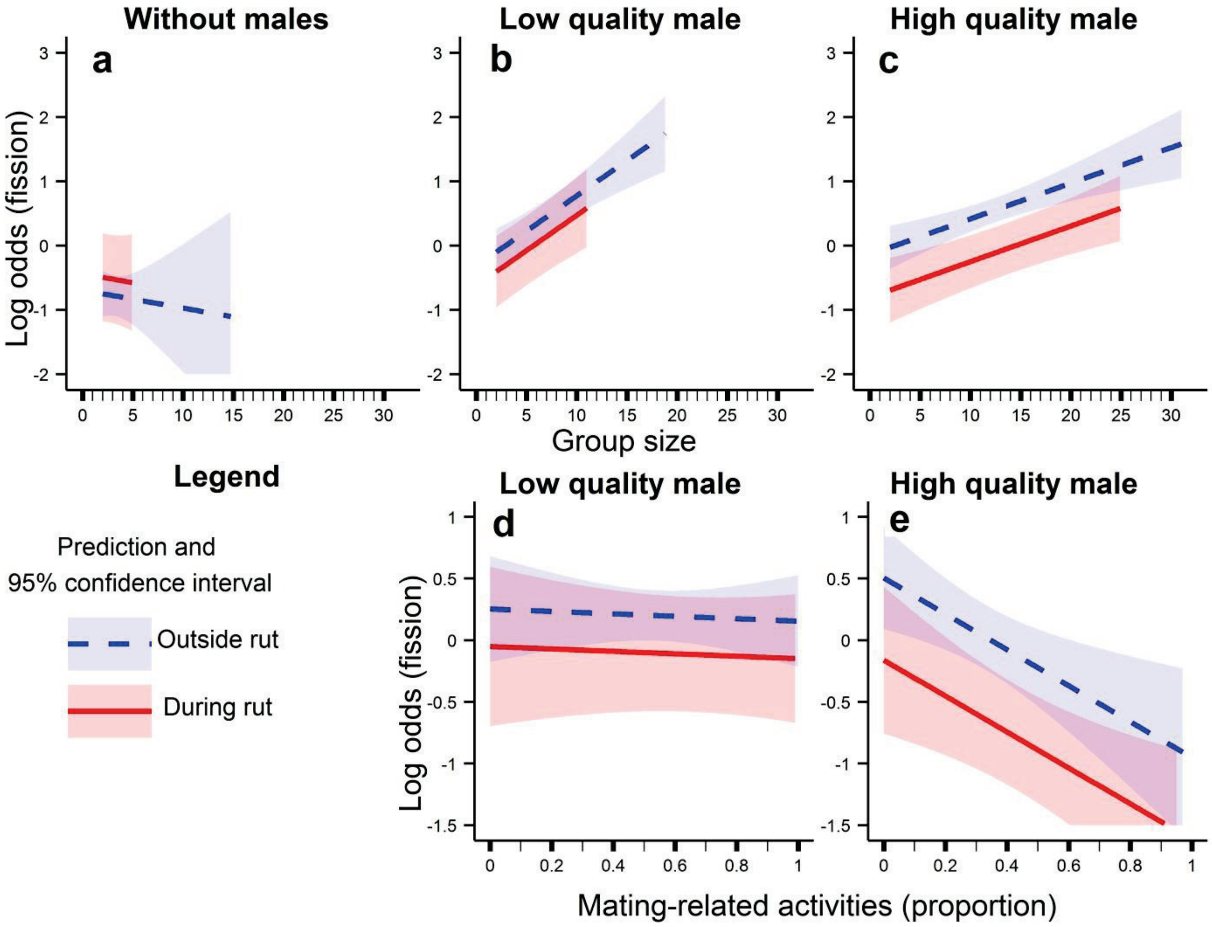


Figure 4.2 Social and behavioral influence on group's fission probability. Partial effect on group's fission probability of the group size (a, b, c) and of the proportion of time the dominant males spent in mating-related activities (d, e) according to the quality of the dominant male of the group : without males (a), low quality male (b, d), high quality male (c, e), and according to the period of the rut: outside the rut (continuous and blue lines) and during the rut period (dashed and red lines). Effects are presented with their 95% confidence intervals

high quality males (Table 4.3, Fig 4.2b vs. Fig 4.2c, Fig 4.2d vs. Fig 4.2e). The model explained about 8% of the variability in the group's fission probability (R^2 = 7.8%), and the model discrimination power had a concordance value of 63.7% ± 2.1.

4.4.3 Temporal synchrony

The beginnings of the peak rut ranged from September 29th to October 13th depending on the year. All years pooled together, behavioural observations happened from 19 days before the beginning of the peak rut to 26 days after (N = 853). The percentage of time spent herding varied throughout the mating season ($p < 0.001$), displaying a dome shape with a maximum at the beginning of the peak rut (Fig 4.3a).

The group's fission probability varied throughout the mating season ($p < 0.001$), displaying an inverse dome shape with a minimum at the beginning of the peak rut (Fig 4.3b). In addition, the mean group size increased the group's fission probability (slope ± SE = 0.16 ± 0.02, $p < 0.001$).

4.5 Discussion

During the breeding season, males may try to increase their mating opportunities by herding females into their harem, and females may continuously be on movement to sample mates, thereby influencing mating groups size, and hence the opportunity for sexual selection (Wade

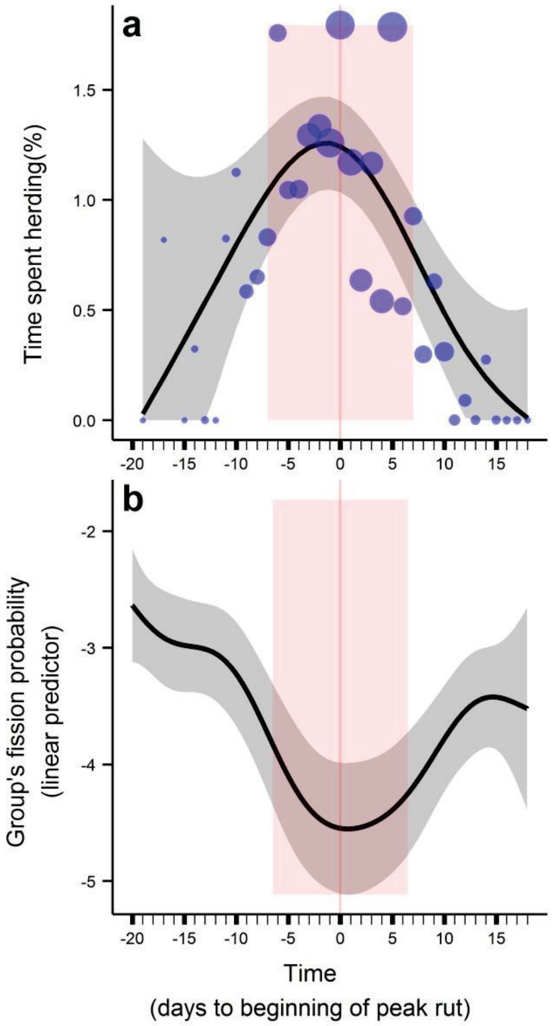


Figure 4.3 Temporal variations of herding frequency and group's fission probability. Temporal variations in the herding frequency of dominant males (a), and in the group's fission probability (b). Black lines represent the predictions and the grey areas surrounding them are their 95% confidence intervals. The red vertical bands represent the period during the rut ("peak rut week" versus "outside rut"), the darker red line the beginning of the peak rut week. Blue dots in panel (a) are the observed daily average of the time spent herding by dominant males, and their sizes are proportional to the number of observations

and Shuster 2004). In this study, we estimated the relative influence of male and female mating tactics on females' movement, using a herd of reindeer exhibiting fission-fusion group dynamics that we followed using GPS. Our results only supported the prediction about male herding ability (P_1), as we found the level of mating-related activities of high quality males to decrease group's fission probability and that, temporal variations of both herding and group's fission probability were exact opposites. Contrary to predictions, we found no evidence for female mate choice (P_2), or for harassment avoidance (P_3).

As males herded females, there was a tendency for groups to be more stable. The resulting decrease in fission rate, induced an increase in average group size (Couzin and Laidre 2009). Larger harems retain more estrous females (L'Italien et al. 2012). Consequently, more efficient is the herding, the greater the number of estrous females a male can have in his harem, depleting mating opportunities from his competitors, and consequently increasing the opportunity for sexual selection (Wade and Shuster 2004). This process is reinforced by the difference in herding ability among males. Low quality males are inefficient herders either due to their lower body condition (Isvaran 2005) or their inexperience for the youngest ones (Mysterud et al. 2003; Holand et al. 2006; Tennenhouse et al. 2011). Herding is expressed mostly at the beginning of the peak-rut. This suggests that males focused their attention, during the peak rut week or toward the end of the peak rut, on other mating behaviours such as courting or tending females. Together with interference competition, which happened through fights for dominance when groups merged together (Holand et al. 2012), herding provides an additional mechanism to explain high sexual selection in reindeer.

We found no evidence that female mate choice influenced the group's fission probability as groups without males were less likely to split than groups with males. However, females are known to change their behaviour during a short period of time around the estrous (Ozoga and Verme 1975; Richard et al. 2008; Stopher et al. 2011), to be choosy only during their estrous (McComb and Clutton-Brock 1994). Females may also express their choice through quick behaviour, such as joining satellite males outside the harem (Byers et al. 1994) or vocalizing when approached by satellite males (Bowyer et al. 2011) to increase agonistic interactions among males. Therefore, we may have to focus more on the estrous period to improve our understanding of the role of mate choice on female ungulates' movement. Moreover, we argue that the question "why females stay within a group" (Naulty et al. 2013; this study) is as important as the question "where are females going" (Byers et al. 1994; Stopher et al. 2011). Female mating tactics are also highly variable among individuals, being experience- and condition dependent (Byers et al. 1994). Therefore, it might be easier to detect female mate choice when studying individual behaviour, rather than group behaviour as we did in this study.

Variables representing harassment had a low statistical support in explaining the fission probability. The increase of fission probability with increasing group size is also inconsistent with the dilution effect of harassment (as observed in red deer *Cervus elaphus*; Carranza and Valencia 1999). Harassment level may, however, be more intense when females are solitary (Clutton-Brock et al. 1992), given also that females prefer to be with other females (Clutton-Brock and McComb 1993; Nefdt and Thirgood 1997). Consequently, females might only lessen the costs of harassment by avoiding being solitary. This is in accordance with earlier findings that the number of solitary females decreased during rut (Clutton-Brock et al. 1982; Kojola 1986).

A recent conceptual framework (Sueur et al. 2011b) hypothesized that social relationships are important in determining group stability. Our results validate this hypothesis as social environment (group size, presence of males, male quality) and social behaviour (herding) influenced group's fission probability. Although herding behaviour seems to be attributed to dominant males during the breeding season, the increase of the group's fission probability with group size is not season-specific (Pays et al. 2007). Indeed, both group size and presence of males decrease the level of synchrony in activity among individuals (Focardi and Pecchioli 2005; Michelena et al. 2008), a key factor explaining group cohesion (Conradt and Roper 2000; Marshall et al. 2012). The resulting negative correlation between group size and group cohesion could be reversed if the relative benefits expected from sociality (i.e. staying in a cohesive group) outweigh the benefits expected from reaching a desired patch (Conradt et al. 2009). In this predator-free reindeer herd, females maintain weak social bonds (Djaković et al. 2012), the group size does not decrease the harassment level, and food patches are widely dispersed. Consequently, there are few benefits expected from social cohesion which may explain the high fission rate observed.

Our study contrasted the relative effect of male and female mating behaviours in a highly sexually dimorphic ungulate, and clearly showed that high quality males, through herding and other mating-related activities, strongly influence females' movement pattern. While studies of female mating tactics are needed in mammals (Clutton-Brock and McAuliffe 2009), we advocate to concurrently evaluate hypotheses derived for both sexes, as sexual coercion is frequent (Bro-Jørgensen 2011), and female choice may be more apparent than real, a lesson learnt from primates (Muller et al. 2011).

4.6 Appendices

Appendix 4.1 Estimation of activity levels from activity sensors

We estimated the proportion of time spent feeding for females and the proportion of time spent in mating-related activities for males during the active bouts of time (i.e. excluding "resting time") using the recursive model (Body et al. 2012). During the rut, running, walking and standing are associated with male mating activities such as chasing males, herding females, threatening, grunting, courting, seeking copulation and being vigilant toward other males. In contrast, feeding is unrelated to mating activities (Tennenhouse et al. 2012).

Methods

Activity level

The dataset for the validation of the activity sensor was composed of a dataset previously reported (Body et al. 2012), in addition to the behavioural observation collected in the field during the 2011 rut season (i.e. additional 39 data for males, 133 data for females in similar condition with the Group B dataset; see Body et al. 2012). We used 80% of the dataset as training dataset to estimate the equation; then we assessed the quality of these models by calculating the bias and spread (see details in Pearce and Ferrier 2000; and an application example in Body et al. 2012) from the remaining 20% (the testing dataset). The final equations were fitted on the whole dataset.

We proceeded in three steps to estimate the proportion of time spent feeding for females, and the proportion of time spent in sexual activities for males: 1) we estimated the proportion of time spent resting for each record; 2) from these estimations, we applied a smoothing procedure to obtain bouts of resting time; 3) we fitted the activity equations on the data that were not included in a resting period (see "estimation of resting period" section, below for an example of these three steps).

Step 1: We estimated the proportion of time spent resting (i.e. rest versus stand/feed/walk/run) for both sexes using a generalized additive model (GAM). As graphical representation of the male equation displayed an obvious error for Y-values (forward-backward movement of the GPS collar) greater than 0.6, adjustment were made on those values.

Step 2: To obtain binary data, we applied a threshold at 60% of time spent resting, which is a conservative measure compared to a threshold at 50%, above which odds are higher to be resting. Due to the variability among collars, some collars were overall estimating low values of the percentage of time spent resting. Therefore, we moved down the threshold to the percentage of time spent resting corresponding to the 90th highest percentile if this one was lower than 60%. This procedure assured to catch the top layer of the estimations of the proportion of time spent resting (see "estimation of resting bouts" section for an easy visual identification of the top layer). Then, we transformed these binary data (rest/not rest) to bouts of resting time. We smoothed the sequential results based on sliding windows of five records centered on the data in focus. If at least four of them agreed with a given value, we attributed this value to the data in focus. Then, we repeated this procedure, but attributing the value of the majority of the five recalculated records to the focal data. This procedure allowed having bouts of resting lasting at least 45 min, i.e. three records. Forty five minutes spent resting allowed us to capture 93% of the resting periods, based on a 5 minutes frequency dataset (see "estimation of resting bouts" section).

Step 3: We estimated the proportion of time spent feeding for females (feed versus stand/walk/run) and the proportion of time spent in sexual activities for males (stand/walk/run versus feed) using a GAM in the recursive model. To set up these equations, we removed from the validation dataset any observation that included resting activity, as step 1 and step 2 removed for the application dataset the occurrence of resting behaviour. We applied these three steps of the procedure to the whole dataset of each individual taken separately.

The mean proportion of time a dominant male spent in sexual activities, disregarding his resting periods, was considered as the *DomAct* variable of that group. At each relocation time *t* we summed the proportion of time spent in sexual activities by satellite males (assuming they would have a cumulative effect on the disturbance level of the group), disregarding their resting periods. The average of these values was considered as the *SatAct* variable of the group. The mean proportion of the time females spent feeding was calculated for each group at each relocation time *t*, disregarding resting females. The average of these values was considered as the *FemEat* variable of the group.

Estimation of resting bouts

We determined the duration of resting bouts using data collected from activity sensors during the 2008 rutting season. During this year, one female was equipped with the same Tellus medium GPS collar as the ones used in this chapter. This collar was however set-up to record positions and the activity data every 5 min and recorded the data from September 20th at 00:00 to October 14th 23:55, i.e. 7193 relocation time.

We applied the two first step of the activity estimation described above to determine resting bouts. This procedure smoothed the resting bout for a minimal duration of three recording times. Using the 5 min data frequency, the minimal resting bout for this estimation was 15 min.

Results

Activity level

We recorded 479 behavioural observations for females and 389 for males, which were used as training and testing dataset to fit the resting models. The relationship between activity sensor values (X_{adj} and Y_{adj}) and the predicted proportion of time spent resting are represented for females (Fig 4.4a) and for males (Fig 4.4b). Both models had no bias ($p = 0.959$ for females; $p = 0.191$ for males), and had a weak spread ($p = 0.001$, slope = 0.86 for females; $p < 0.001$, slope = 0.78 for males; a slope of 1 would mean no spread). The dataset used for fitting the feeding (females) and sexual activity (males) models were composed of 365 (females) and 309 (males) behavioural observations. The relationship between activity sensor values (X_{adj} and Y_{adj}) and the predicted proportion of time females spent feeding is represented in figure 4.4c, and the relationship between activity sensor values and the proportion of time males spent in sexual activities is represented in figure 4.4d. Similarly to the previous validations, both models had no bias ($p = 0.897$ for females; $p = 0.229$ for males) and a weak spread ($p = 0.002$, slope = 0.84 for females; $p = 0.003$, slope = 0.84 for males). The spread observed for each model means that predicted values under 0.5 are slightly underestimating actual values, whereas predicted values over 0.5 are slightly overestimating actual proportion of activity. The absence of bias means that, on average, the estimations were neither overestimated nor underestimated.

Estimation of resting bouts

We present the estimation of the proportion of time spent resting (step 1, Fig 4.5a), the binary calculation of these data (step 2, Fig 4.5b) and the result of the smoothing procedure (step 2, Fig 4.5c) for a randomly chosen subsample of 24h (i.e. 1000 to 1288 recording times). Then we present the distribution of resting bout lengths according to the 2009 dataset, i.e. with a 5

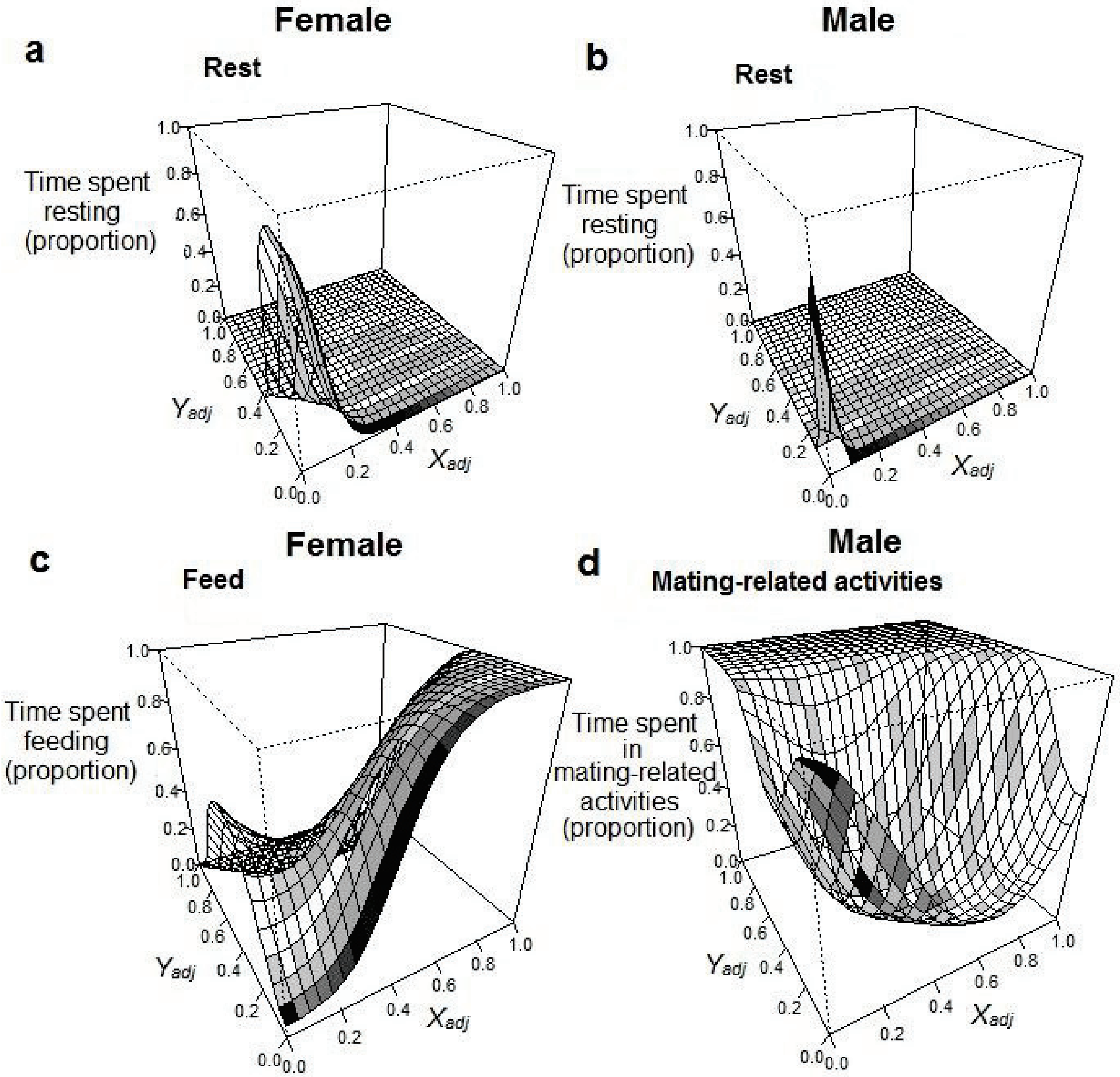


Figure 4.4 Activities probability according to activity sensor records. Relationship between the left-right (X_{adj}), the forward-backward (Y_{adj}) movements of the activity sensor and the proportion of time spent resting for females (a), and males (b), and of the proportion of time spent feeding for females (c), and in mating-related activities for males (d). The darkness of each square is proportional to the observed number of data with the corresponding [X_{adj},Y_{adj}] adjusted values

min data frequency (Fig 4.6). We identified 134 resting bouts, among which 92.5% lasted at least 45 min. The average duration of the resting bouts was 103 ± 51 minutes.

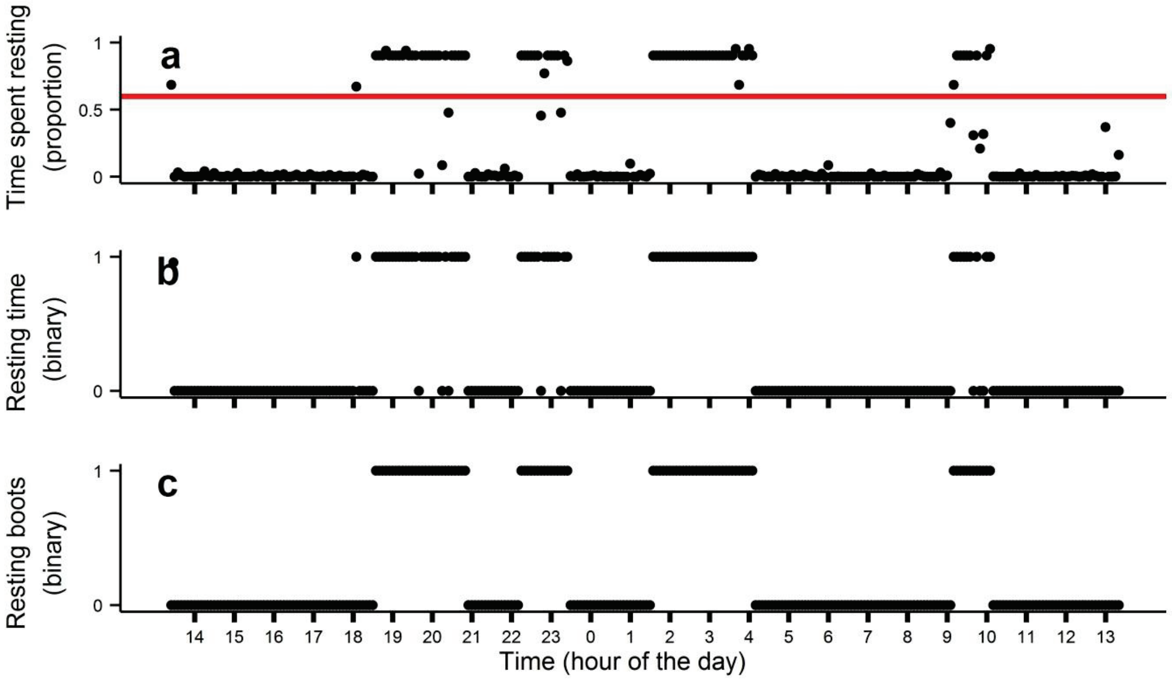


Figure 4.5 Steps of the estimation of the resting bouts. We estimated the proportion of time spent resting from the recursive model (a), then we applied a threshold at 0.6 (red line) to obtain binary resting time (b). We applied a smoothing procedure to clearly identify resting bouts (top layer, c). The calculation of the proportion of time spent feeding for females only applied to records of an active (i.e. excluding "resting") period (i.e. the bottom layer)

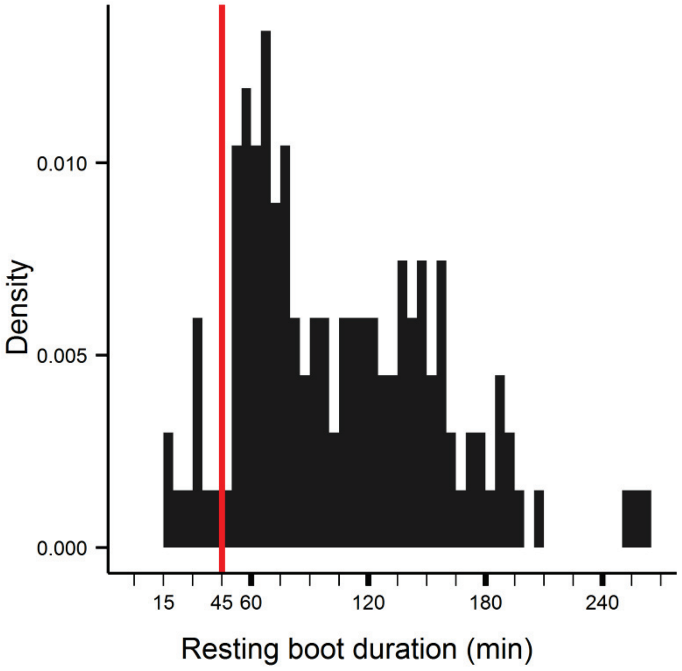


Figure 4.6 Duration of the resting bouts. The vertical red line correspond to the smallest duration of the resting bouts (i.e. 45 min) used in the exploratory variables

Chapter 5 Foraging competition in larger groups overrides the benefits of harassment avoidance during the rut in female reindeer

To be submitted, with credit to authors as follow: Ucchuddu S, Body G, Weladji RB, Holand Ø, Nieminen M

5.1 Abstract

Male harassment toward females during the breeding season may have a negative effect on their reproductive success by disturbing their foraging activity, thereby inducing somatic costs. Accordingly, it is predicted that females will choose mates based on their ability to provide protection or will aggregate into large groups to dilute *per capita* harassment level. Conversely, increasing group size may also lead to a decrease in foraging activity by increasing foraging competition, but this effect has rarely been considered in mating tactic studies. This study examined the importance of two non-exclusive hypotheses in explaining the variations of the female activity budget during the breeding season: the male harassment hypothesis, and the female foraging competition hypothesis. We used focal observations of female activity from known mating groups collected during the breeding season from a long term (15 years) study on reindeer *Rangifer tarandus*. We found that females were more disturbed (i.e. spent less time eating) in the presence of young dominant males, and marginally disturbed in the presence of satellite males, which supports the male harassment hypothesis. We also found that female disturbance level increased with group size, being independent of the adult sex ratio. Consequently, these results rejected the dilution effect, but strongly supported the foraging competition hypothesis. This study therefore highlights a potential conflict in female behaviour. Indeed, any gains from harassment protection were negated by an increase of 14-15 females, since adult males lead larger groups than young males.

5.2 Introduction

Female ungulate body condition is a good predictor of their reproductive success (Festa-Bianchet 1998; Ropstad 2000). Access to food resources is likely the first limiting factor of female reproductive success (Clutton-Brock and Vincent 1991; Hewison et al. 1998), due to the high energetic demands of lactation and gestation (Gittleman and Thompson 1988). Accordingly, female distribution during the mating season is mostly driven by resource distribution rather than male quality (Emlen and Oring 1977; Carranza 1995). Females may also suffer somatic costs from male sexual harassment during the breeding season (Holand et al. 2006). Maximizing access to resources may thus conflict with avoiding sexual harassment. In this study, we investigate whether such a trade-off exists in a highly polygynous ungulate.

Male sexual harassment toward females has been reported, in various taxa, to have negative consequences on female reproductive success (Linklater et al. 1999; McMahon and Bradshaw 2004; Makowicz and Schlupp 2013), and survivorship (Réale et al. 1996). Male sexual harassment is, indeed, an important constraint on female activity budget (Tobler et al. 2011; Weir 2013), and can potentially influence female mate choice (Bierbach et al. 2013). Consequently, it is suggested that female ungulates choose their mate according to their ability to provide protection against other males (Clutton-Brock and McAuliffe 2009; Bro-Jørgensen 2011). Females should avoid young males who are responsible for most of the harassment (Isvaran 2005; Holand et al. 2006), but they may also form larger groups to dilute the *per capita* harassment level (Clutton-Brock et al. 1992; Carranza and Valencia 1999).

However, grouping may also increase foraging competition among females (Danchin et al. 2008; Marshall et al. 2012), thereby reducing foraging payoff for the individuals (Stephens et al. 2007). Foraging competition is therefore in conflict with the dilution effect of harassment. Obtaining protection from adult dominant males is also in conflict with foraging competition, as adult males are known to lead larger groups than young males (L'Italien et al. 2012).

We evaluated the relative importance of (a) the male harassment hypothesis and (b) the female foraging competition hypothesis, on the activity budget of female reindeer *Rangifer tarandus*. Harassment from males and foraging competition between females occur through brief interactions, such as chases, herding and copulation attempts, or head kicks (Holand et al. 2006). These brief interactions, however, impact the main activities of females: eating, standing, and walking. The cost female reindeer suffered due to male harassment was, indeed, related to both an increase in harassment level and a decrease in time females spent eating (Holand et al. 2006). As female reindeer are continuously moving while eating, we can also expect that they would spend less time walking when disturbed. Conversely, we may expect that disturbed females would spend more time standing. Before being disturbed, females may scan for the position of dominant females to avoid aggression (Kojola and Nieminen 1988), for the food items discovered by subdominant females, or for the position of young males. After being disturbed, females may search for new food items or they may calm down after a chase. In the studied population, the variation in proportion of time spent standing can neither be attributed to predator vigilance, as the population is predator free, nor to resting behaviour as female reindeer usually lie down to sleep and to ruminate. Regardless of the exact mechanism (sexual harassment or foraging competition), the level of disturbance females experience in a mating group negatively correlates with their time spent eating and walking, and positively correlates with their time spent standing.

Therefore, we tested the predictions that (1a) the disturbance level should decrease with an increase in mating group size in accordance with the dilution effect related to the male harassment hypothesis, but (1b) should increase under the foraging competition hypothesis (Table 5.1). The dilution effect also predicts that (2a) the disturbance level should increase with the number of males per female (i.e. the adult sex ratio (ASR) of the mating group). In contrast, (2b) the ASR should not influence female disturbance level according to the female foraging competition hypothesis. According to the male harassment hypothesis, we expected an influence of the age of the dominant male on the level of disturbance experienced by females (Table 5.1). We predicted that (3a) females should be less disturbed by adult dominant males than by young dominant males, that (4a) the disturbance level should increase with the presence of satellite males, and that (5a) this increase should be higher in groups controlled by young males than in groups controlled by adult males. In contrast, under the foraging competition hypothesis we predicted (3b) no effect of the age of the dominant male, (4b) no effect of the presence of satellite males, and therefore (5b) no effect of their interaction (Table 5.1).

Table 5.1 Predictions related to (a) the male harassment hypothesis and (b) the female foraging competition hypothesis between female disturbance level, main activities and (1) the mating group size (Group size), (2) the adult sex ratio in the mating group (ASR), (3) the age of the dominant male (Y: young; A: adult), (4) the presence of satellite males, and (5) the effect of the presence of satellite males according to the age of the dominant male. "+" and "-" signs represent the predicted direction of the relation between variables

	(a) Male harassment hypothesis			(b) Female foraging competition		
	Disturbance	Eat/Walk	Stand	Disturbance	Eat/Walk	Stand
(1) Group size	-	+	-	+	-	+
(2) ASR	+	-	+	0	0	0
(3) Age	Y: + A: -	Y: - A: +	Y: + A: -	0	0	0
(4) Satellites	+	-	+	0	0	0
(5) Age × Satellites	Y: ++ A: -	Y: -- A: +	Y: ++ A: -	0	0	0

5.3 Methods

5.3.1 Study area, period and population

The study was conducted at the Kutuharju Field Reindeer Research Station, in Kaamanen, Finland (69°N, 27°E). We collected data from a semi-domestic reindeer population free ranging in two fenced areas: the southeast Sinioivi (13.4 km²) and the northwest Laulavaara (13.8 km²). Birch *Betula spp* and pine *Pinus sylvestris* forests, boggy areas and lakes characterized the enclosures. Data are based on direct behavioural observation collected from 1997 to 2012 (except 1998) during the breeding seasons (mid-September to late October). Herd compositions (a herd is the population in an enclosure in a particular year) varied between 26-92 females, 3-26 males with a sex ratio (percentage of males) between 4% and 28%, and in male age structure (i.e. adult only, young only or mix of young and adult; see Tennenhouse et al. 2012 for herd compositions). Males were fitted with VHF radio collars while females were fitted with colored collars, both with unique identification facilitating mating group composition determination and the monitoring of individual behaviour. As individual identification changed every year, we referred to them using Year-ID; an individual can have different Year-ID across years.

5.3.2 Group records

Within an enclosure, herd members split into unstable groups that undergo intense fission-fusion group dynamics (Holand et al. 2006; L'Italien et al. 2012; Chapter 3). The number of groups in an enclosure, as well as the average group size varied throughout the breeding season, each year (Chapter 3).

We located daily (usually between 09:00h and 18:00h GMT+2 h) each radio-collared male and his eventual mating group. Using Lent (1965)'s definition of a group, a mating group was considered "*an aggregation of individuals separated by some distance from other* *aggregations, showing coordination of activities, such as travelling together or resting and feeding together*", with at least one male and one female. We first recorded the mating group composition: number and identity of females and males, and the identity of the dominant male (see Tennenhouse et al. 2011 for details about dominant male determination).

Daily reports of mating group composition were considered independent due to the fission-fusion group dynamics (L'Italien et al. 2012). Manipulating the composition of herds by increasing the herd ASR, the density of individuals, or the male age composition (i.e. only young males, for instance) allowed us to obtain a larger range of group situation. Large groups dominated by young males in the presence of satellite males would have, for instance, rarely been observed in natural conditions.

5.3.3 Female behaviour records

Female behaviour was observed based on the focal observation technique (Martin and Bateson 2007). A randomly chosen, but active, female was observed for 15 minutes. Every 15 seconds, we recorded the activity of the female (rest, stand, walk, eat) as well as other behaviours (e.g., bush trash, social interactions). When possible, a new female was randomly chosen from the rest of the mating group. When visual contact with the female was lost during the focal observation, the procedure was stopped and the actual duration of the focal reported. Behavioural frequencies were divided by the focal duration to estimate the proportion of time spent performing an activity.

For this study, we only included focal observations that lasted at least 7.5 minutes. We also excluded observations from extreme mating group composition (i.e. 7 records from mating groups with ASR > 0.5; and 13 records from mating groups particularly large with ≥ 45 females) to avoid leverage effect on our analyses (Crawley 2007). These records from extreme mating group composition corresponded to 2.2% of the dataset.

5.3.4 Statistical analysis

We assessed the influence of the mating group composition on time females spent foraging, standing and walking using, for each activity taken separately, a generalized linear model (GLM) fitted with a logistic link function and binomial error structure, weighted by the focal duration. We first adjusted the most complex model including the effect of mating group size (number of females), the adult sex ratio (the percentage of males in a mating group, *sensu* de Jong et al. 2012), the age of the dominant male in the mating group (a categorical variable with two levels; Young: 1.5 or 2.5 years old; Adult: ≥ 3.5 years old) and the presence of satellite males (a binary variable: absence/presence). As we might expect the effect of the presence of satellite males to be modified by the age of the dominant male, we included an interaction term between these categorical variables.

We adopted an all subset approach (Symonds and Moussalli 2010), and therefore we fitted all of the simpler models derived from the above full model, for a total of 20 models. We chose the best model according to the Akaïke Information Criterion (AIC). We retained the most parsimonious model among the competing models that differed in AIC by less than 2 (Burnham and Anderson 2002). We estimated the percentage of deviance explained by a model using equation 5.1 below. All statistical analyses were performed using R 3.0.0 (R Development Core Team 2013).

Eq. 5.1

$$Deviance\ explained\ = (Null\ deviance\ - Residual\ deviance)/Null\ deviance$$

We tried to control for pseudo-replication in our dataset, however, mixed models failed to converge. This was certainly due to pseudo-replication being rare in our dataset (we recorded each Year-ID on average 2.2 times, only two Year-IDs were recorded more than 10 times, 85% of Year-IDs were recorded three or less times, and 47% of Year-IDs were recorded only once), leading us to use a GLM instead of a mixed GLM.

Then, we compared the relative effect size of an increase in mating group size, the change in the age of the dominant male, and the presence of satellite males. Using the model retained to explain the variations in proportion of time females spent eating; we calculated the difference in mating group size that would have the same effect as switching between the four compositions in males of mating groups: young dominant male without or with satellite males; adult dominant male without or with satellite males. This procedure can be quickly performed by dividing the difference among category coefficients by the mating group size coefficient.

5.4 Results

We recorded 908 focal observations of females, for a total of 220.8 hours of observation. These records came from the observation of 414 different females' Year-ID. Mating groups composition ranged from one to 44 females (average mating group size ± sd = 13.12 ± 8.8 females), and from 2% to 50% of males in the mating group (average ASR = 19.7% ± 11.3 of males). We recorded 367 young dominant males versus 541 adult dominant males, and 413 mating groups had no satellite males whereas satellite males were present in 495 mating groups. These values represent the overall dataset structure, not the particular group size and composition distribution of our herds, as multiple focal observations were made on each group.

Overall, females spent 62.0% of their time eating, 20.6% resting, 10.2% walking and 6.0% standing. The time spent resting was negatively correlated with the time spent eating (Pearson correlation: $r = -0.89$), and with the time spent walking ($r = -0.45$). The correlations among all other variables were small (stand and eat: $r = -0.21$; rest and stand: $r = -0.15$; eat and walk: $r = 0.13$; stand and walk = 0.12).

The variation in proportion of time females spent eating and standing were best explained by the model including the effect of group size, age of dominant male, presence of satellite and the interaction between age and satellite presence (Table 5.2), whereas the most complex model best explained the variation in proportion of time spent walking (Table 5.2). The retained models explained 6.79%, 0.81% and 1.34% of the deviance, for the eating, standing and walking models, respectively.

Increasing mating group size decreased the proportion of time spent eating (Table 5.3, Fig 5.1a), increased the time spent standing (Table 5.3, Fig 5.1b), and decreased the time spent walking (Table 5.3, Fig 5.1c). An increase in ASR only decreased the proportion of time spent walking (Table 5.3, Fig 5.2). Females in mating groups dominated by young

Table 5.2 Difference in Akaike Information Criteria (ΔAIC) compared to the model with lowest AIC value obtained for the eating, standing and walking models. We only present models with a ΔAIC ≤ +10, and the selected models are those with ΔAIC = 0

Model	Eat	Stand	Walk
Group size + ASR + Age × Satellites	+ 0.8	+ 1.8	**0**
Group size + Age × Satellites	**0**	**0**	+ 6.2
Group size + ASR + Age + Satellites			+ 7.9

Table 5.3 Coefficients ± standard error of the selected models. “Age” represents the coefficient corresponding to the effect of the age of the dominant male (Y: young; A: adult) in absence of satellite males, and “Satellites” represents the influence of adding satellite males for each age class category. Bold coefficients are significant (all p < 0.001)

	Eat	Stand	Walk
Group size	**-0.034 ± 0.001**	**0.009 ± 0.002**	**-0.013 ± 0.002**
ASR			**-0.53 ± 0.19**
Age classe:			
Young	**0.72 ±0.02**	**-2.72 ±0.04**	**-2.18 ±0.06**
Adult	**1.14 ±0.02**	**-3.07 ±0.05**	**-1.89 ±0.05**
Satellites:			
Young	**-0.23 ±0.03**	**-0.27 ±0.06**	**0.23 ±0.05**
Adult	**0.08 ±0.03**	**0.08 ±0.03**	0.03 ±0.05

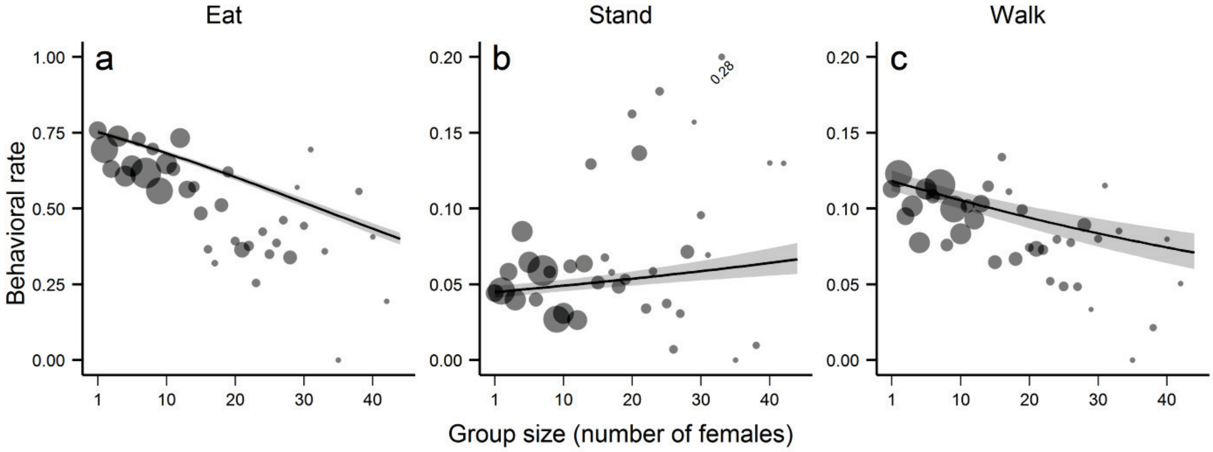


Figure 5.1 Effect of group size on the proportion of time females spent (a) eating, (b) standing, and (c) walking. Dots represent the partial residuals (conditional to ASR = 0.2 and the presence of an adult dominant male without satellite males) which have been averaged per group size. Dot size is proportional to the sample size for the corresponding group size. Lines represent predicted values with their 95% confident intervals

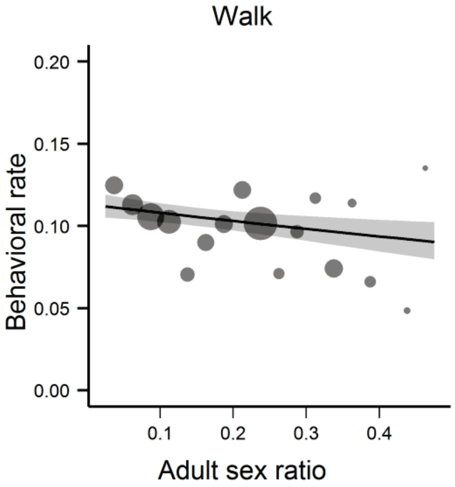


Figure 5.2 Effect of the adult sex ratio (ASR) on the time females spent walking. Dots represent the partial residuals (conditional to group size = 13 and the presence of an adult dominant male without satellite males) which have been averaged per 0.025 ASR intervals. Dot size is proportional to the sample size for the corresponding ASR interval. The line represents predicted values with its 95% confident intervals

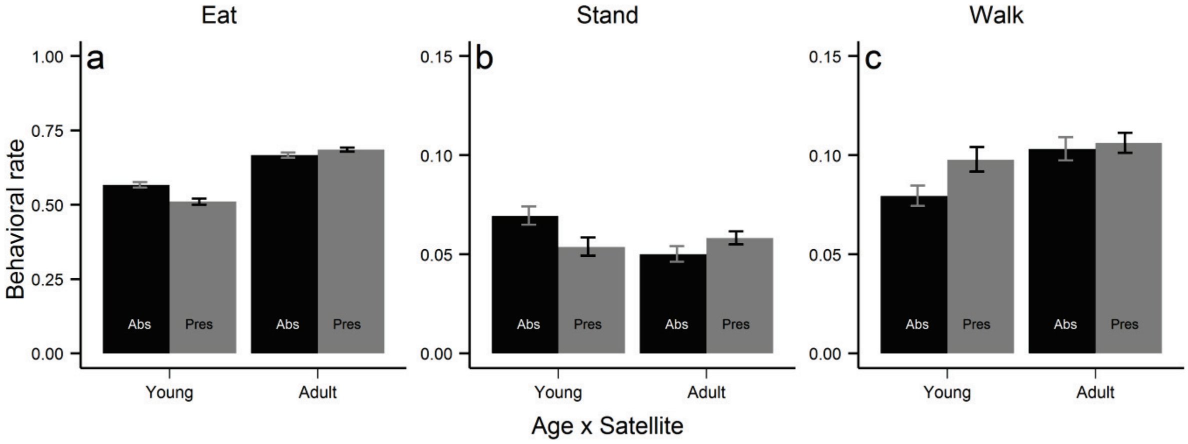


Figure 5.3 Effect of the age group of the dominant male (Young versus Adult), the presence of satellite males [black bars = absent (Abs), grey bars = present (Pres)], and their interaction on the proportion of time females spent (a) eating, (b) standing, and (c) walking. Bars represent the predicted values with their 95% confident intervals (conditional to the average group size = 13, and the average ASR = 0.2)

males, and in absence of satellites, spent less time eating (Table 5.3, Fig 5.3a), more time standing (Table 5.3, Fig 5.3b) and less time walking (Table 5.3, Fig 5.3c), than females in groups dominated by adult males. The effect of the presence of satellite males on female activity budget differed according to the age of the dominant male (Table 5.3, Fig 5.3). Indeed, the presence of satellite males decreased the time females spent eating in groups controlled by young males, but increased time spent eating by females in groups controlled by adult males. Similarly, their presence increased the time females spent standing when the group was controlled by young males, but decreased it when in groups controlled by adult males. The presence of satellite males also increased the proportion of time females spent walking when the group was controlled by young males, but had no significant effect when controlled by adult males.

We found that switching from an adult toward a young dominant male is equivalent to increasing the mating group size by 12 females in the absence of satellite males, and that adding satellites males had the same effect as increasing the mating group size by six females with a young dominant male, whereas it had the same effect as decreasing the mating group size by two females with an adult dominant male. When we compared the two most typical situations occurring in natural populations (i.e. adult dominant male with satellite males versus young dominant male without satellite males), females spent the same proportion of time eating if the mating group of the young dominant male was smaller by 14-15 females. If there were satellite males in both groups, females would spend the same time eating if the difference in group size is 21 females.

5.5 Discussion

In this study, we tested the male harassment hypothesis and the female foraging competition hypothesis by assessing the influence of the mating group size and composition on female activity budget, and we found support for both.

We first assessed the overall activity budget of female reindeer, and found that they spent only 20% of their time resting. This is inconsistent with earlier findings by Rodney and Boert (1985) and Kojola (1986) reporting that females spent about 40% of their time resting during the breeding season. However, when active (i.e. not resting), some of the rates were rather close to earlier work; for instance, we observed that active females spent 80% of their time eating. This is in line with earlier results from this population (see Holand et al. 2006) and with above studies which reported 66% and 80-88%, respectively. The bias on the time spent resting is mostly due to the observers selecting an active female at the beginning of the focal sampling (see method). Since the overall bias was constant among observations, we believe our interpretation of the influence of mating group composition on female activity budget is valid.

We found a strong support for the female foraging competition hypothesis. The disturbance level was, indeed, positively correlated with mating group size. A decrease in foraging activity in larger groups is not specific to the breeding season, and has been reported in many ungulates (bighorn sheep *Ovis canadensis*, Berger 1978**;** pronghorn *Antilocapra americana*, Lipetz and Bekoff 1982; fallow deer *Dama dama*, Focardi and Pecchioli 2005) and in other taxa (Marshall et al. 2012). A negative correlation between group size and travelling activities has also been reported in different ungulates (bison *Bison bison*, Fortin et al. 2009; roe deer *Capreolus capreolus*, Pays et al. 2012). This trend was explained by a conflict in movement direction among individuals, which constrained the speed of large groups. A higher rate of interactions among individuals in larger groups either directly through agonistic behaviours or indirectly through acquisition of social information provides another explanation for constraint of foraging-related activities, i.e. eating and walking.

Our results strongly rejected the dilution effect on harassment level for two reasons. First of all, as explained above, the disturbance level increased, instead of decreasing, with increasing mating group size. Secondly, the adult sex ratio had little or no effect on female activity budget. Although being in a group has been widely reported to decrease the harassment level received by females when compared to solitary females (Clutton-Brock et al. 1992; Byers et al. 1994), our results suggest that females are equally negatively affected in large and small mating groups. The increase of reindeer group size during the breeding season might not be the result of a harassment avoidance female mating behaviour through the dilution effect.

As predicted by the male harassment hypothesis, females were more disturbed when the dominant male was young than when he was an adult. Young males are known to have less developed social rutting behaviours than adult males, and to cause extra stress on females (Clutton-Brock and Parker 1992; Komers et al. 1999; Tennenhouse et al. 2012). Their behaviours when dominant, such as herding females, may also be less efficient (L'Italien et al. 2012) and regularly turn into a chase. Females may also be responsible for this disturbance. Given that adult females are almost as big as young males (Melnycky et al. 2013), they are more likely to resist herding behaviours when performed by young males rather than by adult males. Regular chases of young males by female reindeer have, indeed, been reported (Kojola and Nieminen 1988).

Whether satellite males are harassing females or not remains unclear. Indeed, the presence of satellite males had the same effect on the time spent eating and on the time spent standing, which is incoherent with our definition of disturbance. Accordingly, our discussion from this point forward will only consider the effect of satellite males presence on the time females spent eating, as an overall measure of disturbance. The age of the dominant male seemed to influence the effect of the presence of satellite males. Although the presence of satellite males appeared to induce an additional disturbance on females when the dominant male was young, they actually reduced the disturbance when it was an adult dominant male. It is easier to understand the latter effect if we consider that adult dominant males switched from female directed behaviours in the absence of competitors to male directed behaviours in the presence of competitors, for instance by trading time spent herding females with time spent chasing males. More generally, we can argue that even if satellite males were harassing females, their effect appeared to be marginal.

Our results highlighted a trade-off in female mating tactics. Avoiding harassment from young dominant males by merging with a group dominated by an adult dominant male may not be beneficial, due to the increase in female foraging competition. Such a group fusion would, indeed, only be beneficial if the mating group size did not increase by more than 14-15 females, or 21 females depending on whether or not there were satellite males in the initial group. This value corresponds to a typical adult male's mating group size. Consequently, in many occasions, group fusion would not be beneficial to females, regardless of the age of the dominant male present in their group. Conversely, there are no costs associated with the fission of a large group: the cost of being dominated by a young male being compensated by the decrease in female foraging competition. The existence of this trade-off does not constrain, therefore, fission-fusion group dynamics. An efficient strategy to get around this conflict or trade-off might be to decrease female foraging competition, for instance by maintaining long-term bonds among females as seen in feral horses (Cameron et al. 2009), but this remains to be investigated in reindeer.

Chapter 6 Measuring the intensity fission and fusion from longitudinal group size data

The following chapter is based on the manuscript: Body G, Weladji RB, Holand Ø, Nieminen M Measuring the intensity fission and fusion from longitudinal group size data. Submitted to *Behavioral Ecology and Sociobiology* on 17/03/2014.

6.1 Abstract

The intensity of the group fission and fusion is critical for understanding social dynamics in gregarious species. Indeed, variations in intensity reveal an individual's response to environmental changes. For example, changes to perceived predation risk, or the value of resource such as mates or forage. Moreover, group mixing also influences risk of disease transmission, which may have implications for game management. Here we test a method developed to assess the intensity of the group dynamics from longitudinal records of group composition. This method has raised concerns because it had not been validated. Upon validation of the method here, we propose a correction. Subsequently we validate the corrected model using accurate measures of the frequency of fusion and fission events. We do this using an enclosed herd of reindeer *Rangifer tarandus*. Each individual of this herd was followed by Global Positioning System collars and from field observations. We found that the previous method erroneously recorded changes in the average group size, rather than in the intensity of the group fission and fusion. This error is rectified using our corrected method.

6.2 Introduction

In animal societies, the fission and fusion of groups is increasingly being recognized as an important component of social structure (Aureli et al. 2008; Sueur et al. 2011b). Three components of the fission-fusion group dynamics can be measured: the variation in spatial cohesion, the variation in group size, and the variation in group composition (Aureli et al. 2008). The method to measure the variation in group composition, i.e. the social network analysis (Croft et al. 2008), is powerful (Krause et al. 2007; Sueur et al. 2011a; Godde et al. 2013) and has already provided compelling results (Couzin 2006; Fischhoff et al. 2009; Pearson 2009; Kelley et al. 2011; Parra et al. 2011). Other methods are being developed to study individual spatial cohesion (Schellinck and White 2011; Aureli et al. 2012). But the study of group size variations faces methodological problems, such as the definition of group (Haddadi et al. 2011). Yet, measuring the variation in group size allows researchers to quantify the intensity of the fission-fusion group dynamics, i.e. the overall frequency of fission and fusion events. A low intensity of the group dynamics represents stable groups; whereas high intensity represents highly fluid groups – whether or not populations are structured in multilevel societies. Individual perception of predation risk (Gower et al. 2009) or the outcomes of mating behaviours (Chapter 4) can influence the intensity of the group dynamics, which itself can influence other processes such as the transmission of diseases (Couzin and Laidre 2009). Because the intensity of fission and fusion is central to the study of group size variation, it is critical that the methods used for its quantification to be first validated.

Researchers have used different methods to study the determinant of the intensity of fission and fusion. Recording the duration of particular groups to assess their fusion and fission probability is one of the method (Focardi and Pecchioli 2005; Pays et al. 2007; Pays et al. 2012; King et al. 2012; Aureli et al. 2012). This method best explains the influence of group composition and individual's activity on group fission probability. A second method, that takes well into account the influence of the density of groups on the fusion probability, consists in recording all the fission and fusion among all groups present at a particular location (e.g., Pépin and Gerard 2008; Fishlock and Lee 2013). Development of remote sensors, such as Global Positioning System (GPS), provided a third method by measuring fission and fusion rates among pairs of individuals (e.g., Fortin et al. 2009). This third method efficiently explained landscape variation of group dynamics. Although the three above methods are not limited to the variables they best explain, integrating these different components to explain group dynamics requires particular and highly controlled conditions; for example, a population released into a new environment (Haydon et al. 2008), experimental, or pseudo-experimental conditions (Chapter 3). While these methods improve our understanding of group dynamics, they are not suitable for long-term survey or management purpose because they are too intensive. However, long-term surveys of natural populations often consist in recording group size, within which some individuals are identified. Developing a method to assess the intensity of group dynamics adapted to these field records would both improve the opportunities to study fission-fusion group dynamics, and would enable connecting theoretical findings in this field of research to management policies.

Such a method has been proposed by Gower et al. (2009) on elk *Cervus elaphus* and then used on pronghorn *Antilopacra americana* (White et al. 2012a), but has not yet been evaluated. Gower's method is based on the variation of group size within which one particular individual is found during a particular interval of time. Results of these two studies showed a strong correlation between this measure of the intensity of the group dynamics and the average group size. Such a variance-mean correlation is a common pattern as it corresponds to issues of heteroscedasticity (Crawley 2007). As a result, this method merits further reflection to ensure that the variation of the measured variable (variance in group size) actually reflects a variation in group dynamics and not a variation in average group size, which would have a different biological interpretation.

Consequently, our study has three objectives. 1) To improve Gower's method by using the coefficient of variation instead of the group size variation. 2) To validate the modified-Gower technique using an accurate record of the frequency of group fusion and fission events in a controlled herd. 3) To perform a sensitivity analysis on the model parameters to provide best practice advices when using the modified-Gower technique. To achieve these objectives, we analyzed the temporal variation of the intensity of the group dynamics reported in Chapter 3 in reindeer *Rangifer tarandus* during one breeding season. This enclosed herd of 45 individuals was intensively followed by GPS collars. From this dataset we measured the actual intensity of the group dynamics as the number of group dynamics events (fission, fusion, departure, junction, group appearance and group disappearance) observed during a particular period of time. Group sizes were also measured using this GPS dataset as well as daily field-based observations. These datasets formed the two testing datasets used to compare Gower's method with the modified-Gower method and to validate the modified-Gower method.

6.3 Methods

6.3.1 Study herd and GPS methods

In 2011, we studied a herd of 45 adult reindeer (34 females, 11 males) free to move in a 13.4 km² fenced enclosure (the Sinioivi enclosure of the Kutuharju Field Reindeer Research Station, Finland), during 39 days in the breeding season (September 8th -October 18th). We equipped all individuals, but one female, with GPS collars, which synchronously recorded their positions every 15 minutes (hereafter "recording"). On each recording, we assessed aggregations of individuals based on a chain rule stating that two neighbours were in the same aggregation if they were closer than 89 m (Chapter 3). Based on the aggregation's composition, we then differentiated "groups" (≥ 2 females, any number of males) from "solitary females" (1 female, any number of males), and we disregarded "males group" (0 females, ≥ 1 male). By comparing group composition among successive recordings, we were able to record six group dynamics events: fusion (groups merging together), fission (one group splitting into two or more groups), departure (one female leaving a group), junction (one female joining a group), group appearance (solitary females joining each other), and group disappearance (one group splitting into solitary females, see Chapter 3 for the complete methods). Every day from September 21st to October 17th, we also located groups, and recorded their composition in the field. We recorded identities of individuals that composed the groups using unique collar identification numbers (L'Italien et al. 2012).

6.3.2 Gower's method and the CV method

Gower et al. (2009) calculated their measure of the intensity of group dynamics from the group size variation as: "*the absolute difference between a given group size and the mean group size for [the period in focus]*". The corresponding statistical method was a linear mixed model with the individual ID as the random term, the measure of the group size variation during the studied interval as response variable, and time intervals, among others related to their specific questions, as explanatory variables.

For our purpose, we adapted this method in three ways. First, to enhance the temporal variability, we calculated group size variation on sliding windows of four days instead of on time intervals. These windows are referred to as "the period in focus". Second, when applying the method, we only used the group sizes related to a particular individual experience, and not the population dataset. This approach should match more closely individual exposure to group size (Vander Wal et al. 2013), in line with the individual-based observation of group size (i.e. the typical group size; Jarman 1974). Third, we used the time as a continuous explanatory variable, instead of a categorical variable, and consequently we used a generalized additive mixed model (GAMM) instead of a linear mixed model, as it provided a better fit of the temporal variability. We used a smoothing factor of k = 8; and individual identity as random factor (i.e. intercept only). These adaptations applied to both Gower's method and the modified-Gower method.

The modified-Gower method (hereafter CV method), only differed from Gower's method on the response variable. Instead of the group size variation, we calculated the coefficient of variation (i.e. $\mathrm{CV} = \mathrm{standard\ deviation/mean}$) using the same data (i.e. the group sizes recorded per individual during the sliding window). We compared the actual variation in group dynamics (see paragraph below) to the predictions obtained from both methods when applied to testing datasets (see second paragraph below).

6.3.3 Actual average group size and group dynamics records

Group compositions recorded by GPS on the 15 min frequency formed dataset A. From this dataset we calculated the average group size (Eq. 6.1) for the period in focus and we assessed its temporal variability by fitting a GAM with time as explanatory variable, and a smoothing factor k = 8 (model$_{\mathrm{A1}}$).

Eq. 6.1

$$Average\ group\ size_t = \frac{total\ \#females_{t \to t+1}}{\#groups_{t \to t+1} + \#solitary\ females_{t \to t+1}}$$

Where # means number in the period in focus starting at time *t* and finishing at time *t+1*.

We used dataset A to calculate the intensity of group dynamics as the sum of the frequency of each group dynamics event during the period in focus. To match more closely each individual exposure to group dynamics, we weighted events related to groups (fusion, fission, disappearance) by the proportion of females in groups, and we weighted events related to individual females (junction, departure, appearance) by the proportion of solitary females (Eq. 6.2). To assess temporal variation of the intensity of the group dynamics, we fitted a generalized additive model (GAM) with time as an explanatory variable, and a smoothing factor k = 8 (model$_{A2}$).

Eq. 6.2

$$Intensity_t =$$

$$\frac{\#females\ in\ groups_{t\rightarrow t+1}}{total\ \#females_{t\rightarrow t+1}} \times \sum_{t\rightarrow t+1}(\#fusion + \#fission + \#disappearance\ ) +$$

$$\frac{\#solitary\ females_{t\rightarrow t+1}}{total\ \#females_{t\rightarrow t+1}} \times \sum_{t\rightarrow t+1}(\#junction + \#departure + \#appearance\ )$$

Where # means number in the period in focus starting at time *t* and finishing at time *t+1*.

6.3.4 Testing datasets

Dataset B was a subsample of dataset A, only including group composition every four hours. We did so to match classical GPS dataset. We recorded on each recording of dataset B the individual exposure to group size (in number of females). Dataset C corresponded to group composition recorded daily from direct observation. We recorded the individual exposure to group sizes (in number of females) on each day. Occasionally individual females were recorded multiple times within a day. We retained these records in dataset C. Datasets A, B and C had different timeframes (15 min, 4 h and 1 day, respectively). To compare results based on these datasets, each fitted model was then used to predict a fourth dataset. Dataset D was composed of a single timeframe of four hours frequency and which covered the whole studied period (Fig 6.1).

We applied Gower's method and the CV method on both dataset B (Gower's model$_B$; CV model$_B$) and dataset C (Gower's model$_C$; CV model$_C$). The predicted values on dataset D

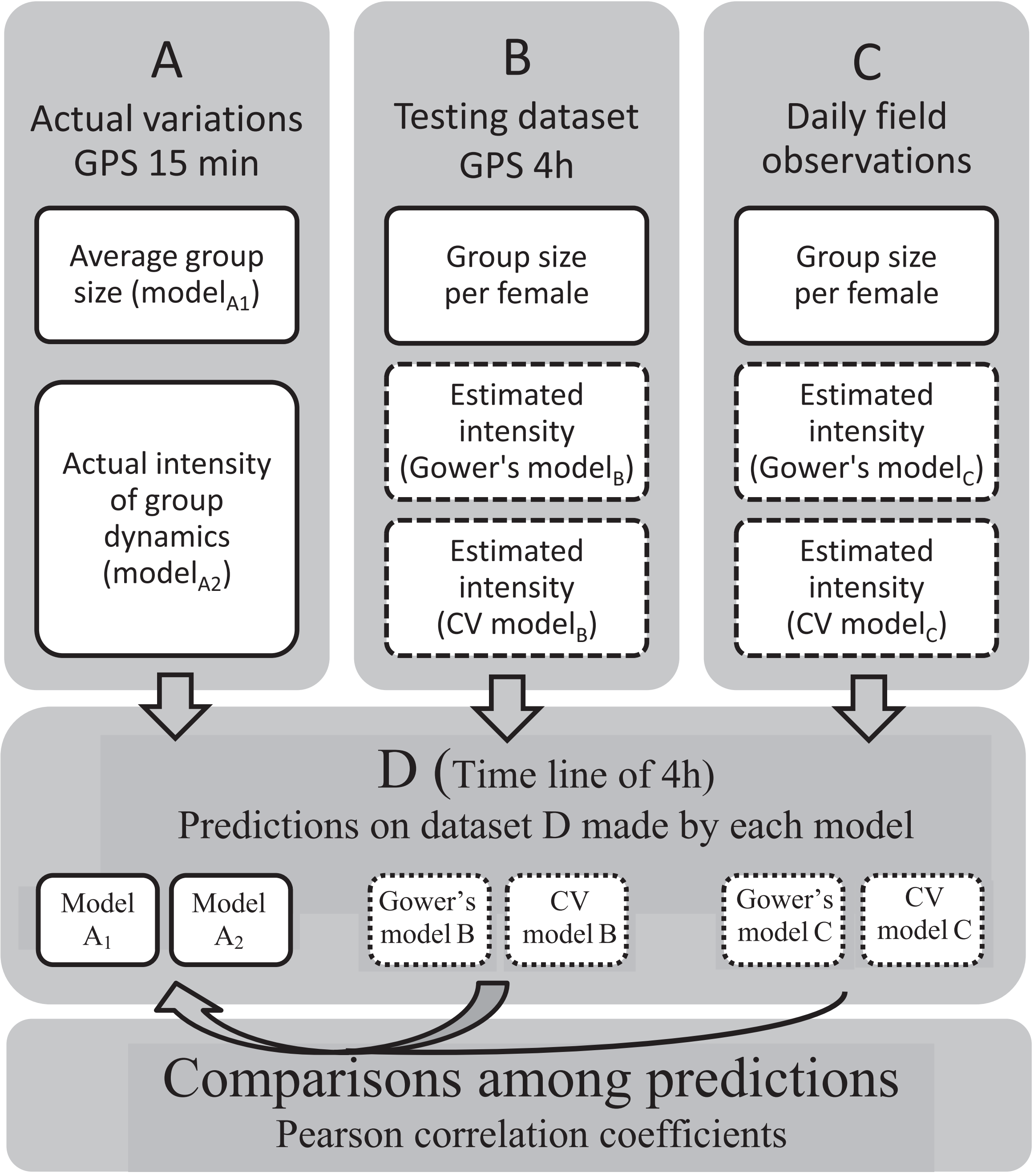


Figure 6.1 Organization of datasets A, B, C, D in methods. Actual variation in group size and intensity of group dynamics is obtained from dataset A; intensity of group dynamics is estimated by the Gower's method and the CV method from dataset B and C; actual and estimated temporal variations are compared based on their prediction made on dataset D. Continuous-line boxes correspond to actual data or their derived prediction, while dashed-line boxes are estimations of the intensity of group dynamics

from model$_A$ were compared to the predicted values on dataset D from models $_B$, and from models $_C$ based on the Pearson correlation coefficient (Fig 6.1).

6.3.5 Sensitivity analysis

We performed the CV method on a range of data frequencies and window lengths based on the dataset A. We rarefied data frequencies to 2, 4, 8, 12, 24, 36, 48 h, and enlarged window lengths to 4, 8, 12, 24, 48, 72, 96, 120, 144, 168, 192, 216, 240 h (i.e. from 4 h to 10 days), with all possible combinations for which there were at least two records per window. We calculated the Pearson correlation between the predictions on dataset D made by the model$_{A2}$ and the predictions on dataset D made by models derived from the combination of the two parameters.

Then, we ran a logistic model (logistic as the Pearson correlation is limited by the interval [0,1]) to assess the effect of the data frequency and the window length on the Pearson correlation coefficient. We included in the logistic model the interaction between the data frequency and the window length to verify whether the number of data per window was a good criterion to improve the Pearson correlation.

Using an arbitrary chosen window length of four days and a data frequency of 12 h, we performed a sensitivity analysis on the number of females followed. We used 33 (i.e. the maximal), 30, 25, 20, 15, 10, 5, and 2 randomly chosen females, with 10 different subsamples for each subsample size (except for 33 females which was the full dataset). Again, we calculated the Pearson correlation between the predictions on dataset D made by model$_A$ and the predictions on dataset D made by the CV models based on each subsample of females. Then, we performed a logistic model to explain the variation of the Pearson correlation coefficients by the number of females followed.

6.4 Results

6.4.1 Methods comparison and validation

Similar results were obtained from the GPS records (dataset B, Fig 6.2) and from the records based on direct observation of the group compositions (dataset C, Fig 6.3). The actual average group size increased until the peak rut week, then decreased (Fig 6.2a, Fig 6.3a). The actual intensity of the group dynamics showed the opposite pattern, decreasing until the beginning of the peak rut, and then increasing (Fig 6.2b, Fig 6.3b). Predictions made by Gower's method were weakly and negatively correlated with the actual intensity of the group dynamics using dataset B ($r = -0.40$, $p < 0.001$, Fig 6.2c), and showed no significant correlation when applied to dataset C ($r = 0.12$, $p = 0.53$, Fig 6.3c). However, predictions made by Gower's method were positively correlated with the variation in average group size ($r = 0.64$, $p < 0.001$; $r = 0.41$, $p = 0.03$, for datasets B and C, respectively).

In contrast, predictions made by the CV method showed a pattern of temporal variation consistent with the actual changes in intensity of the group dynamics. Predictions were strongly and positively correlated with actual variations when applied either to dataset B

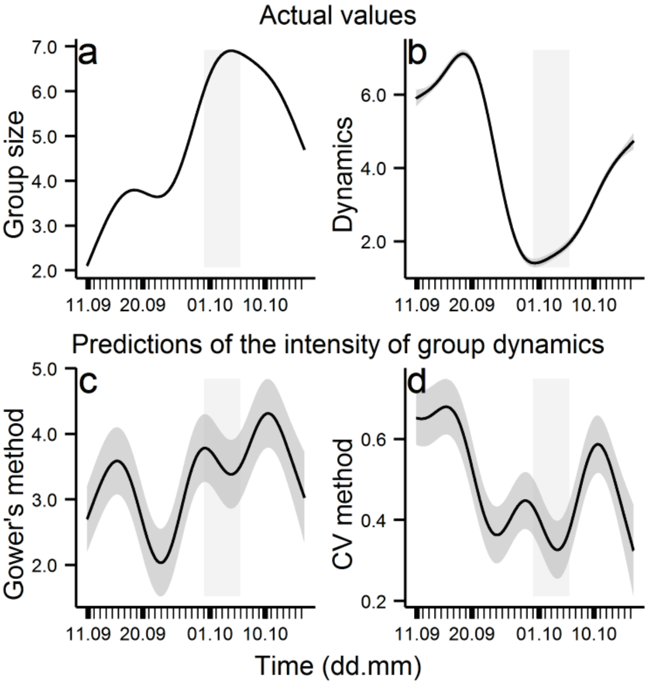


Figure 6.2 Temporal variability of (a) the actual average group size, and (b) the actual intensity of the group dynamics derived from the GPS dataset A. Predictions of the temporal variation of the intensity of the group dynamics are presented according to (c) the Gower's method and (d) the CV method, using the GPS dataset B. The temporal variations are presented with their 95% confidence interval. The grey vertical bar represent the peak rut week

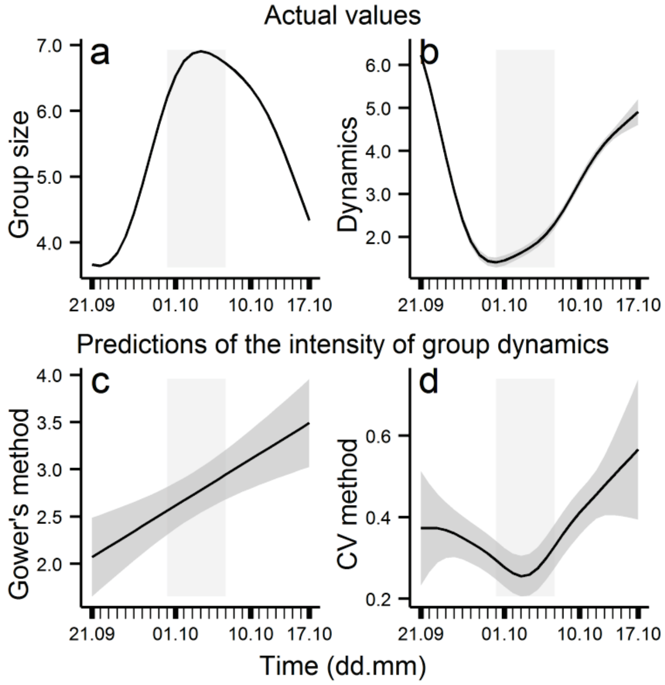


Figure 6.3 Temporal variability of (a) the actual average group size, and (b) the actual intensity of the group dynamics derived from the GPS dataset A. Predictions of the temporal variation of the intensity of the group dynamics are presented according to (c) the Gower's method and (d) the CV method, using the direct field-based observations in dataset C. The temporal variations are presented with their 95% confidence interval. The grey vertical bar represent the peak rut week

(r = 0.68, p < 0.001, Fig 2d) or dataset C (r = 0.75, p < 0.001, Fig 3d). The correlations between these predictions and the average group size were weaker and negative compared to those with the group dynamics intensity (r = -0.52, p < 0.001; r = -0.44, p = 0.02, for datasets B and C, respectively). The statistically significance between the CV method's prediction and the actual average group size is likely to result from the significant negative correlation between the actual variation in group dynamics and the actual variation in average group size (r = -0.70, p < 0.001).

6.4.2 Sensitivity analysis

None of the parameters used in the CV method were statistically influencing the Pearson correlation between the predicted and the actual intensity of the group dynamics (data frequency: p = 0.69; window length: p = 0.40; interaction: p = 0.89; and number of females followed: p = 0.20). Yet, visual inspection of the graphs suggested that the Pearson correlation could be better for data recorded every 48h than for data recorded every 2h (Fig 6.4a), and for the shortest window length (Fig 6.4b). Overall, it meant that windows including only two data were providing the highest Pearson correlation (Fig 6.4c). In addition, increasing the number of females followed seemed to decrease the variability of the obtained Pearson correlation (Fig 6.4d), which was particularly problematic when only two females were followed.

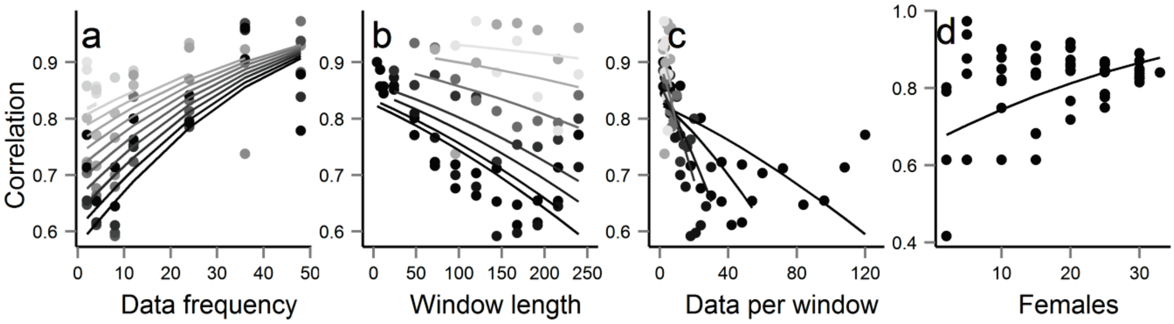


Figure 6.4 Variation of the Pearson correlation between the actual and predicted group dynamics intensity according to the variation of (a) the data frequency, (b) the window length, (c) the number of data per window, and (d) the number of females followed. Points are actual Pearson correlation coefficients and lines are the predictions from the logistic models. The darkness of points and lines are correlated in panel (a) to the window length (lightest: window length = 4h; darkest: window length = 240h), and in panels (b, c) to the data frequency (lightest: data every 48h; darkest: data every 2h)

6.5 Discussion

In this methodological study, we corrected Gower's method and validated the CV method to estimate the intensity of the group dynamics from longitudinal records of group size. Gower's method led to erroneous conclusions, as the estimated variation in intensity of the group dynamics was only recording the variation in average group size. Therefore, we advise that previous studies using Gower's method (Gower et al. 2009; White et al. 2012a) to be updated with the CV method (or modified-Gower method) to validate or correct their interesting findings. The CV method performed well on daily field-based group composition records, and can therefore be applied on studied natural populations where group size records include some recognizable individuals.

We recommend using the shortest possible sliding window, given that windows should include at least two records per individual. We showed also that it was not necessary to increase the frequency of data to improve the fit of statistics. However, the frequency of data should match the temporal scale of the biological question. In our system, we were looking at the temporal variation of the group dynamics intensity during the breeding season. Therefore, daily records were more appropriate than, for instance weekly records, as it would have provided too few data. Researchers should also invest more time obtaining regular records of a subsample of females rather than obtaining irregular records of a large number of females. If the number of females followed is sufficiently large, a valuable addition would be to bootstrap the analysis and obtain confident intervals.

The variation of the intensity of the group dynamics is a pattern representing group stability, not a process. Processes underlying change in this pattern could be assess following self-organization principles (Couzin and Krause 2003) and should be studied at the individual or group level. The influence of both external variables (e.g. landscape, population density), and internal variables (e.g. fear of predation, sociality, mating tactics, see Sueur, King, et al. 2011) should be investigated from the fission or fusion probability of groups or individuals. However, as we illustrated here, to avoid erroneous interpretations it is critical that patterns be accurately quantified prior to testing process.

Chapter 7 General Discussion

The objective of this thesis was to study female mating behaviour in reindeer *Rangifer tarandus*. For that purpose, (1) I developed appropriate methodology (Chapters 2, 3, 4), (2) I analyzed the determinant of the fission-fusion group dynamics (Chapters 3, 4) and I validated field based methods in order to follow such patterns in other population (Chapter 6). (3) I identified proximal processes involved in group dynamics variation, and the associated mating behaviour related to either sex (Chapters 4, 5). Below, I review the main findings on the research as well as their implications for further studies.

7.1 Methodological development

In a recent review Krause et al. (2013) emphasized "*[the promising] significant advances in new basic and applied research on animal social systems*" offered by "*reality mining*", i.e. the systematic and remote study of entire populations. They also advocated collecting extra layers of data, such as remote behavioural observations, in addition to individual proximity records. In this thesis, I developed and applied such methods. I identified individual-level patterns (Chapter 4) that influenced population-level patterns (Chapter 3), partially addressing what Krause et al. (2013) identified as "*one of the major unresolved challenges in the field of animal social network*". Social network *per se* remains to be studied in reindeer, including in the studied population. However, this was not the aim of my research.

Methods to assess animal behaviour from remote sensing records are still in development (Heurich et al. 2012; Augustine and Derner 2013). The application of such methods mostly focused on biological rhythms of activity (van Oort et al. 2007; Loe et al. 2007; Stache et al. 2013; Krop-Benesch et al. 2013). The recursive model, developed in this thesis (Chapter 2), goes further to estimate the proportion of time spent in given activities, for instance the time spent eating or in mating-related activities (Chapter 4). Predicting the proportion of activity rather than binary values also solves Heurich et al.'s (2012) methodological shortcomings. They were predicting a single activity on 5 min activity records, which we found unrealistic, given that a third of our 30 s duration records corresponded to mixed activities (Chapter 2). Overall, our method appeared to be reliable for detecting general patterns of the main behaviours. However, short and infrequent behaviours, such as herding or chasing, must be observed directly. A 15 min sampling rate proved to be ideal to accurately identify resting bouts (Chapter 4), and therefore to avoid a strong source of bias in estimating eating and mating-related activity rates.

Although activity states are classically estimated from methods based on individual movement (see Schwager et al. 2007; Van Moorter et al. 2010; Owen-smith et al. 2012), we did not applied one of them for a few reasons. Firstly, validation of the estimations obtained by these methods through direct observation is difficult, as it is impossible to use data collected on individuals that are spatially constrained (i.e. within small enclosure). These validations are also dependent on the sampling rate as the distance individuals move while eating, for instance, would be different from GPS data obtained on a 15 min frequency (as in 2011) and on a 1 hour frequency (as in 2009). Secondly, these methods are biased towards activities related to movement and they cannot record trade-offs such as spending time foraging versus in vigilance. In addition, they do not detect within group variations, such as the difference between male and female activity budget. Finally, assessing activities only from an accelerometer and not individual movement insured statistical independence with GPS data, allowing further comparison of data collected from both sources (e.g. the variation of the distance travelled while eating).

A further methodological development using remote activity sensors would be the identification of estrous period using variations of female daily activity pattern (Kojola 1986; Firk et al. 2002), in addition to their proximities with males.

The method developed to study the fission-fusion group dynamics (Chapter 3) focused on an intermediate spatial scale, because I was not interested in inter-individual distance among group members. In addition, I was neither looking at inter-group distance, nor at group movement. Therefore, the method could be seen as spatially implicit, as it still relied on geographic distance to identify groups. Early analyses revealed that the fission and fusion rates were more sensitive to position recording rate than to the maximal intra-group distance. For this reason, in addition to the purpose of detecting resting bouts, I recommend to keep the same recording rate on further study based on the finding from this study population.

The full-experiment control provided data adapted to reality mining and allowed an accurate measurement of the fusion and fission intensity. This is particularly interesting for validation of field based methods as performed in Chapter 6, and this experimental set up would be best appreciated by setting up an agent-based model that reflects the study system (Aureli et al. 2008; Sueur et al. 2011b). Ideally, this method should be coupled with social network analysis (Croft et al. 2008) to measure both the second (variation of group size) and third (variation of group composition) axes of the fission-fusion group dynamics framework defined by Aureli et al. (2008).

7.2 The fission-fusion group dynamics

Previous knowledge on reindeer fission-fusion group dynamics relied on daily group composition records (L'Italien et al. 2012; Djaković et al. 2012), which *a posteriori* was a very broad scale. Understanding fission-fusion group dynamics was challenging. Consequently, I first used a data mining approach, resulting in the relatively descriptive nature of Chapter 3. Description of natural systems remains an important part of ecology (in addition to the hypothesis testing approach), and it is a necessary step for newly started study systems. Indeed, I learned that relatively rare events, such as female departure and junction from/to groups were important to maintain the group size distribution (Chapter 3). At the beginning of the mating season, these intense group dynamics resulted in a stable (stationary) distribution of group size in the population (i.e. the average group size in this thesis), which corresponds to self-organization principles discussed by Couzin and Krause (2003).

Using a funneled approach, I showed that the self-organization theory applied to this population. I identified that a decrease of the group splitting propensity was responsible for the increase in average group size (Chapter 3). Then, I demonstrated that male herding behaviour was the main mechanism decreasing the group splitting propensity (Chapter 4). Though I used a top-down approach, from population pattern to individual behaviour, one should remember that the actual process works in the opposite direction. As male herding increases, groups become more cohesive, which results in an increase in group size. In terms of natural selection, genetic material related to herding behaviour is selected based on the fitness advantages it provides: herding increases male fitness as increasing group size increases male ability to monopolize estrous females (Emlen and Oring 1977). This in turn increases the opportunity for sexual selection (Wade and Shuster 2004).

Reversion, or the splitting of groups into sub-groups that quickly merge together, are an important part of fission-fusion group dynamics (Chapter 3, Pays et al. 2007). It can be understood as the natural dynamics of lightly cohesive groups in habitats where they long distance visual contacts can be maintained, such as on plains. Visual contacts are, however, not necessary for reversion. Splitting events increase the local density of groups, which increases the probability that two groups merge together (Chapter 3) through random movement with no intention to keep contact with the other sub-group. The high number of reversions, and more generally the high intensity of group dynamics (as defined in Chapter 6), may raise the question of why reindeer groups are so weakly cohesive.

However, one may equally wonder why females should be cohesive during rut. I found that mate choice did not seem to influence group cohesiveness (Chapter 3). Rather, I showed that the observed increase in group size during the rut (Chapter 3) should decrease the cohesion of the group (Chapter 4) through female foraging competition (Chapter 5). Though protection against harassment proved to be beneficial to females, it cannot induce cohesiveness due to the trade-off with foraging competition (Chapter 5). For instance, in a group dominated by a young male, females would be better if they induce a fission of the group, as they may find themselves in the sub-group free from the young harassing male (Chapter 5). On the other hand, those that stay with the young male do not necessarily suffer higher cost, as there is no dilution effect (Chapter 5). In the case of medium sized groups, merging would yield no advantages, as losses from foraging competition cost would offset any gains from harassment protection (Chapter 5). In large groups dominated by adult male, females would benefit from splitting, as the reduction in foraging cost worth the potential increased harassment cost, as one young satellite male may become dominant in one of the sub-group (Chapter 5).

Using this advantage/benefit approach, one can expect a high splitting rate in the study population, which contradicts with the higher rate of fusion than fission initially reported (Chapter 3). Therefore, in accordance with the self-organization theory, we should consider that groups are not an entity that behave, but that group patterns reflect individual behaviours (but see Couzin (2009) for influences of group size on individual behaviours). Females may not necessarily estimate the group size in which they are, nor the potential benefits of splitting the group in two.

Females could, however, be highly responsive to the solitary state versus within-group situations. The higher merging propensity of groups may come from the merging propensity of individuals, switching from solitary to within group situations (Chapter 3). The cost associated with the solitary situation appears to be high (Appendix 7.1), as has also been reported in fallow deer (Clutton-Brock et al. 1992; see also references in Couzin 2009). Indeed, female fallow deer seemed particularly nervous when solitary, and they preferred joining other females rather than males (Clutton-Brock and McComb 1993; McComb and Clutton-Brock 1994). A similar process could occur in our system, explaining the finding that the number of solitary females decreased first (Chapter 3). According to findings on fallow deer *Dama dama* (McComb and Clutton-Brock 1994), a higher harassment of females outside groups may be specific to estrous females, which may contribute to the highest group cohesion during the peak rut (Chapter 4). Long chases of females, likely in estrous, were indeed observed in the field, but always when females were solitary. However, avoiding being solitary will have a marginal effect on group size and group's splitting probability compared to the herding effect of males (Chapters 3, 4). In conclusion, the absence of a strong pressure on females to increase their cohesiveness is a robust explanation for the intensity of the fission-fusion group dynamics to remain high in the study herd.

7.3 Female mating behaviours

Female harassment avoidance is expected to be a major mating tactic of female ungulates (Clutton-Brock and McAuliffe 2009; Bro-Jørgensen 2011). Although sexual harassment influences female activity budget (Chapters 4, 5, Appendix 7.1) and is costly to females (Holand et al. 2006), the mechanism through which females avoid harassment is not clear. Indeed, we rejected both the dilution effect, i.e. that increasing group size decreases *per capita* harassment level, and that harassment comes from satellite males (Chapters 4, 5). Yet, we found the harassment to be highest when females were solitary (Appendix 7.1) or in the presence of a young dominant male (Chapter 4, 5). The efficient herding behaviour of adult dominant males (Chapter 4) has likely the same negative consequences on females than harassment (Chapter 5). The presence of satellite males may thus be beneficial for females (Chapter 5) by reducing the herding activity of the dominant males.

I reported a negative consequence of large aggregation on female activity budget and I concluded this could not be a female mating tactic to avoid harassment (Chapter 5). However, consequences of harassment can be very different to consequences of foraging competition, as sexual harassment through chases or forced copulations can induce injuries and even death (Réale et al. 1996). Therefore, females may accept to pay the cost of foraging competition to decrease the frequency of extreme harassment events. Increasing group size may increase the number of obstacles and it may therefore reduce duration or length of chases. Increasing group size or the adult sex ratio also increases the proportion of time males spent in competition (i.e. in male-male interactions; Tennenhouse et al. 2011), thus reducing the time they spend herding. Further studies should therefore test whether the frequencies of chases and copulation attempts from both dominant and satellite males decrease with increasing male-male competition.

I found that the herding behaviour of males was a key component of reindeer mating systems, reinforcing sexual selection through scramble competition among males (Chapter 4). Coercive tactics have been widely reported in mammals (Cappozzo et al. 2008; Muller et al. 2011; Bro-Jørgensen 2011), and these male tactics has regularly been associated with female counter tactics: polyandry (Aloise King et al. 2013), and inter females cooperation (Cameron et al. 2009; Möller 2012). Multiple mating in reindeer is rare (Mysterud et al. 2009), and is unlikely a female mating tactic. Female-female cooperation in reindeer cannot be excluded, as association patterns have not been evaluated among unrelated females. Yet, association among related females are limited to mother-daughter (Djaković et al. 2012). Social network analysis may reveal association among non-related females.

If herding is truly the main behaviour shaping male and female mating behaviours and their consequences on fission-fusion group dynamics, one can expect strong variation of characteristics of the group dynamics with variation in the proportion and age of males present in the herd. Such variations could be assessed by applying the method developed in Chapter 6 on group composition records collected in the Kuthuarju Field Reindeer Research Station the last 15 years.

Although female mate choice is expected to be widespread in mammals (Clutton-Brock and McAuliffe 2009), I did not find evidence that female reindeer were choosing their mates (Chapter 4), nor that they were leaving groups to sample mates (Appendix 7.2). This result is consistent with previous studies reporting that female reindeer neither avoid inbreeding (Holand et al. 2007), nor use the major histocompatibility complex as a criterion to select their mates (Djaković 2012). Yet, one should not conclude that female reindeer do not have preference for certain mates: I performed conservative analyses (Appendix 7.2) and, in presence of coercive male tactic, female mate choice may be hard to detect (Bisazza et al. 2001; Muller et al. 2011).

Further attempts to detect female mate choice in reindeer should focus on the estrous day of individual females. During their estrous, we occasionally observed females courting males and they may also change their grouping behaviour. Activity and movement patterns should be refined to individuals, and one should use the exact day of estrous for each female as the reference (see above 7.1 for method suggestion), rather than the peak rut week, to perform these analyses. Experiments may also be performed to assess the preference function of females, and personality analyses may help in understanding variation in choosiness among females.

7.4 Conclusion

This research showed that fission-fusion group dynamics might be shaping male and female mating behaviour. In other mating systems based on stable harems, male herding activity would not have induced an increase in harem size. In our system, we found the increase in harem size to be the results of group fusions. Due to the trade-off between protection from harassment and foraging competition, there were no constraints on females to increase their cohesion, and accordingly no need for females to form social bonds and hence permanent groups. In addition, fission-fusion group dynamics may have benefited females by providing them the opportunity to see different dominant males.

I showed that fission-fusion group dynamics was influenced by male and female mating behaviour. Male herding ability stabilized groups, reducing the intensity of the group dynamics. Female harassment avoidance also decreased the number of solitary females in the population, slightly increasing group sizes and reducing the intensity of group dynamics.

As a follow up to my thesis, I recommend two areas for further research. The first would focus on the variation of mating behaviour among females at the individual level and how it relates to their exact estrous day. The second would focus on the determinant of fission-fusion group dynamics, for instance by evaluating the influence of habitat openness or of the activity synchrony among group members.

7.5 Appendices

Appendix 7.1 Harassment cost

Sexual harassment from males is costly to females (Holand et al. 2006). Within group, both satellite males (Carranza and Valencia 1999) and dominant males (Chapter 5) may harass females, but young males seem to be harassing the most (Holand et al. 2006). Despite harassment occurs within groups, females aggregate during the breeding season (Clutton-Brock and McComb 1993), and join safe territories to avoid harassment (Clutton-Brock et al. 1996; Carranza and Valencia 1999).

Because they may be harassed up to four times more frequently outside a group than within a group as in fallow deer *Dama dama* (Clutton-Brock et al. 1992), females may rather try to avoid being solitary, than trying to dilute *per capita* harassment level by being in larger groups (Chapter 5). Avoiding being solitary could indeed explain why the number of solitary females decreases during rut (Clutton-Brock et al. 1982; Kojola 1986, Chapter 3).

Here, we tested whether female reindeer *Rangifer tarandus* were more harassed by males outside than inside groups. We predicted that the percentage of time females spent eating would decrease when they are solitary with a male, as compared to when they are alone (i.e. solitary with no male) or when they are within a group. This relationship may depend on the quality of the male.

Method

To estimate whether the harassment cost occurred outside groups rather than within groups, we compared the proportion of time females spent eating in four situations (categorical variable *Situation*). We compared three solitary situations (one female without males, with a low quality male, or with a high quality male; see Chapter 4 for the definition of male quality) to the group situation (where the time spent eating by females is independent of the male quality, $p > 0.51$). We included the influence of the rutting season using the variable *Period*, which separate the pre and peak-rut weeks (during the rut) to other period (outside rut, see Chapter 4). The proportion of time spent eating was estimated by the activity sensor values using the same procedure as for the *FemEat* variable in Chapter 4.

We used generalized linear models with binomial errors (as we fitted proportions) to assess the effect of *Situation*, *Period*, as well as their interaction, on the proportion of time spent eating by females. We selected the best model among the full model (*Situation* + *Period* + *Situation:Period*) and the four simpler ones using the AIC (Burnham and Anderson 2002). We selected the most parsimonious model among those with a $\Delta AIC \leq 2$. If we retained the interaction term, we then performed a post-hoc test (a Tuckey test on the arc sinus transformed values) to compare for each situation the effect of the period, and for each period the difference between the group situation and each of the three solitary situations.

In addition, we reported the frequency of each solitary situation for each period to assess the variation of the risk of being followed by a male, especially a low quality one.

Results

We retained the most complex model, including the effect of *Situation*, *Period* and of their interaction; since the second model (without the interaction) had a $\Delta AIC = +10.7$. Overall, it seems that the proportion of time females spent eating only decreased when females were solitary with a low quality male during the peak rut (Fig 7.1).

Indeed, in this situation, females spent less time eating during the rut than outside the rut (-7.9%, $p_{adj} < 0.001$); and during the rut, they spent less time eating than females within a group (-5.6%, $p_{adj} < 0.001$). In addition, whereas 19.4% of solitary females were with a low quality male outside the rut, this proportion increased to 45.1% during the rut.

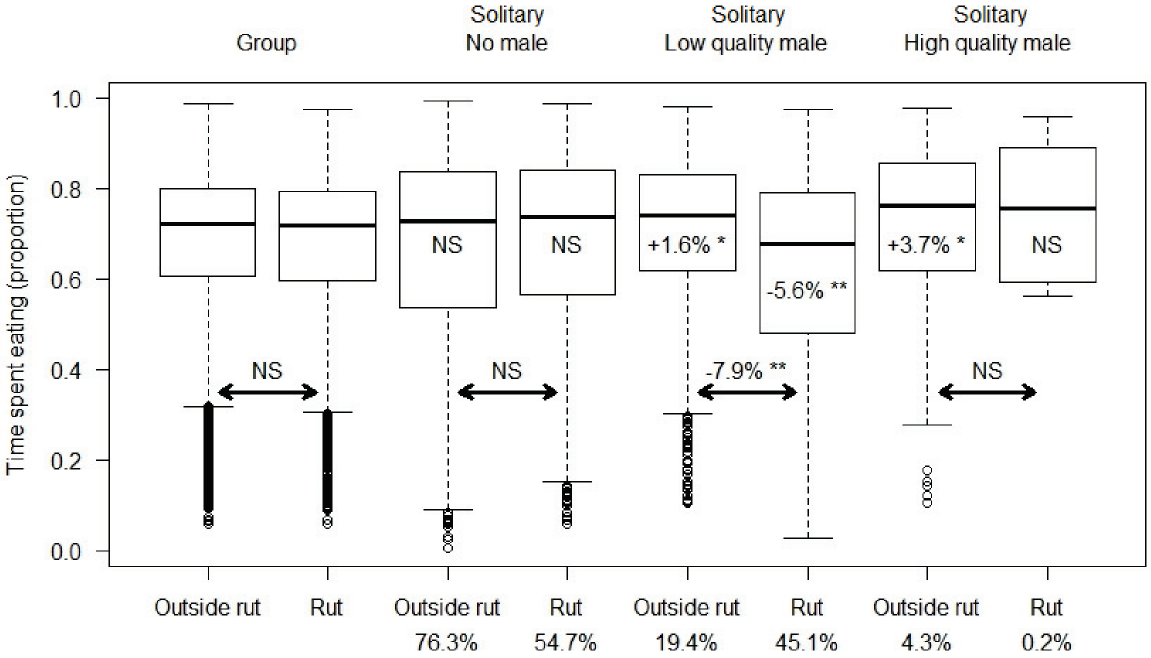


Figure 7.1 Proportion of time female spent eating for each situation (within group, solitary with no males, solitary with a low quality male, solitary with a high quality male) and for each period (outside or within rut influence period). Horizontal arrows represent the test of difference between periods within the same situation, and are associated with the significance level. Significance levels within boxes correspond to the test of the difference between the solitary situation and the group situation for the same period. When significant, the actual difference of the percentage of activity is displayed. The relative frequencies of each solitary situation within periods are displayed at the bottom of the figure, summing up to 100% for each period. Significance levels: NS: $p > 0.05$; *: $p < 0.05$; **: $p < 0.001$

In contrast, the period did not influence the proportion of time spent eating for females in other situations (all $p_{adj} > 0.36$). The group situation was not different from the solitary situation with no males for both period (both $p_{adj} > 0.73$), or from the solitary situation with a high quality male during the rut ($p_{adj} = 0.92$). Outside the rut, solitary females with a low or a high quality male spent more time eating than females within a group (estimated difference: 1.6%, $p_{adj} = 0.025$ and 3.7%, $p_{adj} = 0.021$ for low and high quality males, respectively). For both periods, solitary females were most often without males (76.3% and 54.7% of the records outside and during the rut, respectively), and were rarely followed by a high quality male, especially during the rut (4.3% and 0.2% of the records for outside and during the rut, respectively).

Discussion

We validated that females suffered more from the harassment outside groups, but only during the rut and only from low quality dominant males. As solitary females are more often followed by low quality males during the rut, it is more risky for them to leave a group. Therefore, they should stay within the group, independently of the harassment level they experience within the group (Chapter 5). This change in behaviour from outside to during the rut would result in higher cohesiveness of the group, and would consequently increase the average group size (Chapter 3). We also validated the idea that low quality males induce a higher cost than high quality males, likely due to harassment, as previously reported for young males (Holand et al. 2006, Chapter 5).

In conclusion, females are not evaluating harassment level within a group, but adjust their behaviours to the fear of leaving the group. A parallel with predation risk could be interesting to develop, whereby the avoidance of a risky situation (i.e. being solitary) could explain variation of group size with no need to precisely assess anti-predator benefits of being in group, such as the increase of vigilance, the dilution or the confusion effects.

Appendix 7.2 Sampling strategy

Movement among harems have been interpreted as evidence of mate sampling strategy in female ungulates (pronghorn *Antilopacra americana*: Byers et al. 1994, 2005; reindeer: L’Italien 2010). The main arguments for this interpretation have been that the frequency of these movements increased until copulation (Byers et al. 1994; L’Italien 2010), and that these movements were energetically costly to females (Byers et al. 2005).

These reports are based on the daily variation of group compositions. However, in fission-fusion group dynamics, daily variation of group compositions may not be due solely to individual female movement, which could be interpreted as evidence of sampling strategy, but also to the natural group dynamics *per se* (Chapter 3). Indeed, females, passively following each other, could meet different dominant males when their group merged with another one or when their group split.

The latter pattern provides a unique opportunity to assess whether females actively move among harems to sample males. We tested this sampling strategy hypothesis by predicting that females leaving groups would meet more dominant males than passive females following the fission-fusion group dynamics.

Methods

Area and study herd

We studied a semi-domestic herd of reindeer in Kuthuarju Field Reindeer Research Station in Kaamanen, Finland (69°N, 27°E) during the breeding season 2011 (September 8th -October 18th). The herd, composed of 11 males (from 1.5 to 5.5 years old) and 34 females (from 1.5 to 10.5 years old), was released into the Sinioivi enclosure (13.4 km²) as described in Chapter 4.

Group dynamics

Group dynamics was recorded from GPS collars, which equipped each individual, but one male (ranked 4) and one female. The GPS synchronously recorded the animal positions at each recording time *t*, on a 15 min frequency. Group fission and group fusion events occurred when at groups of at least two females joined or split. Consequently, single female departure and junction did not influence the group dynamics (see complete method in Chapter 3).

Sampling strategy

For each observed female, we recorded the number of different dominant males they met, from the GPS records, using two different sliding windows ending at the recording time *t*. We used a 14 days window as did Byers et al. (1994) for their work on pronghorn; and a two days window which is the duration of the estrous period in reindeer (Ropstad 2000). At each recording time *t*, we calculated the average number of males met by females and we reported the highest and lowest values. For this analysis, we excluded females whose GPS collars had problems (see Chapter 4) to avoid bias due to incomplete records.

We compared these values to 1000 spatially implicit simulated random movement. Simulated females were moving following the fission-fusion group dynamics according to the links among group identities (see Chapter 4). They were not allowed to perform departure nor junction events. In case of fission, the next group was randomly chosen among the groups produced by this fission event. In case of group disappearance, we attributed the female to a randomly selected group, which is a conservative measure (that may allow simulated females to see a larger number of males). Then, for each simulated female, we recorded the number of different dominant males they met using the two sliding windows. At each recording time *t*, we calculated the average number of males met by simulated females, as well as the 95% confidence interval. We compared graphically the observed and simulated distribution of curves for the two sliding window durations.

Results

Simulated females using a random walk within the fission-fusion group dynamics were able to meet as many different dominant males as actual females, for both sliding windows (two days, Fig 7.2a; 14 days, Fig 7.2b). A detectable difference among the two curves only occurred during the pre-rut week with a sliding window of two days (from September 24th to October 1st, Fig 7.2a), where actual females seemed to meet more different dominant males than simulated females.

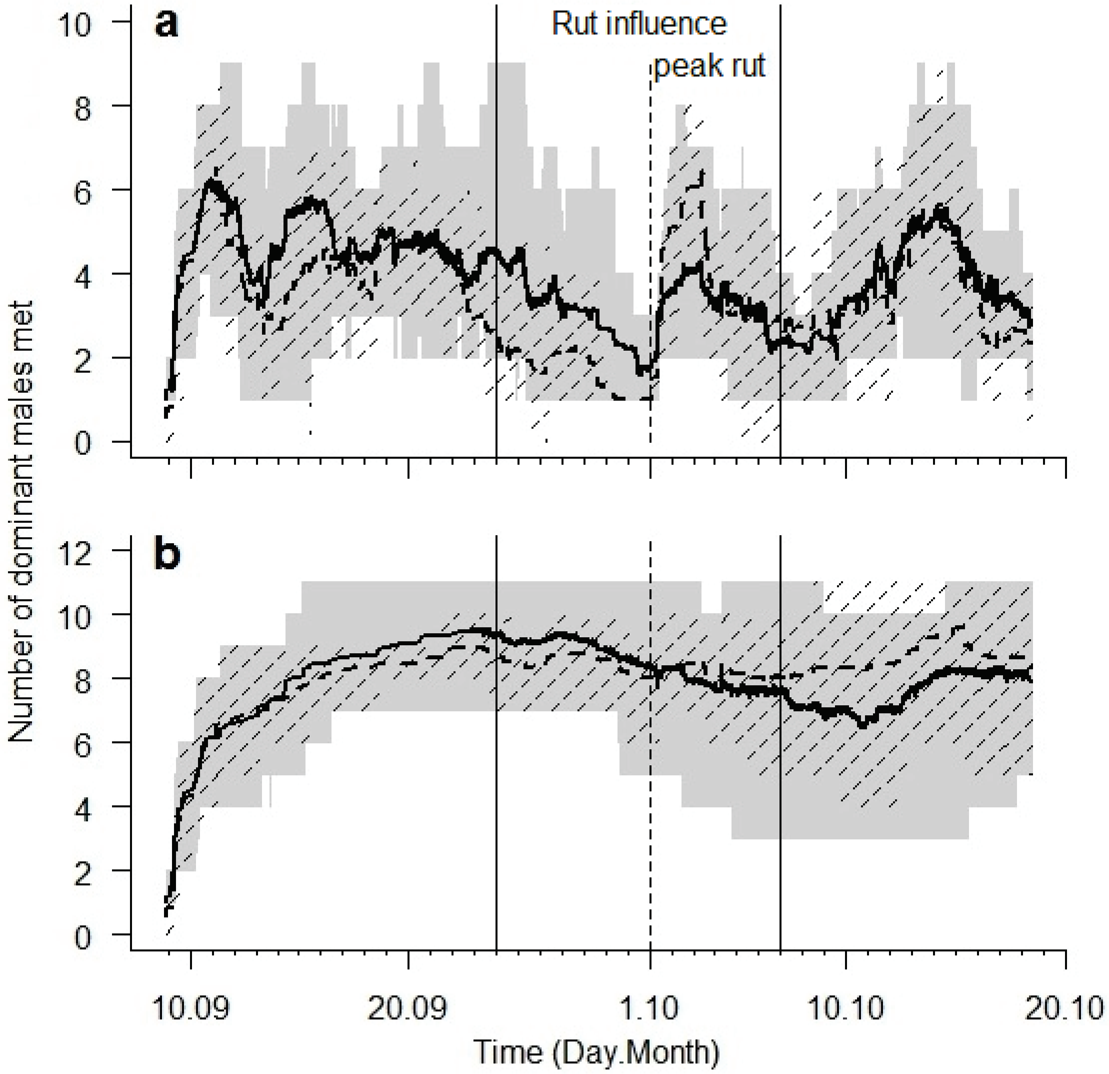


Figure 7.2 Number of different dominant males females met during sliding windows of (a) two days and (b) 14 days, for the 31 observed females (average: continuous lines, minimal and maximal: grey polygons) and for the 1000 simulated females that passively followed the fission-fusion dynamic (average: dashed lines, 95% confidence intervals: polygons made of the diagonal dashed lines)

Discussion

We did not find strong evidence that females were actively sampling males. A difference in the number of dominant males met by simulated (passive) females and actual (active) females was only found during the pre-rut week, using a two days window. By leaving groups, actual females met more different dominant males than what would have been possible by following the fission-fusion group dynamics. However, the number of dominant males met was decreasing during this period, which is inconsistent with a sampling strategy.

Mate sampling strategy is expected to vary among females (Byers et al. 2005). Females in low body condition are indeed not expected to use sampling strategy, in contrast to females in good body condition. Using mate sampling strategy can also vary among years (Byers et al. 2006). A general increase of the sampling cost should decrease the number of females sampling mates (Wiegmann et al. 2013). The lack of evidence of female mate sampling strategy in this herd of reindeer, in 2011, could be due to the fact that we did not take into account individual differences, and we have a single year of study. In addition, female reindeer may not need to actively leave harems to sample mates, as simulated females were still able to meet many different dominant males.

References

Adrados C, Verheyden-Tixier H, Cargnelutti B, Pépin D, Janeau G (2003) GPS approach to study fine-scale site use by wild red deer during active and inactive behaviors. Wildl Soc Bull 31:544–552

Agresti A (1996) An introduction to categorial data analysis. Wiley, New York

Albo MJ, Costa FG (2010) Nuptial gift-giving behaviour and male mating effort in the Neotropical spider *Paratrechalea ornata* (Trechaleidae). Anim Behav 79:1031–1036

Aloise King ED, Banks PB, Brooks RC (2013) Sexual conflict in mammals: Consequences for mating systems and life history. Mamm Rev 43:47–58

Alonso JC, Magaña M, Alvarez-Martinez JM (2012) Male display areas in exploded leks: The importance of food resources for male mating success. Behav Ecol 23:1296–1307

Alonso JC, Magaña M, Palacín C, Martín CA (2010) Correlates of male mating success in great bustard leks: The effects of age, weight, and display effort. Behav Ecol Sociobiol 64:1589–1600

Alonzo SH (2012) Sexual selection favours male parental care, when females can choose. Proc R Soc B 279:1784–1790

Asensio N, Korstjens AH, Aureli F (2009) Fissioning minimizes ranging costs in spider monkeys: A multiple-level approach. Behav Ecol Sociobiol 63:649–659

Augustine DJ, Derner JD (2013) Assessing herbivore foraging behavior with GPS collars in a semiarid grassland. Sensors 13:3711–3723

Aung M, McShea WJ, Htung S, Than A, Soe TM, Montfort S, Wemmer C (2001) Ecology and social organization of a tropical deer (*Cervus eldi thamin*). Am Soc Mammal 82:836–847

Aureli F, Schaffner CM, Asensio N, Lusseau D (2012) What is a subgroup? How socioecological factors influence interindividual distance. Behav Ecol 23:1308–1315

Aureli F, Schaffner CM, Boesch C, Bearder SK, Call J, Chapman CA, Connor R, Fiore A Di, Dunbar RIM, Henzi SP, Holekamp KE, Korstjens AH, Layton R, Lee PC, Lehmann J, Manson JH, Ramos-Fernández G, Strier KB, van Schaik CP (2008) Fission-fusion dynamics: New research frameworks. Curr Anthropol 49:627–654

Baddeley A (2010) Analysing spatial point patterns in R. CSIRO:version 4.1 URL:http://www.csiro.au/files/files/piph.pdf

Baddeley A, Gill RD (1997) Kaplan-Meier estimators of distance distributions for spatial point processes. Ann Stat 25:263–292

Barboza PS, Hartbauer DW, Hauer WE, Blake JE (2004) Polygynous mating impairs body condition and homeostasis in male reindeer (*Rangifer tarandus tarandus*). J Comp Physiol B 174:309–317

Barrette C (1991) The size of axis deer fluid groups in Wilpattu national park, Sri Lanka. Mammalia 55:207–220

Beauchamp G (2011a) Fit of aggregation models to the distribution of group sizes in Northwest Atlantic seabirds. Mar Ecol Prog Ser 425:261–268

Beauchamp G (2011b) Functional relationship between group size and population density in Northwest Atlantic seabirds. Mar Ecol Prog Ser 435:225–233

Beckers OM, Wagner WEJ (2011) Mate sampling strategy in a field cricket: Evidence for a fixed threshold strategy with last chance option. Anim Behav 81:519–527

Bercovitch FB, Berry PSM (2010) Ecological determinants of herd size in the Thornicroft's giraffe of Zambia. Afr J Ecol 48:962–971

Bercovitch FB, Berry PSM (2012) Herd composition, kinship and fission – fusion social dynamics among wild giraffe. Afr J Ecol 51:206–216

Berger J (1978) Group size, foraging, and antipredator ploys: An analysis of bighorn sheep decisions. Behav Ecol Sociobiol 4:91–99

Bergvall UA, Schäpers A, Kjellander P, Weiss A (2011) Personality and foraging decisions in fallow deer, *Dama dama*. Anim Behav 81:101–112

Bierbach D, Sassmannshausen V, Streit B, Arias-Rodriguez L, Plath M (2013) Females prefer males with superior fighting abilities but avoid sexually harassing winners when eavesdropping on male fights. Behav Ecol Sociobiol 67:675–683

Bisazza A, Vaccari G, Pilastro A (2001) Female mate choice in a mating system dominated by male sexual coercion. Behav Ecol 12:59–64

Body G, Weladji RB, Holand Ø (2012) The recursive model as a new approach to validate and monitor activity sensors. Behav Ecol Sociobiol 66:1531–1541

Borkowski JJ, Furubayashi K (1998) Seasonal and diel variation in group size among Japanese sika deer in different habitats. J Zool London 245:29–34

Bourgoin G, Garel M, Van Moorter B, Dubray D, Maillard D, Marty É, Gaillard J-M (2008) Determinants of seasonal variation in activity patterns of mouflon. Can J Zool 86:1410–1418

Bowyer RT, Rachlow JL, Stewart KM, Van Ballenberghe V (2011) Vocalizations by Alaskan moose: Female incitation of male aggression. Behav Ecol Sociobiol 65:2251–2260

Briefer EF, Farrell ME, Hayden TJ, McElligott AG (2013) Fallow deer polyandry is related to fertilization insurance. Behav Ecol Sociobiol 67:657–665

Brockmann HJ (2001) The evolution of alternative strategies and tactics. Adv Study Behav 30:1–51

Bro-Jørgensen J (2007) Reversed sexual conflict in a promiscuous antelope. Curr Biol 17:2157–2161

Bro-Jørgensen J (2011) Intra- and intersexual conflicts and cooperation in the evolution of mating strategies: Lessons learnt from Ungulates. Evol Biol 38:28–41

Brown DD, LaPoint S, Kays R, Heidrich W, Kümmeth F, Wikelski M (2012) Accelerometer-informed GPS telemetry: Reducing the trade-off between resolution and longevity. Wildl Soc Bull 36:139–146

Burnham K, Anderson D (2002) Model selection and multimodel inference: A practical information-theoretical approach. Springer, New York

Burnham KP, Anderson DR, Huyvaert KP (2011) AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons. Behav Ecol Sociobiol 65:23–35

Byers JA, Byers AA, Dunn SJ (2006) A dry summer diminishes mate search effort by pronghorn females: Evidence for a significant cost of mate search. Ethology 112:74–80

Byers JA, Hebets E, Podos J (2010) Female mate choice based upon male motor performance. Anim Behav 79:771–778

Byers JA, Moodie JD, Hall N (1994) Pronghorn females choose vigorous mates. Anim Behav 47:33–43

Byers JA, Wiseman PA, Jones L, Roffe TJ (2005) A large cost of female mate sampling in pronghorn. Am Nat 166:661–668

Cagnacci F, Boitani L, Powell RA, Boyce MS (2010) Animal ecology meets GPS-based radiotelemetry: A perfect storm of opportunities and challenges. Phil Trans R Soc B 365:2157–2162

Cameron EZ, Setsaas TH, Linklater WL (2009) Social bonds between unrelated females increase reproductive success in feral horses. Proc Nat Acad Sci USA 106:13850–13853

Cappozzo HL, Túnez JI, Cassini MH (2008) Sexual harassment and female gregariousness in the South American sea lion, *Otaria flavescens*. Naturwissenschaften 95:625–630

Carbone C, Taborsky M (1996) Mate choice or harassment avoidance? A question of female control at the lek. Behav Ecol 7:370–378

Carranza J (1995) Female attraction by males versus sites in territorial rutting red deer. Anim Behav 50:445–453

Carranza J (2000) Environmental effects on the evolution of mating systems in endotherms. In: Apollonio M, Festa-Bianchet M, Mainardi D (eds.) Vertebrate mating systems. World Scientific, Singapore, pp. 106–139

Carranza J, Garcia-Munoz AJ, Dios Vargas J de (1995) Experimental shifting from harem defence to territoriality in rutting red deer. Anim Behav 49:551–554

Carranza J, Valencia J (1999) Red deer females collect on male clumps at mating areas. Behav Ecol 10:525–532

Carter AJ, Pays O, Goldizen AW (2009) Individual variation in the relationship between vigilance and group size in eastern grey kangaroos. Behav Ecol Sociobiol 64:237–245

Caughley G (1964) Social organization and daily activity of the red kangaroo and the grey kangaroo. J Mammal 45:429–436

Charlton BD, Reby D, McComb K (2007) Female red deer prefer the roars of larger males. Biol Lett 3:382–385

Clutton-Brock TH (1989) Mammalian mating systems. Proc R Soc B 236:339–372

Clutton-Brock TH (2007) Sexual selection in males and females. Science 318:1882–1885

Clutton-Brock TH, Guinness F, Albon SD (1982) Red deer. Behavior and ecology of two sexes. Edinburgh University Press, Edinburgh

Clutton-Brock TH, McAuliffe K (2009) Female mate choice in mammals. Q Rev Biol 84:3–27

Clutton-Brock TH, McComb K (1993) Experimental tests of copying and mate choice in fallow deer (*Dama dama*). Behav Ecol 4:191–193

Clutton-Brock TH, McComb K, Deutsch JC (1996) Multiple factors affect the distribution of females in lek-breeding ungulates : A rejoinder to Carbone and Taborsky. Behav Ecol 7:373–378

Clutton-Brock TH, Parker GA (1992) Potential reproductive rates and the operation of sexual selection. Q Rev Biol 67:437–456

Clutton-Brock TH, Price OF, MacColl ADC (1992) Mate retention, harassment, and the evolution of ungulate leks. Behav Ecol 3:234–242

Clutton-Brock TH, Vincent ACJ (1991) Sexual selection and the potential reproductive rates of males and females. Nature 351:58–60

Conradt L, Krause J, Couzin ID, Roper TJ (2009) “Leading according to need” in self-organizing groups. Am Nat 173:304–12

Conradt L, Roper TJ (2000) Activity synchrony and social cohesion: A fission-fusion model. Proc R Soc B 267:2213–8

Coulombe M-L, Massé A, Côté SD (2006) Quantification and accuracy of activity data measured with VHF and GPS telemetry. Wildl Soc Bull 34:81–92

Couzin ID (2006) Behavioral Ecology : Social organization in fission – fusion societies. Curr Biol 16:169–171

Couzin ID (2009) Collective cognition in animal groups. Trends Cogn Sci 13:36–43

Couzin ID, Krause J (2003) Self-organization and collective behavior in Vertebrates. Adv Study Behav 32:1–75

Couzin ID, Laidre ME (2009) Fission – fusion populations. Curr Biol 19:633–635

Cramer JS (1999) Predictive performance of the binary logit model in unbalanced samples. Stat 48:85–94

Crawley MJ (2007) The R book. JW& S Ltd Ed., Chichester

Croft DP, James R, Krause J (2008) Exploring animal social networks. Princeton University Press, Princeton

Danchin E, Giraldeau L-A, Cézilly F (2008) Behavioural ecology. Oxford University Press, Oxford

Darwin C (1859) The origin of species by means of natural selection. John Murray, London

Darwin C (1871) The descent of man, and selection in relation to sex. John Murray, London

Djaković N (2012) Female mating strategy in reindeer (*Rangifer tarandus*): The importance of male age structure and availability in the social organisation , behaviour and mate choice. PhD thesis, Norwegian University of Life Science, Ås

Djaković N, Holand Ø, Hovland AL, Røed KH, Weladji RB, Fjeldstad E, Nieminen M (2012) Association patterns and kinship in female reindeer (*Rangifer tarandus*) during rut. Acta Ethol 15:165–171

Drury JP (2010) Immunity and mate choice: A new outlook. Anim Behav 79:539–545

Dubois F, Cézilly F (2002) Breeding success and mate retention in birds: A meta-analysis. Behav Ecol Sociobiol 52:357–364

Emlen ST, Oring LW (1977) Ecology, sexual selection, and the evolution of mating systems. Science 197:215–223

Espmark Y (1964) Rutting behaviour in reindeer (*Rangifer tarandus* L.). Anim Behav 7:159–163

Ezenwa V, Jolles A (2008) Horns honestly advertise parasite infection in male and female African buffalo. Anim Behav 75:2013–2021

Farrell ME, Briefer EF, McElligott AG (2011) Assortative mating in fallow deer reduces the strength of sexual selection. PLoS One 6:e18533

Festa-Bianchet M (1998) Condition-dependent reproductive success in bighorn ewes. Ecol Lett 1:91–94

Firk R, Stamer E, Junge W, Krieter J (2002) Automation of oestrus detection in dairy cows: A review. Livest Prod Sci 75:219–232

Fischhoff IR, Dushoff J, Sundaresan SR, Cordingley JE, Rubenstein DI (2009) Reproductive status influences group size and persistence of bonds in male plains zebra (*Equus burchelli*). Behav Ecol Sociobiol 63:1035–1043

Fishlock V, Lee PC (2013) Forest elephants: Fission–fusion and social arenas. Anim Behav 85:357–363

Focardi S, Pecchioli E (2005) Social cohesion and foraging decrease with group size in fallow deer (*Dama dama*). Behav Ecol Sociobiol 59:84–91

Forchhammer MC, Clutton-brock TH, Lindstrom J, Albon SD (2001) Climate and population density induce long-term cohort variation in a northern ungulate. J Anim Ecol 70:721–729

Fortin D, Fortin M-E, Beyer HL, Duchesne T, Courant S, Dancose K (2009) Group-size-mediated habitat selection and group fusion-fission dynamics of bison under predation risk. Ecology 90:2480–2490

Galimberti F, Boitani L, Marzetti I (2000) The frequency and costs of harassment in southern elephant seals. Ethol Ecol Evol 12:345–365

Geist V (1999) Deer of the world, their evolution, behaviour and ecology. Swan Hill Press, Shrewsbury

Geist V, Bayer M (1988) Sexual dimorphism in the Cervidae and its relation to habitat. J Zool London 214:45–53

Gerard J-F, Bideau E, Maublanc M-L, Loisel P, Marchal C (2002) Herd size in large herbivores: Encoded in the individual or emergent? Biol Bull 202:275–282

Gervasi V, Brunberg S, Swenson JE (2006) Individual-based method to measure animal activity levels: A test on brown bears. Wildl Soc Bull 34:1314–1319

Gittleman JL, Thompson SD (1988) Energy allocation in mammalian reproduction. Am Zool 28:863–875

Godde S, Humbert L, Côté SD, Réale D, Whitehead H (2013) Correcting for the impact of gregariousness in social network analyses. Anim Behav 85:553–558

Gottardi E, Tua F, Cargnelutti B, Maublanc M-L, Angibault J-M, Saïd S, Verheyden H (2010) Use of GPS activity sensors to measure active and inactive behaviours of European roe deer (*Capreolus capreolus*). Mammalia 74:355–362

Gower CN, Garrott RA, White PJ, Cherry S, Yoccoz NG (2009) Elk group size and wolf predation: A flexible strategy when faced with variable risk. In: Garrott RA, White PJ, Watson FGR (eds.) The ecology of large mammals in central Yellowstone, sixteen years of integrated field studies. Terrestria. Academic Press, New York, pp. 401:422

Griggio M, Hoi H (2011) An experiment on the function of the long-term pair bond period in the socially monogamous bearded reedling. Anim Behav 82:1329–1335

Haddadi H, King AJ, Wills AP, Fay D, Lowe JC, Morton AJ, Hailes S, Wilson AM (2011) Determining association networks in social animals: Choosing spatial–temporal criteria and sampling rates. Behav Ecol Sociobiol 65:1659–1668

Hamel S, Côté SD, Gaillard J-M, Festa-Bianchet M (2009) Individual variation in reproductive costs of reproduction : High-quality females always do better. J Anim Ecol 78:143–151

Hand DJ, Vinciotti V (2003) Local versus global models for classification problems: Fitting models where it matters. Am Stat 57:124–131

Hasegawa M, Arai E, Watanabe M, Nakamura M (2012) Female mate choice based on territory quality in barn swallows. J Ethol 30:143–150

Hawley DM, Sydenstricker KV, Kollias GV, Dhondt AA (2005) Genetic diversity predicts pathogen resistance and cell-mediated immunocompetence in house finches. Biol Lett 1:326–329

Haydon DT, Morales JM, Yott A, Jenkins DA, Rosatte R, Fryxell JM (2008) Socially informed random walks: Incorporating group dynamics into models of population spread and growth. Proc R Soc B 275:1101–1109

He H, Garcia EA (2009) Learning from imbalanced data. IEEE Trans Knowl Data Eng 21:1263–1284

Heatta MJ (2009) The mating strategy of female Svalbard reindeer (*Rangifer tarandus platyrhynchus*). MSc thesis, University of Tromsø, Tromsø

Henshaw J (1970) Consequences of travel in the rutting of reindeer and caribou (*Rangifer tarandus*). Anim Behav 18:256–258

Heurich M, Traube M, Stache A, Löttker P (2012) Calibration of remotely collected acceleration data with behavioral observations of roe deer (*Capreolus capreolus* L.). Acta Theriol 57:251–255

Hewison AJM, Vincent JP, Reby D (1998) Social organization of European roe deer. In: The European roe deer: The biology of success. Scandinavian University Press, Oslo, pp. 189–219

Hirotani A (1989) Social relationships of reindeer *Rangifer tarandus* during rut: Implications for female choice. Appl Anim Behav Sci 24:183–202

Hirotani A (1994) Dominance rank, copulatory behaviour and estimated reproductive success in male reindeer. Anim Behav 48:929–936

Hirsch BT (2011) Within-group spatial position in ring-tailed coatis: Balancing predation, feeding competition, and social competition. Behav Ecol Sociobiol 65:391–399

Hoare DJ, Couzin ID, Godin J-GJ, Krause J (2004) Context-dependent group size choice in fish. Anim Behav 67:155–164

Holand Ø, Askim KR, Røed KH, Weladji RB, Gjøstein H, Nieminen M (2007) No evidence of inbreeding avoidance in a polygynous ungulate: The reindeer (*Rangifer tarandus*). Biol Lett 3:36–39

Holand Ø, L'Italien L, Weladji RB, Djaković N, Røed KH, Hovland AL, Nieminen M (2012) Shit happens – a glimpse into males' mating tactics in a polygynous ungulate - the reindeer. Rangifer 32:65–72

Holand Ø, Røed KH, Mysterud A, Kumpula J, Nieminen M, Smith ME (2003) The effect of sex ratio and male age structure on reindeer calving. J Wildl Manage 67:25–33

Holand Ø, Weladji RB, Gjøstein H, Kumpula J, Smith ME, Nieminen M, Røed KH (2004) Reproductive effort in relation to maternal social rank in reindeer (*Rangifer tarandus*). Behav Ecol Sociobiol 57:69–76

Holand Ø, Weladji RB, Røed KH, Gjøstein H, Kumpula J, Gaillard J, Smith ME, Nieminen M (2006) Male age structure influences females' mass change during rut in a polygynous ungulate: The reindeer (*Rangifer tarandus*). Behav Ecol Sociobiol 59:682–688

Hosmer D, Lemeshow S (2000) Applied logistic regression. Wiley, New York

Huk T, Winkel W (2007) Testing the sexy son hypothesis-a research framework for empirical approaches. Behav Ecol 19:456–461

Isvaran K (2005) Variation in male mating behaviour within ungulate populations: Patterns and processes. Curr Sci 89:1192–1199

Jarman P (1974) The social organisation of antelope in relation to their ecology. Behaviour 48:215–267

Jedrzejewski W, Spaedtke H, Kamler JF, Jedrzejewska B, Stenkewitz U (2006) Group size dynamics of red deer in Białowieza Primeval Forest, Poland. J Wildl Manage 70:1054–1059

Jennions MD, Petrie M (1997) Variation in mate choice and mating preferences: A review of causes and consequences. Biol Rev 72:283–327

Johnson CN (1983) Variations in group size and composition in red and western grey kangaroos, *Macropus rufus* (Desmarest) and *M. fuliginosus* (Desmarest). Aust Wildl Res 10:25–31

Jong K de, Forsgren E, Sandvik H, Amundsen T (2012) Measuring mating competition correctly: Available evidence supports operational sex ratio theory. Behav Ecol 23:1170–1177

Juanico DEO (2009) Herding tendency as an aggregating factor in a binary mixture of social entities. Ecol Modell 220:3521–3526

Kelley JL, Morrell LJ, Inskip C, Krause J, Croft DP (2011) Predation risk shapes social networks in fission-fusion populations. PLoS One 6:e24280

King AJ, Wilson AM, Wilshin SD, Lowe JC, Haddadi H, Hailes S, Morton AJ (2012) Selfish-herd behaviour of sheep under threat. Curr Biol 22:561–562

King G, Zeng L (2001) Logistic regression in rare events data. Polit Anal 9:137–163

Kojola I (1986) Rutting behaviour in an enclosured group of wild forest reindeer. Rangifer 1:173–179

Kojola I, Nieminen M (1988) Aggression and nearest neighbour distances in female reindeer during the rut. Ethology 77:217–224

Kokko H, Jennions MD (2008) Parental investment, sexual selection and sex ratios. J Evol Biol 21:919–948

Komers PE, Birgersson B, Ekvall K (1999) Timing of estrus in fallow deer is adjusted to the age of available mates. Am Nat 153:431–436

Kozakai C, Koike S, Yamazaki K, Furubayashi K (2008) Examination of captive Japanese black bear activity using activity sensors. Mammal Study 33:115–119

Krause J, Croft DP, James R (2007) Social network theory in the behavioural sciences: Potential applications. Behav Ecol Sociobiol 62:15–27

Krause J, Krause S, Arlinghaus R, Psorakis I, Roberts S, Rutz C (2013) Reality mining of animal social systems. Trends Ecol Evol 28:541–551

Krause J, Ruxton GD (2002) Living in groups. Oxford University Press, New York.

Krop-Benesch A, Berger A, Hofer H, Heurich M (2013) Long-term measurement of roe deer (*Capreolus capreolus*) (Mammalia: Cervidae) activity using two-axis accelerometers in GPS-collars. Ital J Zool 80:69–81

Kvarnemo C, Simmons LW (2013) Polyandry as a mediator of sexual selection before and after mating. Phil Trans R Soc B 368:20120042

L'Italien L (2010) Dynamique des harems chez le renne: Influence de l'âge et du rang social du mâle. MSc thesis, Université Laval, Québec

L'Italien L, Weladji RB, Holand Ø, Røed KH, Nieminen M, Côté SD (2012) mating group size and stability in reindeer *Rangifer tarandus*: The effects of male characteristics, sex ratio and male age structure. Ethology 118:783–792

Lent P (1965) Rutting behaviour in a barren-ground caribou population. Anim Behav 13:259–264

Linklater W, Cameron E, Minot E, Stafford K (1999) Stallion harassment and the mating system of horses. Anim Behav 58:295–306

Lipetz VE, Bekoff M (1982) Group size and vigilance in pronghorns. Z Tierpsychol 58:203–216

Loe LE, Bonenfant C, Mysterud A, Severinsen T, Oritsland NA, Langvatn R, Stien A, Irvine RJ, Stenseth NC (2007) Activity pattern of arctic reindeer in a predator-free environment: No need to keep a daily rhythm. Oecologia 152:617–624

Loe LE, Irvine RJ, Bonenfant C, Stien A, Langvatn R, Albon SD, Mysterud A, Stenseth NC (2006) Testing five hypotheses of sexual segregation in an arctic ungulate. J Anim Ecol 75:485–496

Löttker P, Rummel A, Traube M, Stache A, Šustr P, Müller J, Heurich M (2009) New possibilities of observing animal behaviour from a distance using activity sensors in GPS-collars: An attempt to calibrate remotely collected activity data with direct behavioural observations in red deer *Cervus elaphus*. Wildlife Biol 15:425–434

Lovari S, Bartolommei P, Meschi F, Pezzo F (2008) Going out to mate: Excursion behaviour of female roe deer. Ethology 114:886–896

Lung MA, Childress MJ (2007) The influence of conspecifics and predation risk on the vigilance of elk (*Cervus elaphus*) in Yellowstone national park. Behav Ecol 18:12–20

Luttbeg B (1996) A comparative Bayes tactic for mate assessment and choice. Behav Ecol 7:451–460

Luttbeg B (2002) Assessing the robustness and optimality of alternative decision rules with varying assumptions. Anim Behav 63:805–814

Maalouf M (2011) Logistic regression in data analysis : An overview. Int J Data Anal Tech Strateg 3:281–299

Maalouf M, Trafalis TB (2011) Robust weighted kernel logistic regression in imbalanced and rare events data. Comput Stat Data Anal 55:168–183

Maher CR (1997) Group stability, activity budgets, and movements of pronghorn females. Southwest Nat 42:25–32

Mainguy J, Côté SD, Cardinal E, Houle M (2008) Mating tactics and mate choice in relation to age and social rank in male mountain goats. J Mammal 89:626–635

Mainguy J, Côté SD, Festa-Bianchet M, Coltman DW (2009) Father-offspring phenotypic correlations suggest intralocus sexual conflict for a fitness-linked trait in a wild sexually dimorphic mammal. Proc R Soc B 276:4067–4075

Makowicz AM, Schlupp I (2013) The direct costs of living in a sexually harassing environment. Anim Behav 85:569–577

Markusson E, Folstad I (1997) Reindeer antlers: Visual indicators of individual quality? Oecologia 110:501–507

Marshall HH, Carter AJ, Rowcliffe JM, Cowlishaw G (2012) Linking social foraging behaviour with individual time budgets and emergent group-level phenomena. Anim Behav 84:1295–1305

Martin P, Bateson P (2007) Measuring behaviour: An introductory guide. Cambridge University Press, Cambridge

Mays HL, Albrecht T, Liu M, Hill GE (2008) Female choice for genetic complementarity in birds: A review. Genetica 134:147–58

McComb K, Clutton-Brock TH (1994) Is mate choice copying or aggregation responsible for skewed distributions of females on leks? Proc R Soc B 255:13–19

McMahon CR, Bradshaw CJA (2004) Harem choice and breeding experience of female southern elephant seals influence offspring survival. Behav Ecol Sociobiol 55:349–362

Melnycky N, Weladji RB, Holand Ø, Nieminen M (2013) Scaling of antler size in reindeer (*Rangifer tarandus*): Sexual dimorphism and variability in resource allocation. J Mammal 6:1371–1379

Michelena P, Gautrais J, Gérard J-F, Bon R, Beneubourg J-L (2008) Social cohesion in groups of sheep: Effect of activity level, sex composition and group size. Appl Anim Behav Sci 112:81–93

Millspaugh J, Marzluff J (2001) Radio tracking and animal populations. Academic Press, San Diego

Moen R, Pastor J, Cohen Y (1996) Interpreting behavior from activity counters in GPS collars on moose. Alces 32:101–108

Möller LM (2012) Sociogenetic structure, kin associations and bonding in delphinids. Mol Ecol 21:745–64

Moreau M, Siebert S, Buerkert A, Schlecht E (2009) Use of a tri-axial accelerometer for automated recording and classification of goats' grazing behaviour. Appl Anim Behav Sci 119:158–170

Morrell LJ, Ruxton GD, James R (2010) Spatial positioning in the selfish herd. Behav Ecol 22:16–22

Muller MN, Thompson ME, Kahlenberg SM, Wrangham RW (2011) Sexual coercion by male chimpanzees shows that female choice may be more apparent than real. Behav Ecol Sociobiol 65:921–933

Mysterud A, Holand Ø, Røed KH, Gjøstein H, Kumpula J, Nieminen M (2003) Effects of age, density and sex ratio on reproductive effort in male reindeer (*Rangifer tarandus*). J Zool 261:341–344

Mysterud A, Langvatn R, Stenseth NC (2004) Patterns of reproductive effort in male ungulates. J Zool 264:209–215

Mysterud A, Røed KH, Holand Ø, Yoccoz NG, Nieminen M (2009) Age-related gestation length adjustment in a large iteroparous mammal at northern latitude. J Anim Ecol 78:1002–1006

Naulty F, Harty H, Hayden TJ (2013) Freedom to choose : Unconstrained mate-searching behaviour by female fallow deer (*Dama dama*). Folia Zool 62:143–154

Naylor LM, Kie JG (2004) Monitoring activity of Rocky Mountain elk using recording of accelerometers. Wildlife Biol 32:1108–1113

Nefdt RJC, Thirgood SJ (1997) Lekking, resource defense, and harassment in two subspecies of lechwe antelope. Behav Ecol 8:1–9

Oommen T, Baise LG, Vogel RM (2011) Sampling bias and class imbalance in maximum-likelihood logistic regression. Math Geosci 43:99–120

Owen-smith N, Goodall V, Fatti P (2012) Applying mixture models to derive activity states of large herbivores from movement rates obtained using GPS telemetry. Wildl Res 39:452–462

Ozoga JJ, Verme LJ (1975) Activity patterns of white-tailed deer during estrus. J Wildl Manage 39:679–683

Parra GJ, Corkeron PJ, Arnold P (2011) Grouping and fission–fusion dynamics in Australian snubfin and Indo-Pacific humpback dolphins. Anim Behav 82:1423–1433

Pays O, Benhamou S, Helder R, Gerard J-F (2007) The dynamics of group formation in large mammalian herbivores: An analysis in the European roe deer. Anim Behav 74:1429–1441

Pays O, Fortin D, Gassani J, Duchesne J (2012) Group dynamics and landscape features constrain the exploration of herds in fusion-fission societies: The case of European roe deer. PLoS One 7:e34678

Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. Ecol Modell 133:225–245

Pearson HC (2009) Influences on dusky dolphin (*Lagenorhynchus obscurus*) fission-fusion dynamics in Admiralty Bay, New Zealand. Behav Ecol Sociobiol 63:1437–1446

Pépin D, Gerard J-F (2008) Group dynamics and local population density dependence of group size in the Pyrenean chamois, *Rupicapra pyrenaica*. Anim Behav 75:361–369

Pépin D, Renaud P, Decuq F (2006) Identifying activity patterns from activity counters in ETHOSYS® collars on red deer. Appl Anim Behav Sci 96:103–114

Pereira PHC, Feitosa JLL, Medeiros D V, Ferreira BP (2013) Reef fishes foraging facilitation behavior: Increasing the access to a food resource. Acta Ethol 16:53–56

Pérez-Barbería FJ, Yearsley JM (2010) Sexual selection for fighting skills as a driver of sexual segregation in polygynous ungulates: An evolutionary model. Anim Behav 80:745–755

Proffitt KM, Gude JA, Shamhart J, King F (2012) Variations in elk aggregation patterns across a range of elk population sizes at Wall Creek, Montana. J Wildl Manage 76:847–856

R Development Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Rakotomalala R (2011) Pratique de la régression logistique: Régression logistique binaire et polytomique. version 2.0 http://eric.univ-lyon2.fr/~ricco/cours/cours/pratique_regression_logistique.pdf

Réale D, Boussès P, Chapuis J (1996) Female-biased mortality induced by male sexual harassment in a feral sheep population. Can J Zool 74:1812–1818

Reimers E (1983) Reproduction in wild reindeer in Norway. Can J Zool 61:211–217

Reimers E, Røed KH, Colman JE (2012) Persistence of vigilance and flight response behaviour in wild reindeer with varying domestic ancestry. J Evol Biol 25:1543–54

Relyea RA, Ortega IM, Demarais S (1994) Activity monitoring in mule deer : Assessing telemetry accuracy. Wildl Soc Bull 22:656–661

Richard E, Morellet N, Cargnelutti B, Angibault J-M, Vanpé C, Hewison AJM (2008) Ranging behaviour and excursions of female roe deer during the rut. Behav Processes 79:28–35

Rintamäki PT, Alatalo R, Höglund J, Lundberg A (1995) Mate sampling behaviour of black grouse females (*Tetrao tetrix*). Behav Ecol Sociobiol 37:209–215

Ripley, B D (1988) Statistical inference for spatial processes. Cambridge University Press, New York

Rodney D, Boert J (1985) Seasonal activity of the Denali caribou herd, Alaska. Rangifer 5:32–42

Røed KH (2005) Refugial origin and postglacial colonization of holarctic reindeer and caribou. Rangifer 25:19–30

Røed KH, Flagstad O, Nieminen M, Holand Ø, Dwyer MJ, Røv N, Vilà C (2008) Genetic analyses reveal independent domestication origins of Eurasian reindeer. Proc R Soc B 275:1849–55

Røed KH, Holand Ø, Gjøstein H, Hansen H (2005) Variation in male reproductive success in a wild population of reindeer. J Wildl Manage 69:1163–1170

Røed KH, Holand Ø, Mysterud A, Tverdal A, Kumpula J, Nieminen M (2007) Male phenotypic quality influences offspring sex ratio in a polygynous ungulate. Proc R Soc B 274:727–33

Røed KH, Holand Ø, Smith ME, Gjøstein H, Kumpula J, Nieminen M (2002) Reproductive success in reindeer males in a herd with varying sex ratio. Mol Ecol 11:1239–1243

Ropstad E (2000) Reproduction in female reindeer. Anim Reprod Sci 60-61:561–570

Sánchez-Prieto CB, Carranza J, Pulido FJ (2004) reproductive behavior in female Iberian red deer: Effects of aggregation and dispersion of food. J Mammal 85:761–767

Schellinck J, White T (2011) A review of attraction and repulsion models of aggregation: Methods, findings and a discussion of model validation. Ecol Modell 222:1897–1911

Schwager M, Anderson D, Butler Z, Rus D (2007) Robust classification of animal tracking data. Comput Electron Agric 56:46–59

Shuster SM, Wade MJ (2003) Mating systems and strategies. Princeton University Press, Princeton

Skogland T (1989) Comparative social organization of wild reindeer in relation to food, mates and predator avoidance. Adv Ethol 29:1:77

Stache A, Heller E, Hothorn T, Heurich M (2013) Activity patterns of European roe deer (*Capreolus capreolus*) are strongly influenced by individual behaviour. Folia Zool 62:67–75

Stephens DW, Brown JS, Ydenberg RC (2007) Foraging: Behavior and ecology. University of Chicago Press, Chicago

Stockley P, Bro-Jørgensen J (2011) Female competition and its evolutionary consequences in mammals. Biol Rev 86:341–366

Stopher K V., Nussey DH, Clutton-Brock TH, Guinness F, Morris A, Pemberton JM (2011) The red deer rut revisited: Female excursions but no evidence females move to mate with preferred males. Behav Ecol 22:808–818

Sueur C, Jacobs A, Amblard F, Petit O, King AJ (2011) How can social network analysis improve the study of primate behavior? Am J Primatol 73:703–719

Sueur C, King AJ, Conradt L, Kerth G, Lusseau D, Mettke-Hofmann C, Schaffner CM, Williams L, Zinner D, Aureli F (2011) Collective decision-making and fission-fusion dynamics: A conceptual framework. Oikos 120:1608–1617

Symonds MRE, Moussalli A (2010) A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behav Ecol Sociobiol 65:13–21

Tennenhouse EM, Weladji RB, Holand Ø, Nieminen M (2012) Timing of reproductive effort differs between young and old dominant male reindeer. Ann Zool Fennici 49:152–160

Tennenhouse EM, Weladji RB, Holand Ø, Røed KH, Nieminen M (2011) Mating group composition influences somatic costs and activity in rutting dominant male reindeer (*Rangifer tarandus*). Behav Ecol Sociobiol 65:287–295

Tobler M, Schlupp I, Plath M (2011) Costly interactions between the sexes: Combined effects of male sexual harassment and female choice? Behav Ecol 22:723–729

Tryjanowski P, Hromada M (2005) Do males of the great grey shrike, *Lanius excubitor*, trade food for extrapair copulations? Anim Behav 69:529–533

Umstätter C, Waterhouse A, Holland JP (2008) An automated sensor-based method of simple behavioural classification of sheep in extensive systems. Comput Electron Agric 64:19–26

Ungar ED, Henkin Z, Gutman M, Dolev A, Genizi A, Ganskopp DC (2005) Inference of animal activity from GPS collar data on free-ranging cattle. Rangel Ecol Manag 58:256–266

Van Moorter B, Visscher DR, Jerde CL, Frair JL, Merrill EH (2010) Identifying movement states from location data using cluster analysis. J Wildl Manage 74:588–594

van Oort BEH, Tyler NJC, Gerkema MP, Folkow L, Stokkan K-A (2007) Where clocks are redundant: Weak circadian mechanisms in reindeer living under polar photic conditions. Naturwissenschaften 94:183–194

van Oort BEH, Tyler NJC, Storeheier P, Stokkan K-A (2004) The performance and validation of a data logger for long-term determination of activity in free-ranging reindeer, *Rangifer tarandus* L.. Appl Anim Behav Sci 89:299–308

Vander Wal E, van Beest FM, Brook RK (2013) Density-dependent effects on group size are sex-specific in a gregarious ungulate. PLoS One 8:e53777

Vanpé C, Gaillard J-M, Kjellander P, Mysterud A, Magnien P, Delorme D, Van Laere G, Klein F, Liberg O, Hewison AJM (2007a) Antler size provides an honest signal of male phenotypic quality in roe deer. Am Nat 169:481–493

Vanpé C, Kjellander P, Galan M, Cosson J-F, Aulagnier S, Liberg O, Hewison AJM (2007b) Mating system, sexual dimorphism, and the opportunity for sexual selection in a territorial ungulate. Behav Ecol 19:309–316

Velando A (2004) Female control in yellow-legged gulls: Trading paternity assurance for food. Anim Behav 67:899–907

Wade MJ, Shuster SM (2004) Sexual selection: Harem size and the variance in male reproductive success. Am Nat 164:83–89

Wade MJ, Shuster SM (2010) Bateman (1948): Pioneer in the measurement of sexual selection. Heredity 105:507–508

Weir LK (2013) Male–male competition and alternative male mating tactics influence female behavior and fertility in Japanese medaka (*Oryzias latipes*). Behav Ecol Sociobiol 67:193–203

Weladji RB, Holand Ø, Almøy T (2003) Use of climatic data to assess the effect of insect harassment on the autumn weight of reindeer (*Rangifer tarandus*) calves. J Zool 260:79–85

Weladji RB, Loison A, Gaillard J-M, Holand Ø, Mysterud A, Yoccoz NG, Nieminen M, Stenseth NC (2008) Heterogeneity in individual quality overrides costs of reproduction in female reindeer. Oecologia 156:237–247

White PJ, Gower CN, Davis TL, Sheldon JW, White JR (2012) Group dynamics of Yellowstone pronghorn. J Mammal 93:1129–1138

White PJ, Proffitt KM, Lemke TO (2012) Changes in elk distribution and group sizes after wolf restoration. Am Midl Nat 167:174–187

Wiegmann DD, Angeloni LM, Seubert SM, Wade GJ (2013) Mate choice decisions by searchers. Curr Zool 59:184–199

Wiegmann DD, Seubert SM, Wade, Gordon J (2010) Mate choice and optimal search behavior: Fitness returns under the fixed sample and sequential search strategies. J Theor Biol 262:596–600

Wiley RH, Poston J (1996) Indirect mate choice, competition for mates, and coevolution of the sexes. Evolution 50:1371–1381