# Introduction

**Francesca Cagnacci, Mathieu Basille, Anne Berger, Sarah Cain Davidson, Holger Dettki, Bram van Moorter and Ferdinando Urbano**

F. Cagnacci
Biodiversity and Molecular Ecology Department, Research and Innovation Centre, Fondazione Edmund Mach, via E. Mach 1, 38010 S.Michele all'Adige, TN, Italy
e-mail: francesca.cagnacci@fmach.it

M. Basille
Fort Lauderdale Research and Education Center, University of Florida, 3205 College Avenue, Fort Lauderdale, FL 33314, USA
e-mail: basille@ase-research.org

A. Berger
Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Straße 17, D-10315 Berlin, Germany
e-mail: berger@izw-berlin.de

S. C. Davidson
Max Planck Institute for Ornithology, Am Obstberg 1, 78315 Radolfzell Germany and Department of Civil, Environmental and Geodetic Engineering, The Ohio State University, 2070 Neil Ave, Columbus, OH 43210, USA
e-mail: sdavidson@orn.mpg.de

H. Dettki
Umeå Center for Wireless Remote Animal Monitoring (UC-WRAM),
Department of Wildlife, Fish, and Environmental Studies,
Swedish University of Agricultural Sciences (SLU), Skogsmarksgränd,
SE-901 83 Umeå, Sweden
e-mail: holger.dettki@slu.se

B. van Moorter
Norwegian Institute for Nature Research (NINA), Høgskoleringen 9,
7034 Trondheim, Norway
e-mail: Bram.van.moorter@gmail.com

F. Urbano
Università Iuav di Venezia, Santa Croce 191 Tolentini, 30135 Venice, Italy
e-mail: ferdi.urbano@gmail.com

## The Long and Winding Road to Movement Ecology: Let's put Things in Order, First

Animal movement is a proximate response to local environmental conditions, such as climate, chemophysical parameters, resources, and the presence of mates or predators; in other words, the context in which animals are born, survive, reproduce, and die. At the same time, movement is the result of complex evolutionary mechanisms, uniquely blending gene expression into physiological and behavioural responses. Animal movement has been described as the glue that ties ecological processes together (Turchin 1998; Nathan et al. 2008) and as an important mechanistic link between ecology and evolution (Cagnacci et al. 2010). Moreover, the increasing concern for rapidly changing ecosystems, under climate and land use change, brings with it an urgent and heightened interest in the capacity of animals to respond to such changes. An exciting perspective offered by animal movement is the possibility to understand this phenomenon at different spatial scales and levels of organisation, from individuals, to populations over landscapes and to the distribution range of species (Mueller and Fagan 2008). Such stimulating theoretical framework and widely relevant applied perspective clearly justify the ever-increasing interest in animal tracking studies, especially those based on animal-borne data sets, i.e. those obtained through the deployment of tracking units on individual animals. The innovation in Global Positioning System (GPS) technology, combined with systems for remote data transfer, has particularly favored GPS-based animal telemetry to become a standard in wildlife tracking (Cagnacci et al. 2010).

The advancement of a movement ecology theoretical framework has been paralleled by technological progress that allows ecologists to obtain a huge amount and diversity of empirical animal movement data sets. However, this fast-growing and expanding process has not been followed by an equally rapid development of procedures to manage and integrate animal movement data sets, thus leaving a gap between the acquisition of data and the overarching scientific questions they have the potential to address (Urbano et al. 2010).

The ideal objective of any movement ecology study is rooted in relevant ecological questions that can contribute to theory and inform conservation and management actions. For example, a study on natal dispersal can help in identifying barriers to gene flow, or an analysis on environmental characteristics affecting population performance can support decisions on protection areas and conservation corridors. Whatever the question, the second necessary step should be the evaluation of the appropriate methodology, and specifically deciding whether individual marking with Global Navigation Satellite Systems (GNSS) devices such as GPS (or other sensors) is the most effective and informative approach to pursuing the final goal (see discussion in Hebblewhite and Haydon 2010). If the answer is yes, GPS units should be ideally deployed according to an *a priori* sampling design (e.g. number, sex ratio and spacing of sampled individuals) and relevant information on marked individuals should be collected (e.g. age and body mass), as well as on the population of reference and the environmental context (e.g. population density, presence of competitor species/predators, human activity pressure). Collection of GPS locations of marked animals should finally flow, and can be complemented with environmental parameters (e.g. weather, habitat types and vegetation indexes based on remote sensing).

The reason for this book comes at this stage of the scientific method: what to do with these data? How to handle, manage, store and retrieve them, and how to eventually feed them to analysis tools such as statistics packages or Geographic Information Systems (GIS) and test scientific hypotheses? These operations, which might be assumed to be secondary compared to the overarching goal of answering scientific questions, can instead become overwhelming and hamper the efficiency and consistency of the whole process. Animal ecologists, wildlife biologists and managers, to whom this book is mainly addressed, are rarely exposed to the basics of computer science in their training, and may select common tools such as spreadsheets, or operate in a 'flat-files' fashion. However, the quantity and complexity of GPS and other bio-logged data require a proper software architecture to be fully exploited and not wasted or, worse, misused. This book is a guide to manage and process wildlife tracking data with an advanced software platform based on a spatial database. It is neither a manual on database programming nor on wildlife tracking; instead, it aims to fill the gap between the state-of-the-art knowledge on data management and its application to wildlife tracking data. This problem-solving oriented guide focuses on how to extract the most from GPS animal tracking data, while preventing error propagation and optimizing the performance of analysis. Using databases to manage tracking data implies a considerable initial effort for those who are not familiar with these tools; however, the time spent learning will pay off in time saved for the management and processing of the data, and in the quality of the analysis and final results. Moreover, once a consistent database structure is built to code the storing and management of data, more familiar tools can be used as interfaces.

Another important advantage of using a structured and consistent software platform for management of wildlife tracking data is the ever-increasing importance of cooperative projects and data sharing. Deploying sensors on animals can be expensive, both in terms of capture logistics and actual cost of tracking devices, and implies some amount of stress for the marked animals. These costs pose financial and even moral incentives to maximize the effective use of animal-borne data. Moreover, many large-scale questions, such as evaluating the effects of climate and land use change on animal distributions, can be properly addressed only by studying multiple populations, or by integrating data from several species or time periods. Data requirements for such studies are virtually impossible for any single research institution to meet, and can only be achieved by cooperative research and data sharing (Whitlock 2011). Spatial databases are essential tools to assure data standards and interoperability, along with a safe multiple users operational environment.

In this manual, a sequential set of exercises guides readers step by step through setting up a spatial database to manage GPS tracking data together with ancillary

**Fig. 1** F10, one of the sample animals studied in this book: an adult female roe deer (*Capreolus capreolus*), wearing a Vectronic Aerospace GmbH GPS-Plus yellow collar

information (e.g. characteristics of animals and sensors), environmental layers (e.g. land cover, altitude and vegetation index), and data from other bio-logged sensors (e.g. accelerometers). Data from five GPS and one activity sensors deployed on roe deer (*Capreolus capreolus*; Fig. 1) in the Italian Alps (Fig. 2) are used as the test data set[1]. In each chapter, data management problems are contextualized in relation to issues that scientists and wildlife managers face when they work with tracking data. The goal of the book is to illustrate conceptual and technical solutions and to offer a practical example of general validity for most of the animal species target of movement ecology studies.

The world of animal-borne telemetry, or bio-logging, comprises information not limited to location data of marked animals. The first natural extension is 3D coordinates, where two-dimensional coordinates (i.e. latitude and longitude) are complemented by altitude for avian species or depth for marine species. Other increasingly used parameters are triaxial acceleration (measured by accelerometers) and magnetic bearing (measured by magnetometers), which can integrate the movement trajectories with information on the type of body movement and therefore activity of animals. A diversity of other variables can be measured by bio-logging devices, providing information on the internal state of animals (biomechanics, energetics), and also their external state, i.e. the environment (pressure, chemophysical parameters) (see Ropert-Coudert and Wilson 2005 for an

[1] The test data set (trackingDB_datasets.zip) is available at http://extras.springer.com.

**Fig. 2** Typical habitat characterizing the alpine area where the sample animals were tracked by means of GPS collars: open alpine pastures and mixed woodlands

early review on this topic). When justified by the scientific questions underpinning the empirical study, the use of multi-sensor platforms deployed on individual animals can be a powerful tool to obtain a complex and diversified picture of the animal in its environment. With the miniaturization of technology and adaptation of devices to an increasing number of species, multi-sensor platforms and their resulting data sets are likely to be even more common in the future (see, for example, deployment of miniature video cameras on birds, Rutz and Troscianko 2013). The presence of multiple sources of information (i.e. the sensors) fitted on the same animal does not represent a challenge *per se*, if each type of information is consistently linked to the animal and integrated with the other data sources. However, this requires a dedicated and consistent management structure, an expansion of the database requirements for tracking data alone. Although this book mainly deals with spatial data obtained from individual animals tracked with GPS, examples are provided to show how most of the conceptual background can be exported to other sensors or to multi-sensor platforms.

PostgreSQL and its spatial extension PostGIS are the proposed software platform to build the wildlife tracking database. This spatial database will allow management of virtually any volume of data from wildlife GPS tracking in a unique repository, while assuring data integrity and consistency, avoidance of error propagation, and limiting data duplication and redundancy. Moreover, this software platform offers the opportunity of automation of otherwise very time-consuming processes, such as, for example, the association between GPS locations, animals and the environmental variables, and high performance of data retrieval, which can become cumbersome when data sets are very large. Databases are designed for being remotely accessed and used concurrently by multiple users, assuring data integrity and security (i.e. ability to define a diversified access policy). Last but not least, once the storing and management of data is coded in a robust database, data can still be visualized and accessed with common tools such as office packages (e.g. spreadsheets, database front-ends) or GIS desktop. Therefore, after the initial effort of designing the structure and populating the database, most users should be able to easily access the data.

The proposed SQL code can be directly used to create an operational test database and, at a later stage, as a template to create a more customized data management system. R is used as the reference statistical environment. The examples, illustrated in a MS Windows environment, can be easily adapted to the commonly used operating systems, such as Mas OS X and Linux, which are all supported by both PostgreSQL/PostGIS and R.

Chapter 1 is a review of the opportunities and challenges that are posed by GPS tracking from a data management perspective, with a brief introduction to (spatial) databases, which are proposed as the best available tool for wildlife tracking data management. Chapter 2 guides readers through the initial creation of a (PostgreSQL) database to store GPS data. Chapter 3 shows how to integrate ancillary information on both animals and GPS devices. Chapter 4 presents a solution to associating GPS positions to animals in an automated fashion. Chapter 5 explores the features offered by PostGIS, the spatial extension of PostgreSQL, transforming the coordinates received from the GPS sensors into 'spatial objects', i.e. points embedding spatial characteristics. Chapter 6 uses the spatial tools of the database to add a set of spatial ancillary information in the form of vectors (points, lines, and polygons) and rasters and to connect these data sets with GPS tracking data. Chapter 7 focuses on the integration of ancillary information that captures the temporal variability of the environment, with a practical example based on the remote sensing image time series of Normalized Difference Vegetation Index (NDVI), a proxy of vegetation biomass. Chapter 8 is dedicated to the detection and management of erroneous location and potential outliers. Chapter 9 introduces methods for representing and analysing tracking data (e.g. trajectories, home range polygons, probability surfaces) using database functions. Chapter 10 explains how to connect R, a powerful software environment for statistical computing and graphics, with the database and describes a set of typical tracking data analyses performed with R. Chapter 11 introduces a more direct and efficient approach to integrate R and PostgreSQL: Pl/R. Chapter 12 discusses the integration of GPS tracking data with other bio-logged sensors, with a practical example based on accelerometers. Chapter 13 gives an overview of data standards and software interoperability, with special reference to data sharing.

Apart from Chaps. 1, 12 and 13, which introduce more general and theoretical topics, each chapter proposes exercises that are developed sequentially. Although it is possible to perform an exercise from a later chapter on its own by restoring the dump of the whole database[2], we recommend completing the exercises in the order in which they are presented in the book to have a full understanding of the whole process. We also suggest acquiring some general understanding of databases and spatial databases before reading this book.

## References

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

Hebblewhite M, Haydon DT (2010) Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. Philos Trans R Soc B 365:2303–2312

Mueller T, Fagan WF (2008) Search and navigation in dynamic environments—from individual behaviours to population distributions. Oikos 117:654–664

Nathan R, Getz WM, Revilla E, Holyoak M, Kadmon R, Saltz D, Smouse PE (2008) A movement ecology paradigm for unifying organismal movement research. Proc Natl Acad Sci U.S.A. 105:19052–19059

Ropert-Coudert Y, Wilson R (2005) Trends and perspectives in animal-attached remote sensing. Front Ecol Environ 3(8):437–444

Rutz C, Troscianko J (2013) Programmable, miniature video-loggers for deployment on wild birds and other wildlife. Methods Ecol Evol 4:114–122

Turchin P (1998) Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants. Sinauer Associates, Inc., Sunderland

Urbano F, Cagnacci F, Calenge C, Cameron A, Neteler M (2010) Wildlife tracking data management: a new vision. Philos Trans R Soc B 365:2177–2186

Whitlock MC (2011) Data archiving in ecology and evolution—best practices. Trends Ecol Evol 26(2):61–65. doi:10.1016/j.tree.2010.11.006

[2] The dump (trackingDB_dump.zip) is available at http://extras.springer.com.