



**HAL**  
open science

## The exploratory analysis of autocorrelation in animal-movement studies

S. Dray, M. Royer-Carenzi, C. Calenge

► **To cite this version:**

S. Dray, M. Royer-Carenzi, C. Calenge. The exploratory analysis of autocorrelation in animal-movement studies. *Ecological Research*, 2010, 25 (3), pp.673-681. 10.1007/s11284-010-0701-7. hal-00539473

**HAL Id: hal-00539473**

**<https://hal.science/hal-00539473>**

Submitted on 18 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 The exploratory analysis of autocorrelation in animal  
2 movement studies

3 Stéphane Dray<sup>1</sup>, Manuela Royer-Carenzi<sup>2</sup> and Clément Calenge<sup>3</sup>

4 <sup>1</sup>*Université de Lyon ; université Lyon 1 ; CNRS ; UMR 5558, Laboratoire de Biométrie  
5 et Biologie Evolutive, 43 boulevard du 11 novembre 1918, Villeurbanne F-69622, France.*

6 <sup>2</sup>*LATP, CNRS-UMR 6632, Evolution biologique et modélisation, case 5, Université de  
7 Provence, Place Victor Hugo, 13331 Marseille Cedex 3, France.*

8 <sup>3</sup>*Direction des études et de la recherche, Office national de la chasse et de la faune  
9 sauvage. Saint Benoist, 78610 Auffargis, France.*

10 Corresponding author:

11 S. Dray ([dray@biomserv.univ-lyon1.fr](mailto:dray@biomserv.univ-lyon1.fr))

12 Tel: +33 4 72 43 27 57

13 Fax: +33 4 72 43 13 88

14

15       **Abstract:**

16       Studies of animal movements has been popularized for many large and shy  
17       species by the increasing use of radio telemetry methods (VHF and GPS tech-  
18       nologies). Data are collected with high sampling frequency, and consist of  
19       successive observations of the position of an individual animal. The statistical  
20       analysis of such data poses several problems due to the lack of independence  
21       of successive observations. However, the statistical description of the tempo-  
22       ral autocorrelation between successive steps is rarely performed by ecologists  
23       studying the patterns of animals movements. The aim of this paper is to warn  
24       ecologists against the consequences of failing to consider this aspect. We discuss  
25       the various issues related to analysing autocorrelated data, and show how the  
26       exploratory analysis of autocorrelation can both reveal important biological in-  
27       sights, and help to improve the accuracy of movement models. We suggest some  
28       tools that can be used to measure, test and adjust for temporal autocorrelation.  
29       A short ecological illustration is presented.

30       *Keywords: autocorrelation function; GPS; independence test; radio telemetry;*  
31       *permutation test*

## 32 1 Introduction

33 Radio telemetry *has become increasingly popular as a convenient means of studying move-*  
34 *ments of wildlife* (Dunn and Gipson 1977) and it can be used to collect data that consist of  
35 successive observations of the locations of an individual animal. The recent development  
36 of GPS technology makes it possible to collect relocation data automatically, at short,  
37 regular intervals. However, as the time interval decreases, the dependence between succes-  
38 sive observations usually increases so that the problem of autocorrelation must be taken  
39 into considered in the analysis. Autocorrelation can be defined as a property of random  
40 variables, which means that values from samples taken near to each other tend to be ei-  
41 ther more similar (positive autocorrelation) or less similar (negative autocorrelation) than  
42 would be expected from a random arrangement. This nearness can be defined in space  
43 (spatial autocorrelation), or in time (temporal autocorrelation). The main difference is  
44 that an observation is only influenced by past values in temporal autocorrelation whereas  
45 spatial dependence extends in all directions. Autocorrelation is often seen as an obsta-  
46 cle, as it interferes with standard statistical hypothesis testing (Legendre 1993), and many  
47 approaches have been developed to eliminate it by restricting data prior to analysis. How-  
48 ever, autocorrelation is often an intrinsic property of biological data, so that eliminating  
49 it could reduce the relevance of ecological studies. In the context of animal movement,  
50 which is by definition a non-independent phenomenon, De Solla et al. (1999) summarized  
51 the situation clearly: *Animals typically move in a non-random fashion, and thus ecolo-*  
52 *gists are frequently faced with strongly autocorrelated data sets, particularly when frequent*  
53 *observations are collected using radiotelemetry.*

54 Solow (1990) distinguished two approaches in the literature on animal movements: one  
55 focusing on relocations, and the other on the movements between relocations. The question  
56 of autocorrelation must be (and indeed has been) considered in different ways in these two  
57 approaches. Any temporal autocorrelation between the coordinates of successive reloca-  
58 tions must be taken into account when the analysis is based on relocations. For instance,  
59 in the context of home range estimation, this problem has been discussed in a number of

60 papers (e.g., Dunn and Gipson 1977, Swihart and Slade 1985a;b, De Solla et al. 1999), and  
61 so we will not consider it any further here. On the other hand, the analysis of movements  
62 between relocations raises the question of the temporal autocorrelation between succes-  
63 sive steps (i.e. the displacement between two successive relocations). The development of  
64 models of realistic animal movement is a very active field of research at present (Patterson  
65 et al. 2008, Schick et al. 2008). The aim of these studies is to understand why, how and  
66 where animals move, and to see how this could affect the spatial dynamics at higher levels  
67 of organization (population, community, etc). Recent methodological developments have  
68 improved the ability of models to handle the biological complexity of animal movements  
69 (Nathan et al. 2008), and to account for multiple behavior modes (e.g. searching, resting,  
70 etc), characteristics of landscape (e.g. patchiness) or environmental variations (e.g. spatial  
71 distribution of resources). The various internal (e.g. behavioral) or external (e.g. environ-  
72 mental) factors influencing organisms could induce temporal autocorrelation in movement  
73 data. For instance, autocorrelation in steps could be due to the animal's switching between  
74 different movement strategies. Hence, autocorrelation is usually considered as an emergent  
75 property in a modelling framework and its exploratory analysis is rarely performed by  
76 authors.

77 The interest of the empirical study of autocorrelation has been affirmed in the recent  
78 special issue of PNAS on movement ecology by Wittemyer et al. (2008). They stated  
79 that *the role of autocorrelation, at different temporal scales, in the movement pathways of*  
80 *animals is an important understudied phenomenon that is critical for predictive modeling*  
81 *of population spatial properties*. Following this idea, the aim of this paper is to demonstrate  
82 that movement ecology could benefit from a deeper exploratory analysis of the empirical  
83 data, with a special emphasis on temporal autocorrelation. We present the various problems  
84 linked to the analysis of autocorrelated data, and show how the study of autocorrelation  
85 can reveal important biological insights, help to develop new biological hypotheses, and  
86 thus make it possible to improve the accuracy of movement models. We suggest some tools  
87 that can be used to measure, test and adjust for temporal autocorrelation. An ecological  
88 illustration is presented.

## 89 2 Description of movement paths

90 The analysis of movements involves recording the path of an individual through space and  
91 time. Animal movement is a continuous process, but its observation using radio telemetry  
92 technologies leads to a discrete representation. The data gathered consist of a sequence  
93 of successive relocations  $(x_i, y_i)$  of an individual collected at time  $i$ . Steps are defined  
94 as the displacement between two successive relocations. As this displacement is usually  
95 represented as a straight-line, a path is represented as a series of straight-lines. Following  
96 Marsh and Jones (1988), several descriptors (Calenge et al. 2009) can be associated with  
97 each step to quantify animal movements (see Fig. 1). For example, the  $n$ -th step can be  
98 characterized by:

- 99 • its length ( $d_n$ )
- 100 • the increment in the X and Y directions ( $\delta x_n$  and  $\delta y_n$ )
- 101 • the absolute angle between the step and the X direction ( $\alpha_n$ )
- 102 • the relative (or turning) angle between the  $n$ -th step and the previous one ( $\rho_n$ )
- 103 • the squared net displacement after  $n$  consecutive steps ( $R_n^2$ )

104 The parameters  $d_n$  and  $\rho_n$  have been widely used since the early work of Siniff and Jessen  
105 (1969), while  $R_n^2$  has been popularized by Kareiva and Shigesada (1983). GPS collars are  
106 often programmed to collect regular data, so that we assume from here on that the time  
107 lag between two relocations is constant. Despite improvements in this technology, GPS  
108 telemetry can still result in missing data in the form of failed location attempts. If this  
109 happens, data are still considered to be regular, as the steps connecting a relocation to a  
110 missing value are missing, but still characterized by the same time interval as the observed  
111 steps (see Calenge et al. 2009). The presence of missing data can be due for example to  
112 the behaviour of the animal or to the habitat structure (e.g., Frair et al. 2004, Graves and  
113 Waller 2006, Bourgoin et al. 2009). It is then important to deal with missing observations

114 when analysing movement data because their structure could be related to important bio-  
115 logical variables and thus gives some information about movement characteristics. In this  
116 context, Calenge et al. (2009) suggests that *the analysis of the patterns of missing values*  
117 *should be a part of trajectory analysis* (p. 36). As missing data are a key component of  
118 GPS data, we propose several procedures that have the ability to deal with them when  
119 analysing autocorrelation structures.

120 *Illustration* — Fig. 2a shows an example of a movement path. The data set contains  
121 the 336 relocations (extending over one week, with a recording interval of 30 minutes) of a  
122 female brown bear monitored using a GPS collar in July 2004 in Sweden.

### 123 **3 Autocorrelation and the analysis of movement paths**

#### 124 3.1 Definition of theoretical models

125 Understanding how animals move is often based on a confrontation of the observed data  
126 with theoretical models. The observed distribution of some of the parameters defined  
127 above may be compared to the distribution expected under the assumption of the theoret-  
128 ical model. These comparisons could be performed using graphical tools (e.g., Kareiva and  
129 Shigesada 1983) or traditional testing procedures. The random walk (RW) and the corre-  
130 lated random walk (CRW) are the models most often used (e.g., Marsh and Jones 1988).  
131 RW corresponds to a succession of random steps, whereas the CRW model assumes that  
132 the movement of animals exhibits directional persistence (“correlated”). In a mathematical  
133 context, the CRW model differs from the RW model, because it assumes a unimodal and  
134 symmetric distribution (about zero) of the turning angles whereas the RW model assumes  
135 a uniform distribution. Both models also require the independence of successive turning  
136 angles and step lengths. Turchin (1998, p. 135) stated that *CRW formulation assumes that*  
137 *the move durations, speeds, and turning angles are not serially correlated, an assumption*  
138 *that should be tested statistically.* This advice has sometimes been followed in the liter-  
139 ature (e.g., Bergman et al. 2000, Nolet and Mooij 2002, Revilla et al. 2004, Bailey and  
140 Thompson 2006, Dumont et al. 2007). For instance, Root and Kareiva (1984) tested the

141 autocorrelation for distances and turning angles using Pearson's correlation coefficients. If  
142 no tests of independence are performed, the probability that a model (e.g. RW, CRW)  
143 will be rejected when it is false (i.e. statistical power) is lower than expected, because it is  
144 assumed *a priori* that independence assumptions are met.

### 145 3.2 Computation of mean squared net displacement and related tools

146 Kareiva and Shigesada (1983) proposed an explicit formula of the expected mean squared  
147 net displacement  $E(R_n^2)$  for (C)RW models. They insisted on the notion of independence  
148 of parameters (e.g. p. 235, the word "independent" is italicized) to obtain a correct  
149 derivation of  $E(R_n^2)$ . The algorithm proposed by McCulloch and Cain (1989) to compute  
150 the variance in squared displacement  $V(R_n^2)$  for a CRW, and the expected value of the  $\Delta$   
151 statistic presented by Marsh and Jones (1988) to distinguish between the CRW and the  
152 biased random walk (BRW) models are also based on the assumption of independence of  
153 some parameters. Marsh and Jones (1988, p. 129) stated that the information given in  
154 their paper can be used *to construct a simple test to distinguish between the two models*  
155 but pointed out that *we must, of course, have other reasons for restricting consideration*  
156 *to just these two types of model*. One of these reasons is that the underlying assumptions  
157 of independence have been confirmed. In this context, using these tools without checking  
158 the independence of parameters (see e.g., Odendaal et al. 1989, Firlie et al. 1998, Morales  
159 and Ellner 2002) makes no sense.

### 160 3.3 Standard hypothesis tests on movement descriptors

161 The various descriptors (see the previous section) are routinely used in standard statis-  
162 tical procedures (e.g. ANOVA, t-test for linear descriptors; Rayleigh test of uniformity  
163 for angular descriptors) to test hypotheses concerning animal movements. Usually, these  
164 procedures require the independence of observations. If data are positively autocorrelated,  
165 classical tests are biased because *computed statistics are too often declared significant un-*  
166 *der the null hypothesis* (Legendre 1993, p. 1660). Hence, the hypothesis of independence  
167 (i.e., absence of autocorrelation) should be tested and not rejected before using standard



168 hypothesis tests on movement descriptors.

169 In his note on the analysis of angular data, Cain (1989, p. 1542) stated that *statistical*  
170 *tests [...] are based on the assumption that observed angles are independent, a condition*  
171 *that may not be satisfied when multiple angles are recorded from a single individual.* He  
172 conducted a survey of the literature and showed that independence tests are rarely per-  
173 formed before hypothesis tests on angular data, and concluded that *a lack of independence*  
174 *among successive angles not only violated hypothesis test assumptions, it also has important*  
175 *movement consequences: organisms that alternate right and left turns, for example, will*  
176 *displace farther than those whose turning angles are independent.* A short review of the  
177 recent literature demonstrated that some authors did test independence (e.g., Fortin et al.  
178 2005, Bailey and Thompson 2006), but many others did not check the independence as-  
179 sumptions before carrying out standard hypothesis tests (e.g., Bowne et al. 1999, Cardona  
180 et al. 2005, Lorch et al. 2005, Parks et al. 2006). More surprisingly, several authors (e.g.,  
181 Bergman et al. 2000, Austin et al. 2004, Lancaster et al. 2006) performed independence  
182 tests (usually to check CRW assumptions), rejected the hypothesis of independence and  
183 then used wrongly standard tests on autocorrelated data.

184 Testing for the independence of path descriptors is essential, and probably the first step  
185 that should be carried out when analyzing movement data. This can be done by means of  
186 several procedures.

## 187 **4 Independence testing procedures**

188 Several approaches have been used in the literature to test the independence of descriptors.  
189 Among the different descriptors, angular data require a particular attention due to their  
190 nature and must be analyzed using appropriate circular statistics (Cain 1989).

### 191 4.1 The autocorrelation function

192 In the literature, some authors (Bergman et al. 2000, Austin et al. 2004, Bailey and Thomp-  
193 son 2006), following the recommendations of Turchin (1998), used autocorrelation functions  
194 (ACF, Diggle 1990), a standard tool for the analysis of time series which consists of plot-

195 ting the values of autocorrelation measured for different time lags against the corresponding  
 196 lags. For a real-valued random variable  $X$  taking values  $x_1, \dots, x_N$ , the sample autocorre-  
 197 lation function for a given time lag  $h$  is:

$$\hat{\rho}(h) = \frac{\sum_{i=1}^{N-h} z_i \cdot z_{i+h}}{\sum_{i=1}^N z_i^2} \text{ where } z_i = x_i - \frac{1}{N} \sum_{i=1}^N x_i.$$

198 For a significance level  $\alpha$ , we conclude that data are not independent if  $\sqrt{n}|\hat{\rho}(h)|$  exceeds  
 199 the standard gaussian  $(1 - \frac{\alpha}{2})$ -quantile. Note that the autocorrelation function for lag 1  
 200 (i.e.  $\hat{\rho}(1)$ ) is sufficient to test for independence (Diggle 1990). For the first order lag (i.e.,  
 201  $h = 1$ ), we have:

$$\begin{aligned} \hat{\rho}(1) &= \frac{\sum_{i=1}^{N-1} z_i \cdot z_{i+1}}{\sum_{i=1}^N z_i^2} \\ &= 1 - \frac{1}{2N\hat{\sigma}^2} \left( z_1^2 + z_N^2 + \sum_{i=1}^{N-1} (z_i - z_{i+1})^2 \right) \end{aligned}$$

202 where  $\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N z_i^2$ .

203 The foregoing equations allow us to link the ACF to the two methods generally used to  
 204 measure the autocorrelation. The first method considers the products  $z_i \cdot z_{i+1}$  (Wald and  
 205 Wolfowitz 1943), while the second considers the local differences  $(z_{i+1} - z_i)^2$  (von Neumann  
 206 et al. 1941, von Neumann 1941). In this paper, we will focus on the second approach which  
 207 has been implemented in the adehabitat package (Calenge 2006) for the R software.

208 ACF is not designed to handle missing data. Consequently, the previous testing proce-  
 209 dure cannot be used for data sets with missing observations. Fortunately, similar tools have  
 210 been developed in the context of spatial analysis, particularly the correlogram (Legendre  
 211 and Legendre 1998, p. 714), which can easily be modified to accept missing values. In the  
 212 following sections, we propose statistics and procedures adapted to movement data with

213 missing observations for linear and angular descriptors.

## 214 4.2 Linear descriptors

215 Among the different linear descriptors, tests of independence are often used for the length  
 216 of the step. Several authors have used the Pearson product-moment correlation coefficient  
 217 and its associated  $t$ -test (Nolet and Mooij 2002, Banks and Yasenak 2003, Lancaster et al.  
 218 2006) to evaluate the significance of the correlation of the variable of interest (e.g. the  
 219 length) and a time-shifted version of itself. Others have preferred to use Spearman's rank  
 220 correlation coefficient (Cain et al. 1991, Dumont et al. 2007) on the same data. As explained  
 221 below, standard tests (including  $t$ -test) assume independence of observations, and this  
 222 kind of approach violates this basic assumption. A correct way to test for independence is  
 223 provided by permutation procedure. For a path with  $N$  steps, a measure of autocorrelation  
 224 (von Neumann 1941, von Neumann et al. 1941) of step lengths is given by:

$$S_L = \frac{1}{N_d} \sum_{n=1}^{N-1} (d_{n+1} - d_n)^2$$

225 where  $N_d = N - 1$  is equal to the number of summed squared differences. Our presen-  
 226 tation focused on step lengths, but the method can also be used for other linear descriptors  
 227 (e.g.,  $\delta x_n$ ,  $\delta y_n$ ). The principles of the randomization procedure are as follows:

- 228 (i). Compute a reference value of  $S_L$  using original data.
- 229 (ii). Randomly permute the values of the step lengths, and recompute the  $S_L$  statistic.  
 230 This operation is repeated a number of times (e.g., 999), to provide a set of values of  
 231 the statistic under the null hypothesis of independence ( $H_0$ ).
- 232 (iii). Compare the observed statistic to the distribution containing the values obtained by  
 233 permutation as well as the reference value; compute the associated probability and  
 234 take the appropriate statistical decision (reject or fail to reject the null hypothesis of  
 235 independence).

236 Depending on the alternative hypothesis ( $H_1$ ), the third step will involve different pro-  
 237 cedures. If the alternative hypothesis is the presence of positive autocorrelation, using a  
 238 significance level equal to  $\alpha$ , a p-value is estimated as: (number of random values equal to  
 239 or less than the observed value + 1)/(number of permutations + 1). The null hypothesis  
 240 is rejected if the p-value is below the threshold of significance ( $\alpha$ ). If the alternative hy-  
 241 pothesis is the presence of negative autocorrelation, a p-value is estimated as: (number of  
 242 random values equal to or greater than the observed value + 1)/(number of permutations  
 243 + 1). Again, the null hypothesis is rejected if the p-value is below the threshold of signifi-  
 244 cance ( $\alpha$ ). Lastly, if the alternative hypothesis is the presence of autocorrelation (positive  
 245 or negative), a two-sided test can be used. The estimation of the p-value is equivalent to  
 246 the second procedure described above, except that random and observed values are first  
 247 centered (using the average of random values), and then transformed into their absolute  
 248 values.

249 In the case of missing data, the computation of  $S_L$  is restricted to the  $N_d$  pairs of  
 250 successive observed data (i.e. if  $d_n$  is missing,  $(d_n - d_{n-1})^2$  and  $(d_{n+1} - d_n)^2$  are not  
 251 computed) and only observed data are permuted (i.e. the structure of the missing data is  
 252 kept constant under permutation).

253 Our procedure can also be used to construct a correlogram by plotting the statistic  
 254  $S_L(h) = \frac{1}{N_{d(h)}} \sum_{n=1}^{N-h} (d_{n+h} - d_n)^2$  against the time lag  $h$ . The significance of  $S_L(h)$  is  
 255 assessed using the permutation procedure, as described above. Since tests are performed  
 256 for various time lags, results could be corrected using a multiple adjustment method such  
 257 as Bonferroni or Holm's sequential procedure (Legendre and Legendre 1998).

258 *Illustration* — A correlogram (Fig. 2b) computed on the bear step length data indicated  
 259 a significant positive autocorrelation for the three first lags, and also for several lags between  
 260 14 and 27 (7-13h30) and lag 43 (21h30). This seems to reflect a recurrent periodic structure  
 261 related to diel movement patterns.

### 4.3 Angular descriptors

In the literature, angular autocorrelation has been analyzed using both qualitative and quantitative measurements (Turchin 1998). There have been a few attempts to use quantitative measures: Lancaster et al. (2006) used the circular coefficient of correlation of Jammalamadaka and Sengupta (2001) while Nolet and Mooij (2002) used the coefficient of Jupp and Mardia (1980). Tests based on qualitative measurements involve transforming angles as right (R) or left (L) deviations from the previous direction for turning angles, or from the X axis for absolute angles (Cain 1989). Then one can use a runs test on the sequence of R and L (e.g., Bailey and Thompson 2006). An alternative approach is to construct the contingency table with the four possible values of successive steps (LL, RR, LR, RL) as entries. Independence is then tested using a  $\chi^2$  test with one degree of freedom (e.g., Cain et al. 1991, Banks and Yassenak 2003). Schaeffer and Messier (1997) transformed left and right deviations into 0 and 1, and then computed an ACF on this binary variable (see also Bergman et al. 2000). Transforming angular data into qualitative information undoubtedly leads to a loss of information. Moreover, we do not think that the R-L transformation, which considers only two directions, is really suitable for the analysis of movement data. Indeed, if an animal moves in a given direction with slight deviations to the left and then to the right, this would lead to a negative association (LR) but we would assume a positive autocorrelation as the movement is always in the same main direction. On the other hand, if an animal moves in given direction and then in the opposite with slight deviations to the left for both steps, this would lead to a positive association (LL) but we would assume a negative autocorrelation as the main direction has completely changed.

If we consider the turning angles, it seems natural to assume that positive autocorrelation means that the direction taken for the  $n$ -th step is similar to the previous one, while negative autocorrelation means that the direction of the  $n$ -th step is opposite to that of the previous step. Hence, it would be preferable to use a qualitative measurement that considers the four possible directions corresponding to different values of turning angles (Fig. 3): turn right (R,  $-\pi/2$ ), turn left (L,  $\pi/2$ ), go forwards (F, 0) or go backwards (B,  $\pi$

290 or  $-\pi$ ). It makes sense to consider the sequences LL, RR, FF or BB as positive associations,  
 291 and LR, RL, FB or BF as negative associations.

292 As an alternative, we propose to use a quantitative coefficient (the chord distance) to  
 293 measure angular autocorrelation (Fig. 3b). Chord distance is low when two successive  
 294 angles are similar. In the case of sequences LR, RL, FB or BF, the distance reaches its  
 295 maximum, and is equal to 2. For a sequence of  $N$  turning angles, we propose measuring  
 296 the autocorrelation as the sum of the squared chord distances:

$$S_A = \frac{1}{N_d} \sum_{n=1}^{N-1} [(\cos \rho_{n+1} - \cos \rho_n)^2 + (\sin \rho_{n+1} - \sin \rho_n)^2] = \frac{2}{N_d} \sum_{n=1}^{N-1} 1 - \cos(\rho_{n+1} - \rho_n)$$

297 Like the linear descriptors, the angular autocorrelation can be tested using the permu-  
 298 tation procedures described above. A correlogram can also be computed.

299 *Illustration* — Correlograms of absolute and turning angles computed on the bear data  
 300 did not highlight any autocorrelation structure, and are not reported here.

## 301 5 Understanding autocorrelation and detrending

302 Autocorrelation is a very general property of ecological variables. Unfortunately, it is also  
 303 an obstacle in ecological studies as it violates the basic assumptions of standard statistical  
 304 hypothesis testing. Autocorrelation is usually viewed as a form of bias, and a common  
 305 procedure is to eliminate it before analysing the data by subsampling (Turchin 1998, p.  
 306 130). However, we do not think that subsampling of relocations or steps should be routinely  
 307 used before the analysis of movement paths. The first reason for this is that computing  
 308 some descriptors requires two relocations (e.g., distance) or two steps (turning angle),  
 309 and in such circumstances, subsampling will reduce the accuracy of the estimates. The  
 310 second reason is that autocorrelation is often the consequence of ecological processes. This  
 311 is particularly true for long-time surveys obtained with GPS monitoring, for which the  
 312 animal's "switching behaviour" is often responsible of the presence of autocorrelation. In  
 313 some cases, this autocorrelation can be generated by a process unrelated to the focus issue  
 314 (e.g., the mechanistic aspects of the locomotion process), which would justify eliminating it

315 prior to the analysis. However, we stress that the preliminary analysis of the autocorrelation  
316 and identifying its causes are essential if we are to gain a better understanding of the  
317 patterns of animal movements.

318 Legendre (1993) stated that spatial structures observed in ecological communities can  
319 arise from two independent processes. Environmental factors that influence species dis-  
320 tributions are spatially structured and, as a result of an indirect process, communities of  
321 species are also spatially structured (induced spatial dependence). Spatial autocorrelation  
322 can also be created directly at the community level as a result of contagious biotic processes.  
323 By analogy, two processes can be considered to be a potential cause of temporal autocorre-  
324 lation of successive steps (Martin et al. 2008). If movement parameters are dependent on  
325 local environmental conditions, then environmental variability can induce temporal auto-  
326 correlation of movement descriptors (Turchin 1998, p. 135). Here, "environment" is used in  
327 a very broad sense, and includes both abiotic (e.g., altitude) and biotic factors (e.g., distri-  
328 bution of preys and predators). Temporal autocorrelation can also be the consequence of  
329 internal (individual) processes (e.g., circadian activity). When the origins of autocorrela-  
330 tion have been identified, it can then be eliminated before carrying out the other analyses.  
331 This requires tools to describe temporal patterns explicitly in order to detrend the data,  
332 and perform subsequent analyses on the residual (or detrended) data. Several methods  
333 exist to detrend time series including the variate difference method, Fourier series (Witte-  
334 myer et al. 2008) or harmonic regression (see for instance Legendre and Legendre 1998,  
335 Ch. 12 for details). Tools developed for the analysis of spatial autocorrelation (e.g., Dray  
336 et al. 2006) can also be useful. The reader could consult Frair et al. (2005) for an example  
337 of detrending of animal movement data. Detrending of angular data is more complicated,  
338 and involves the use of circular regression methods (Fisher 1996, Ch. 6).

339 *Illustration* — The daily pattern of step lengths is represented in Fig. 2c, and is  
340 characterized by two peaks of activity corresponding to long step lengths. This circadian  
341 pattern with two peaks is common, and has been demonstrated for several species (Aschoff  
342 1966). Hence, this circadian pattern must be taken into account before attempting any  
343 other analysis of step lengths.

344 We modelled the temporal constraint as a linear graph with 336 nodes, and computed  
345 Moran's eigenvector maps (Dray et al. 2006) for the associated weighting matrix. This  
346 approach can be used to construct a set of predictors in order to model an autocorrelated  
347 signal in ecological data (Peres-Neto 2006). Detrending was performed using the procedure  
348 proposed in Griffith and Peres-Neto (2006) with a significance level equal to 0.05. The  
349 correlogram computed on detrended data (Fig. 2d) demonstrated that the autocorrelation  
350 structure corresponding to the circadian rhythm has been removed. The correlogram for  
351 the trend (Fig. 2e) corresponds to a periodic pattern (the period is around 10-14h). The  
352 detrended data can then be used for subsequent analysis.

## 353 6 Concluding remarks

354 In these days of state-space and hierarchical Bayesian modelling (Patterson et al. 2008,  
355 Schick et al. 2008), our approach may look rather old-fashioned. What is the point of data  
356 exploration when models of realistic movements are available? Answers to this question  
357 can be found in the old debate initiated by Tukey (1977; 1980) about exploratory and  
358 confirmatory analyses. Data exploration allows one to extract what the data says, without  
359 any *a priori* hypothesis. It can be useful for generating new questions/ideas, improving  
360 the sampling design or guiding the choice of subsequent analyses. In the field of animal  
361 movement, autocorrelation is an important property of the data that could reveal infor-  
362 mation about how animals move. Patterns revealed by its exploratory analysis help to  
363 generate new hypotheses about the behavioral processes related to animal mobility. These  
364 hypotheses can then be tested by introducing new parameters into animal movement mod-  
365 els. For instance, our example suggests that a model integrating the time of the day (e.g.  
366 Preisler et al. 2004) should be used to analyse brown bear movement data. Exploratory  
367 analysis and modelling of animal movements should therefore not be opposed, but rather  
368 considered as two complementary approaches. The reader could consult Brillinger et al.  
369 (2004) for a nice example of an exploratory analysis of animal movement data.

370 We hope that this paper will help to improve the way we tackle the question of auto-



371 correlation in animal movement analysis. An appropriate approach would consist of the  
372 following steps:

- 373 • measure and test autocorrelation using appropriate procedures;
- 374 • if autocorrelation is detected, its structure must be studied as it can reflect ecological  
375 processes of interest.

376 Then, depending on the type of subsequent analysis, one can:

- 377 • detrend the data so that autocorrelation (corresponding to processes analyzed in the  
378 previous step) is eliminated;
- 379 • use detrended data in the analysis. Standard statistical tools can now be used, as  
380 the observations are now independent.

381 or

- 382 • incorporate new parameters in movement models that account for the hypothesised  
383 processes.

384 In order to help ecologists, functions that can be used to store and manage animal move-  
385 ment paths and to measure and test autocorrelation have been included in the `adehabitat`  
386 package (Calenge 2006) for the R software. The reader could consult Calenge et al. (2009)  
387 for more details about these functions. Appendix 1 provides a complete example corre-  
388 sponding to the brown bear illustration used in this paper.

389 **Acknowledgments:** We are grateful to Jon Swenson and to the Scandinavian Brown  
390 Bear Research Project for kindly providing the data. Financial support has been provided  
391 by the ANR (project Mobilité ANR-05-BDIV-008) and the ONCFS. We would like to thank  
392 Jean-Michel Gaillard, Jodie Martin and Dominique Allainé for their comments on earlier  
393 drafts of this paper.

## 394 References

395 Aschoff J (1966) Circadian activity pattern with two peaks. *Ecology* 47:657–662

- 396 Austin D, Bowen W, McMillan J (2004) Intraspecific variation in movement patterns:  
397 modeling individual behaviour in a large marine predator. *Oikos* 105:15–30
- 398 Bailey H, Thompson P (2006) Quantitative analysis of the bottlenose dolphin movement  
399 patterns and their relationship with foraging. *J Anim Ecol* 75:456–465
- 400 Banks J, Yasenak C (2003) Effects of plot vegetation diversity and spatial scale on *Coc-*  
401 *cinella septempunctata* movement in the absence of prey. *Entomol Exp Appl* 108:197–204
- 402 Bergman C, Schaeffer J, Luttich S (2000) Caribou movement as a correlated random walk.  
403 *Oecologia* 123:364–374
- 404 Bourgoin G, Garel M, Dubray D, Maillard D, Gaillard J (2009) What determines global  
405 positioning system fix success when monitoring free-ranging mouflon? *European Journal*  
406 *of Wildlife Research* 55:603–613
- 407 Bowne D, Peles J, Barrett G (1999) Effects of landscape spatial structure on movement  
408 patterns of the hispid cotton rat (*Sigmodon hispidus*). *Landsc Ecol* 14:53–65
- 409 Brillinger D, Preisler H, Ager A, Kie J (2004) An exploratory data analysis (EDA) of the  
410 paths of moving animals. *J Stat Plan Infer* 122:43–63
- 411 Cain M (1989) The analysis of angular data in ecological field studies. *Ecology* 70:1540–  
412 1543
- 413 Cain M, Carson W, Root R (1991) Long-term suppression of insect herbivores increases  
414 the production and growth of *Solidago atissima* rhizomes. *Oecologia* 88:251–257
- 415 Calenge C (2006) The package "adehabitat" for the R software: a tool for the analysis of  
416 space and habitat use by animals. *Ecol Model* 197:516–519
- 417 Calenge C, Dray S, Royer-Carenzi M (2009) The concept of animals trajectories from a  
418 data analysis perspective. *Ecol Inform* 4:34–41

- 419 Cardona L, Revelles M, Carreras C, San Félix M, Gazo M, Aguilar A (2005) Western  
420 Mediterranean immature loggerhead turtles: habitat use in spring and summer assessed  
421 through satellite tracking and aerial surveys. *Mar Biol* 147:583–591
- 422 De Solla S, Bonduriansky R, Brooks R (1999) Eliminating autocorrelation reduces biolog-  
423 ical relevance of home range estimates. *J Anim Ecol* 68:221–234
- 424 Diggle P (1990) Time series: a biostatistical introduction. Oxford University Press
- 425 Dray S, Legendre P, Peres-Neto P (2006) Spatial modeling: a comprehensive framework  
426 for principal coordinate analysis of neighbor matrices (PCNM). *Ecol Model* 196:483–493
- 427 Dumont C, Himmelman J, Robinson S (2007) Random movement pattern of the sea urchin  
428 *Strongylocentrotus droebachiensis*. *J Exp Mar Biol Ecol* 340:80–89
- 429 Dunn J, Gipson P (1977) Analysis of radio telemetry data in studies of home range. *Bio-*  
430 *metrics* 33:85–101
- 431 Firlé S, Bommarco R, Ekbom B, Natiello M (1998) The influence of movement and resting  
432 behavior on the range of three carabid beetles. *Ecology* 79:2113–2122
- 433 Fisher N (1996) Statistical analysis of circular data. Cambridge University Press, Cam-  
434 bridge
- 435 Fortin D, Morales J, Boyce MS (2005) Elk winter foraging at fine scale in Yellowstone  
436 National Park. *Oecologia* 145:335–346
- 437 Frair J, Merrill E, Visscher D, Fortin D, Beyer H, Morales J (2005) Scales of movement by  
438 elk (*Cervus elaphus*) in response to heterogeneity in forage resources and predation risk.  
439 *Landsc Ecol* 20:273–287
- 440 Frair J, Nielsen SE, Merrill E, Lele S, Boyce MS, Munro R, Stenhouse G, Beyer H (2004)  
441 Removing GPS collar bias in habitat selection studies. *J Appl Ecol* 41:201–212
- 442 Graves T, Waller J (2006) Understanding the causes of missed global positioning system  
443 telemetry fixes. *J Wildl Manag* 70:844–851

- 444 Griffith DA, Peres-Neto P (2006) Spatial modeling in ecology: the flexibility of eigenfunc-  
445 tion spatial analyses. *Ecology* 87:2603–2613
- 446 Jammalamadaka S, Sengupta A (2001) *Topics in Circular Statistics*. World Scientific Press,  
447 Singapore
- 448 Jupp P, Mardia K (1980) A general correlation coefficient for directional data and related  
449 regression problems. *Biometrika* 67:163–173
- 450 Kareiva P, Shigesada N (1983) Analyzing insect movement as a correlated random walk.  
451 *Oecologia* 56:234–238
- 452 Lancaster J, Buffin-Bélanger T, Reid I, Rice S (2006) Flow- and substratum-mediated  
453 movement by a stream insect. *Freshw Biol* 51:1053–1069
- 454 Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* 74:1659–  
455 1673
- 456 Legendre P, Legendre L (1998) *Numerical Ecology*. Elsevier Science, Amsterdam, 2nd  
457 edition
- 458 Lorch P, Sword G, Gwynne D, Anderson G (2005) Radiotelemetry reveals differences in  
459 individual movement patterns between outbreak and non-outbreak Mormon cricket pop-  
460 ulations. *Ecol Entomol* 30:548–555
- 461 Marsh L, Jones R (1988) The form and consequences of random walk movement models.  
462 *J Theor Biol* 133:113–131
- 463 Martin J, Calenge C, Quenette P, Allainé D (2008) Importance of movement constraints  
464 in habitat selection studies. *Ecol Model* 213:257–262
- 465 McCulloch C, Cain M (1989) Analyzing discrete movement data as a correlated random  
466 walk. *Ecology* 70:383–388
- 467 Morales J, Ellner S (2002) Scaling up animal movements in heterogeneous landscapes: the  
468 importance of behavior. *Ecology* 83:2240–2247

- 469 Nathan R, Getz W, Revilla E, Holyoak M, Kadmon R, Saltz D, Smouse P (2008) A  
470 movement ecology paradigm for unifying organismal movement research. *Proc Natl*  
471 *Acad Sci* 105:19052–19059
- 472 Nolet B, Mooij W (2002) Search paths of swans foraging on spatially autocorrelated tubers.  
473 *J Anim Ecol* 71:451–462
- 474 Odendaal F, Turchin P, Stermitz F (1989) Influence of host-plant density and male ha-  
475 rassment on the distribution of female *Euphydras anicia* (Nymphalidae). *Oecologia*  
476 78:283–288
- 477 Parks E, Derocher A, Lunn N (2006) Seasonal and annual movement patterns of polar  
478 bears on the sea ice of Hudson Bay. *Can J Zool* 84:1281–1294
- 479 Patterson T, Thomas L, Wilcox C, Ovaskainen O, Matthiopoulos J (2008) State-space  
480 models of individual animal movement. *Trends Ecol Evol* 23:87–94
- 481 Peres-Neto P (2006) A unified strategy for estimating and controlling spatial, temporal  
482 and phylogenetic autocorrelation in ecological models. *Oecologia Brasiliensis* 10:105–119
- 483 Preisler H, Ager A, Johnson B, Kie J (2004) Modeling animal movements using stochastic  
484 differential equations. *Environmetrics* 15:643–657
- 485 Revilla E, Wiegand T, Palomares F, Ferreras P, Delibes M (2004) Effects of matrix hetero-  
486 geneity on animal dispersal: from individual behavior to metapopulation-level parame-  
487 ters. *Am Nat* 164:130–153
- 488 Root R, Kareiva P (1984) The search of resources by cabbage butterflies (*Pieris rapae*):  
489 Ecological consequences and adaptive significance of markovian movements in a patchy  
490 environment. *Ecology* 65:147–165
- 491 Schaeffer J, Messier F (1997) Footedness in foraging muskoxen *Ovibos moschatus*. *Acta*  
492 *Theriol* 42:335–338

- 493 Schick R, Loarie S, Colchero F, Best B, Boustany A, Conde D, Halpin P, Joppa L, McClellan  
494 C, Clark J (2008) Understanding movement data and movement processes: current and  
495 emerging directions. *Ecol Lett* 11:1338–1350
- 496 Siniff D, Jessen C (1969) A simulation model of animal movement patterns. *Adv Ecol Res*  
497 6:185–219
- 498 Solow AR (1990) A note on the statistical properties of animal locations. *J Math Biol*  
499 29:189–193
- 500 Swihart R, Slade N (1985a) Testing for independence of observations in animal movements.  
501 *Ecology* 66:1176–1184
- 502 Swihart RK, Slade NA (1985b) Influence of sampling interval on estimates of home-range  
503 size. *J Wild Manage* 49:1019–1025
- 504 Tukey J (1977) *Exploratory data analysis*. Addison-Wesley, Reading
- 505 Tukey J (1980) We need both exploratory and confirmatory. *Am Stat* 34:23–25
- 506 Turchin P (1998) *Quantitative analysis of movement: measuring and modeling population*  
507 *redistribution in plants and animals*. Sinauer Associates, Sunderland, MA
- 508 von Neumann J (1941) Distribution of the ratio of the mean square successive difference  
509 to the variance. *Ann Math Stat* 12:367–395
- 510 von Neumann J, Kent R, Bellinson HR, Hart B (1941) The mean square successive differ-  
511 ence. *Ann Math Stat* 12:153–162
- 512 Wald A, Wolfowitz J (1943) An exact test for randomness in the non-parametric case based  
513 on serial correlation. *Ann Math Stat* 14:378–388
- 514 Wittemyer G, Polansky L, Douglas-Hamilton I, Getz W (2008) Disentangling the effects  
515 of forage, social rank, and risk on movement autocorrelation of elephants using Fourier  
516 and wavelet analyses. *Proc Natl Acad Sci* 105:19108–19113

517 **7 Figure legend**

518 FIG. 1: Parameters used to describe steps. The  $n$ -th step is associated to its length ( $d_n$ ),  
 519 the increments in the X and Y directions ( $\delta x_n$  and  $\delta y_n$ ), the absolute angle ( $\alpha_n$ ), the  
 520 turning angle ( $\rho_n$ ) and the squared net displacement ( $R_n^2$ ).

521 FIG. 2: Analysis of brown bear data. (a) representation of the movement path. (b)  
 522 correlogram of step lengths ( $S_L(h)$ ). The gray area corresponds to a 95% confidence  
 523 interval obtained by permutation procedure (999 permutations). The gray dotted line  
 524 depicts the median of the values of the statistic computed on permuted data. Empty  
 525 circles represent non-significant autocorrelation, while black squares correspond to  
 526 significant values. (c) Boxplots of the daily pattern of step lengths. The line with white  
 527 circles corresponds to average values. Correlogram for the detrended step lengths (d) and  
 528 the trend (e). The trend corresponds to the predictions of the linear model with step  
 529 length as response variable and 9 Moran's eigenvectors maps as explanatory variables;  
 530 detrended data correspond to the residuals of this model.

531 FIG. 3: Using the chord distance to measure the autocorrelation for the turning angles.  
 532 (a) A path of 4 steps encompasses 3 successive turning angles  $\rho_1$ ,  $\rho_2$  and  $\rho_3$ . (b) Small  
 533 chord length corresponds to positive association between successive angles  $\rho_1$  and  $\rho_2$  and  
 534 large chord length corresponds to "negative association" between successive angles  $\rho_2$  and  
 535  $\rho_3$ .

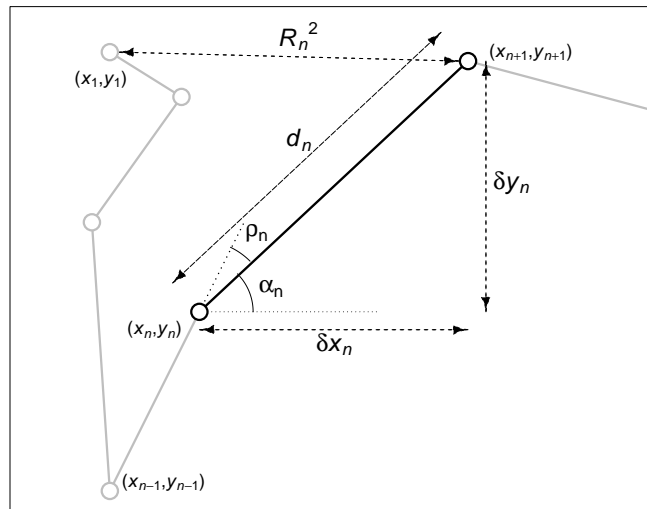


Fig. 1



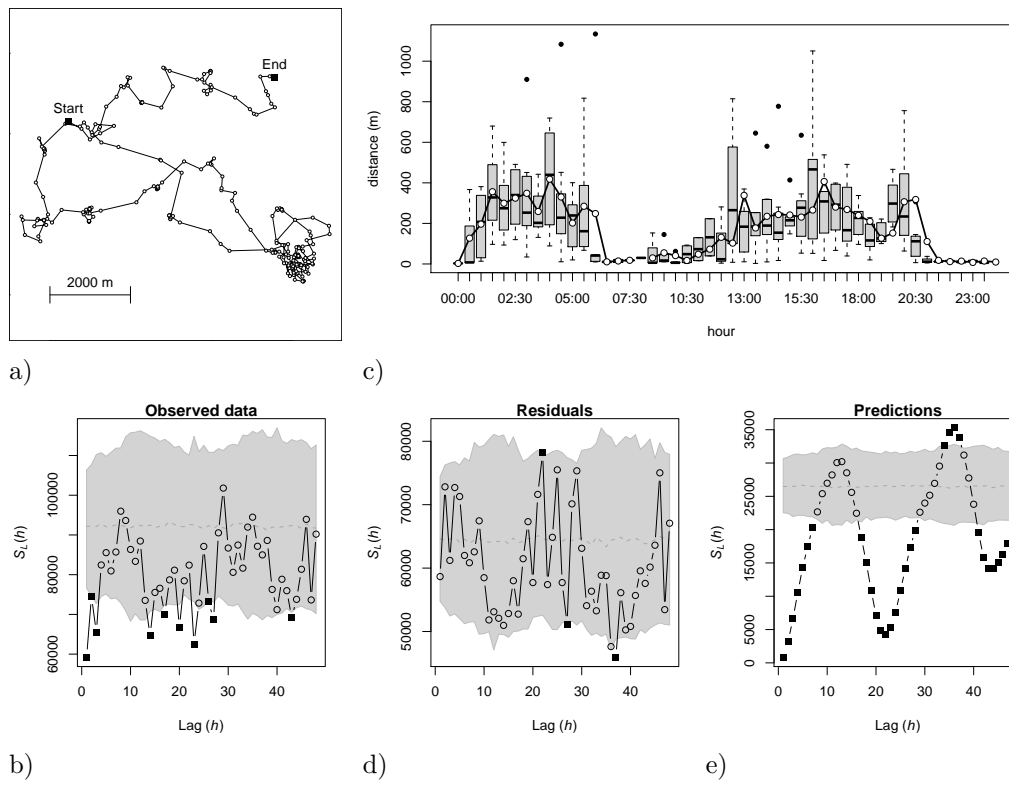
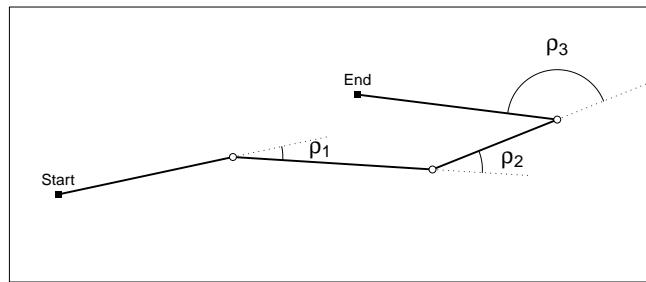
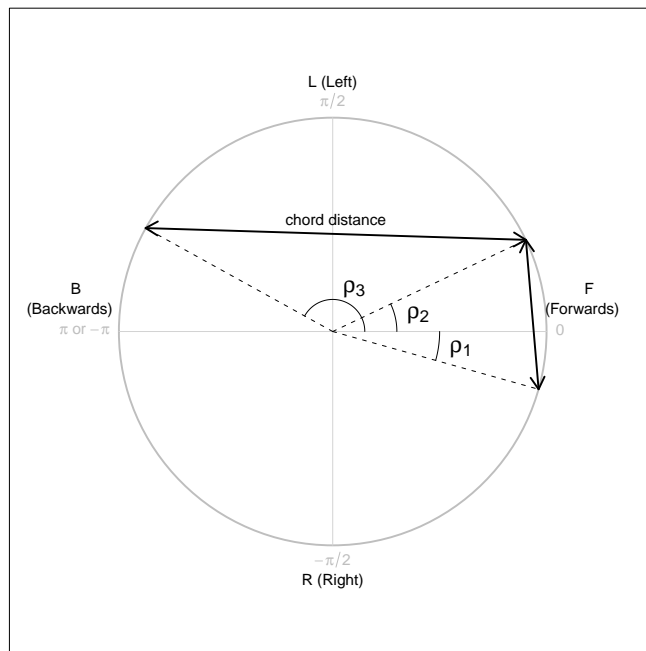


Fig. 2



a)



b)

Fig. 3