

1
2
3
4
5
6
7
8

9
10

A Review of Quantitative Methods for Movement Data

Jed A. Long*, Trisalyn A. Nelson

*Spatial Pattern Analysis & Research Lab, Department of Geography, University of
Victoria, Victoria, BC, Canada*

*corresponding author: jlong@uvic.ca

Pre-print of published version.

Reference:

Long, JA and TA Nelson. 2013. A review of quantitative methods for movement data. *International Journal of Geographical Information Science*. 27(2). 292-318.

DOI:

<http://dx.doi.org/10.1080/13658816.2012.682578>

Disclaimer:

The PDF document is a copy of the final version of this manuscript that was subsequently accepted by the journal for publication. The paper has been through peer review, but it has not been subject to any additional copy-editing or journal specific formatting (so will look different from the final version of record, which may be accessed following the DOI above depending on your access situation).

11 **Abstract**

12 The collection, visualization, and analysis of movement data is at the forefront of
13 geographic information science research. Movement data are generally collected by
14 recording an object's spatial location (e.g., XY coordinates) at discrete time intervals.
15 Methods for extracting useful information, for example space-time patterns, from these
16 increasingly large and detailed datasets have lagged behind the technology for generating
17 them. In this article we review existing quantitative methods for analyzing movement
18 data. The objective of this article is to provide a synthesis of the existing literature on
19 quantitative analysis of movement data while identifying those techniques that have merit
20 with novel datasets. Seven classes of methods are identified: 1) time geography, 2) path
21 descriptors, 3) similarity indices, 4) pattern and cluster methods, 5) individual-group
22 dynamics, 6) spatial field methods, and 7) spatial range methods. Challenges routinely
23 faced in quantitative analysis of movement data include difficulties with handling space
24 and time attributes together, representing time in GIS, and using classic statistical testing
25 procedures with space-time movement data. Areas for future research include:
26 investigating equivalent distance comparisons in space and time, measuring interactions
27 between moving objects, development of predictive frameworks for movement data,
28 integrating movement data with existing geographic layers, and incorporating theory
29 from time geography into movement models. In conclusion, quantitative analysis of
30 movement data is an active research area with tremendous opportunity for new
31 developments and methods.

32 **1 – Introduction**

33 The study of movement in geographic information science (GISci) has followed a
34 similar trajectory to the discipline of geography, whereby early work relied heavily on
35 qualitative methods. In the 1960's and 70's the discipline of geography experienced a
36 quantitative revolution whereby theory and methods were developed for explaining how
37 place and space could be modeled as quantitative entities. The quantitative revolution
38 produced developments in statistical methods designed specifically for spatial data, for
39 instance spatial autocorrelation measures (Cliff and Ord 1973). Only later in the
40 quantitative revolution did theoretical frameworks for quantitative analysis of movement
41 emerge; most notably Hägerstrand's (1970) time geography. As the quantitative
42 revolution in geography sputtered in the late 1970's (Johnston 1997) Hägerstrand's ideas
43 were primarily used as context for examining human behavior (e.g., Parkes and Thrift
44 1975, Pred 1981), rather than as an analytical toolkit for quantitative research. An
45 exception is the work of Lenntorp (1976) and Burns (1979), which represent seminal
46 pieces using time geography in quantitative analysis.

47 In the 1990's, triggered by the development of geographic information systems
48 (GIS), quantitative analysis again moved to the forefront of the geographic literature
49 (Sheppard 2001). The term geographic information science (GISci) was coined to refer
50 collectively to the science behind the collection, storage, representation, and analysis of
51 geographic datasets (Goodchild 1992). The term amalgamated those interested in the
52 study of geographic information including geographers, computer scientists, and
53 statisticians. As technologies for recording the paths of moving objects have evolved
54 (e.g., video, cell-phone, and GPS tracking) contemporary GIScientists have found new

55 opportunities for quantitative analysis using time geography with GISci (e.g., Miller
56 1991, Kwan 1998). Other quantitative methods for analyzing movement have stemmed
57 from geography's strong legacy in spatial point pattern analysis (e.g., Gao et al. 2010), as
58 movement data are commonly represented by a sequence of points. Computational
59 geometry has played a leading role in recent advances in analyzing movement data (e.g.,
60 Laube et al. 2005). As well, methods for representing movement data using areal data
61 formats, for example polygons (Downs and Horner 2009) or fields (Downs 2010), remain
62 ongoing research areas. The study of movement is of interest in many applications
63 outside of GISci, for example wildlife ecology (Nathan et al. 2008), urban planning
64 (Drewe 2005), and military applications (Wells 1981). Further, the study of movement
65 has a long history in physics. Even Hägerstrand's time geography was strongly
66 influenced by the ideas of physicists from the early 20th century (Rose 1977, Hallin
67 1991). For example, the diagram of the space-time cone from time geography can be
68 clearly related to the past and future light-cones used in Einstein's relativity.

69 Movement is a complex process that operates through both space and time.
70 Representing the temporal dimension in geographic studies has presented a challenge for
71 GISci to move beyond static (map-based) representations of space (Chrisman 1998,
72 Laube et al. 2007). Despite notable advances at incorporating temporal dynamics in
73 GISci (e.g., Pultar et al. 2010), integrating the study of space and time remains at the
74 forefront of GISci research, as evidenced by the special symposium on space-time
75 integration in GISci at the 2011 annual meeting of the Association of American
76 Geographers. How to effectively integrate time into the quantitative analysis of
77 movement, specifically movement data stored in a GIS, is at the core of this review.

78 The growth of spatial methods for quantitative analysis of movement data has
79 been facilitated by developments in movement databases that now provide efficient
80 methods for storing, indexing, and querying movement data (Güting and Schneider
81 2005). Despite the large body of existing literature on the topic of moving object
82 databases, it remains an active area of research as new tools (e.g., Güting et al. 2010a)
83 and applications (e.g., Jensen et al. 2010) continue to develop. Data visualization
84 methods have developed alongside these readily available movement databases; in GISci
85 this practice is termed geovisualization (Dykes et al. 2005). Given the sheer volume of
86 data often contained in movement databases, geovisualization can be a powerful tool for
87 identifying patterns in movement databases – a process referred to as visual analytics
88 (Thomas and Cook 2005). A complete treatment of either of these topics is beyond the
89 scope of this review, and we restrict the contents of this review to, as the title suggests,
90 those methods for analyzing movement data that are *quantitative* in nature. We would
91 point those interested in more information on movement databases to the comprehensive
92 book by Güting and Schneider (2005) and a recent special issue on data management for
93 mobile services (*VLDB Journal*, **20**(5), Güting and Mamoulis 2011). For those interested
94 in more information on visual analytics for movement data we refer readers to Andrienko
95 and Andrienko (2007), and to the special issue from *IJGIS* entitled geospatial visual
96 analytics: focus on time (*IJGIS*, **24**(10), Andrienko et al. 2010).

97 The objective of this review is to provide an unbiased evaluation of the usefulness
98 and shortcomings of existing quantitative methods for movement data, while highlighting
99 techniques that have particular merit with emerging movement datasets. Challenges to the
100 development and application of quantitative methods with movement data are identified

101 in an attempt to locate avenues for future research. An outline of this article is as follows;
102 section 2 contains a brief introduction to the properties of movement data, and how
103 movement data is typically represented in a GIS. In section 3 we review the existing
104 literature on quantitative analysis of movement data separated into seven classes of
105 methods: 1) time geography, 2) path descriptors, 3) similarity indices, 4) pattern and
106 cluster methods, 5) individual-group dynamics, 6) spatial field methods, and 7) spatial
107 range methods. With section 4 we provide a discussion of the challenges routinely faced
108 in GISci when analyzing movement data and, what we feel are, some future directions for
109 quantitative movement analysis. Lastly, we close with some conclusions.

110 **2 – Movement Data**

111 Movement is a continuous process that operates in both the spatial and temporal
112 domains. Movement data are used to represent the continuous process of movement for
113 geographical analysis. Due to existing geospatial data collection and storage techniques,
114 movement data are most commonly represented as a collection of spatial point objects
115 with time stored as an attribute. A more formal definition of movement data is the
116 collection $\{M_t\}$ of $t = 1 \dots n$ ordered records each comprised of the triple $\langle ID, S, T \rangle$,
117 where ID is a unique object identifier, S are spatial coordinates, and T a sequential (non-
118 duplicated) time-stamp (Hornsby and Egenhofer 2002). A number of terms are used
119 synonymously for movement data (see Table 1); here we use the term *path* to represent
120 the ordered sequence of records portraying individual/object movement, the term *fix*
121 when discussing a single record from a path, and the term *movement database* to describe
122 a collection of paths. The term *movement data* is used in broader contexts when

123 discussing the study of movement, to refer generally to fixes, paths, and movement
124 databases.

125 <approximate location Table 1>

126 While movement data have historically been collected using a variety of
127 techniques, most current acquisition schemes use some form of wireless sensor (e.g.,
128 GPS, cellular phone records, radio telemetry). Calenge et al. (2009) identify two types of
129 sampling commonly employed in the collection of movement data – regular and
130 irregular. Regular paths are those where fixes are acquired at an even temporal interval,
131 for example recording one fix per minute. Irregular paths are those where fixes are
132 acquired at unequal temporal intervals, for example paths collected from cell phone call
133 records. The term *granularity* is used to refer to the resolution of a path (Hornsby and
134 Egenhofer 2002). Finer granularities are associated with frequent sampling intervals, and
135 provide a detailed representation of movement. Conversely, coarser granularities
136 correspond to sparse sampling and less-detailed representation of movement.

137 Technological developments now facilitate finer sampling intervals in movement paths
138 (e.g., 1 fix / second), and movement data can be used to represent a (near) continuous
139 movement path (Laube et al. 2007). However, these sensor-specific sampling designs
140 may not be suitable for all analysis questions, requiring the use of re-sampling (up- or
141 down-sampling) to fit a given research need (see Turchin 1998, and Hornsby and
142 Egenhofer 2002 for a more thorough discussion of changing granularity).

143 Spaccapietra et al. (2008) present an alternative view of movement data
144 granularity, defining a path as consisting of stops and moves separating a path into
145 periods of movement and stationary behavior. This conforms with the event-based model

146 for movement data outlined by Stewart Hornsby and Cole (2007) which contrasts with
147 the coordinate-based representation of movement typically employed. An event based
148 model for movement data still allows for the detection of movement patterns, but with
149 focus placed on combinations or sequences of events that identify a specific behavior,
150 such as an exodus of objects out of a zone or region (Stewart Hornsby and Cole 2007).
151 Further, event based models allow for enriching movement data with the geographic
152 information associated with events, for instance if events are related to spatial regions the
153 attributes of each region.

154

155 **3 – Review of Methods**

156 This section contains a review of quantitative analysis methods that exist within
157 seven areas of movement research; 1) time geography, 2) path descriptors, 3) path
158 similarity indices, 4) pattern and cluster methods, 5) individual-group dynamics, 6)
159 spatial field methods, and 7) spatial range methods. We emphasize techniques we feel
160 have particular merit for analysis with novel and emerging movement datasets.

161 **3.1 – Time Geography**

162 The concept of time geography was first presented in the 1960's and 1970's by
163 Torsten Hägerstrand at the Research Group for Process and System Analysis in Human
164 Geography at the University of Lund, Sweden (Lenntorp 1999). Time geography
165 (Hägerstrand 1970) represents a framework for investigating the constraints, such as an
166 object's maximum travel speed, on movement in both the spatial and temporal
167 dimensions. Hägerstrand expanded on the purely physical limitations of movement,
168 identifying three other types of constraints: *capability*, *coupling*, and *authority*

169 *constraints*. *Capability constraints* limit the activities of the individual because of their
170 biological construction and abilities, for example the necessity to eat and sleep. *Coupling*
171 *constraints* represent specific locations in space-time an individual must visit that limit
172 movement possibilities. *Authority constraints* are opposite of coupling constraints,
173 locations in space time an individual cannot visit, for example a mall after it has closed.
174 Contemporaries expanded on Hägerstrand's work providing both theoretical (Parkes and
175 Thrift 1975, Pred 1981) and applied (Lenntorp 1976, Burns 1979) extensions. Originally,
176 time geography was used solely to investigate the movement of humans, but has since
177 been reformulated for use with transportation networks (Miller 1991) and wildlife
178 ecology (Baer and Butler 2000).

179 Time geography uses volumes (Figure 1) capable of capturing the movement
180 limits of an object. A 3-D space (often termed cube, Kraak 2003, or aquarium, Kwan
181 2004), with two spatial axes representing geographic space and a third orthogonal axis for
182 time, is used to develop time geography volumes. The space-time cone (Figure 1a)
183 identifies the future movement possibilities of an object. A space-time prism (Figure 1b)
184 is used to quantify movement possibilities between known start and end locations. The
185 potential path area is the projection of the space-time prism onto geographic space
186 (Figure 1c), and is a purely spatial measurement of movement capability. A path is used
187 to portray the trajectory of movement through space-time. Bundling (Figure 1d) occurs
188 when multiple paths coincide in space and time, for example taking the same bus to
189 work. Typically, time geography is discussed qualitatively in terms of the aforementioned
190 volumes, but Miller (2005) has provided mathematical definitions for time geography
191 concepts that can be used in more rigorous quantitative analyses.

192 <approximate location Figure 1>

193 Recently, with advances in GISc and movement data, time geography is
194 experiencing a resurgence (Miller 2003). Lenntorp (1999) explains how time geography
195 has reached ‘the end of it’s beginning’, suggesting that current and future research using
196 GIS and novel movement datasets will present new and exciting developments in time
197 geography. Examples include using time geography to investigate mobility data on a
198 network (Miller and Wu 2000), factoring in uncertainty (Neutens et al. 2007), field-based
199 time geography (Miller and Bridwell 2009, further discussed in S3.6), and the
200 development of a probabilistic time geography (Winter 2009, further discussed in S3.6).

201 Time geography represents a useful tool for quantitative analysis of movement as
202 it contains a framework for measuring space-time bounds on movement. Movement
203 models that fail to consider the constraints provided by space and time often result in
204 misleading conclusions (Long and Nelson 2012). Methods that explicitly consider time
205 geography principles, even unknowingly (e.g., Yu and Kim 2006), avoid such deceptions.

206 **3.2 – Path Descriptors**

207 Path descriptors are measurements of path characteristics, for example velocity,
208 acceleration, and turning azimuth. Typically path descriptors may be calculated at each
209 point in a movement dataset, and can be scaled appropriately to represent interval or
210 global averages. Dodge et al. (2008) categorize a number of path descriptors as *primitive*
211 *parameters*, *primary derivatives*, or *secondary derivatives* based on simple measurements
212 in space, time, and space-time (see Table 2). Ecologists routinely use simple path
213 descriptors in the study of wildlife movement (Turchin 1998). Measures of movement
214 tortuosity have also been developed for the study of wildlife, for example path entropy

215 (Claussen et al. 1997), sinuosity (Benhamou 2004), and fractal dimension (Dicke and
216 Burrough 1988). Related to these are stochastic movement models (i.e., models where
217 fixes are obtained via random draws from distributions for movement displacement and
218 turning angle) such as Lévy flights (Viswanathan et al. 1996) and correlated random
219 walks (Kareiva and Shigesada 1983). When movement data are statistically fit to such
220 models, interpretation of model parameters can provide useful quantitative inference.

221 <approximate location Table 2>

222 **3.3 – Path Similarity Indices**

223 Path similarity indices are routinely used to quantify the level of similarity
224 between two movement trajectories. It is desirable for similarity indices to take the form
225 of a metric distance function, as metric functions are able to distinguish objects on an
226 interval scale of measurement (Sinha and Mark 2005). A metric distance function (d) is
227 one that computes a generalized scalar distance between two objects while satisfying the
228 following four properties (Duda et al. 2001):

- 229 (i) Non-negativity: $d(x, y) \geq 0$;
- 230 (ii) Reflexivity (uniqueness): $d(x, y) = 0$, iff $x = y$;
- 231 (iii) Symmetry: $d(x, y) = d(y, x)$;
- 232 (iv) Triangle Inequality: $d(x, z) \leq d(x, y) + d(y, z)$

233

234 The simplest similarity metric is a Euclidean measurement. Sinha & Mark (2005)
235 implement a time-weighted distance metric where spatial proximity (Euclidean) is
236 weighted by its temporal duration. Sinha & Mark (2005) also present a modified version
237 of the time-weighted distance metric for the situation where the two objects move over
238 different time intervals. Because the time-weighting is based on the duration an object
239 spends at a given spatial location, this index works best with movement data defined as a
240 series of stops and moves such as suggested by Spaccapietra et al. (2008). Yanagisawa et

241 al. (2003) present an alternative Euclidean-based similarity index that focuses on the
 242 shape of the movement path by normalizing the spatial coordinates of a path to a
 243 common plane. Euclidean measurements in the normalized spatial plane are used to
 244 identify similarly shaped movement paths. Euclidean distance is appropriate for
 245 comparisons in the spatial *or* temporal domains. However, Euclidean measurements are
 246 limited when data are represented with different scales (spatial and temporal). That is,
 247 what is the temporal equivalent to a 1 km distance in space? Despite these limitations,
 248 Euclidean distance similarity indices are frequently implemented by fixing either space or
 249 time and considering Euclidean distance in the other dimension, such as the above
 250 examples.

251 Other distance metrics may be more appropriate for assessing path similarities.
 252 The Hausdorff distance is a shape comparison metric commonly used to evaluate the
 253 similarity of two point sets (Huttenlocher et al. 1993), which has also been used to
 254 measure the similarity of movement paths. Given two movement paths M^a and M^b , the
 255 Hausdorff distance is defined as:

$$256 \quad H(M^a, M^b) = \max(h(M^a, M^b), h(M^b, M^a)) \quad [1]$$

$$257 \quad \text{with } h(M^a, M^b) = \max_{t \in T} \left(\min_{s \in S} d(M_t^a - M_s^b) \right) \quad [2]$$

258 where t and s are used to index fixes from M^a and M^b respectively, and d is a distance
 259 operator (e.g., Euclidean). Not originally designed for movement data, the Hausdorff
 260 distance performs poorly when analyzing movement paths as it fails to consider the
 261 ordering of points (Zhang et al. 2006), and is sensitive to outliers and data noise (Shao et
 262 al. 2010). As such, modified versions of the Hausdorff distance metric have been

263 designed specifically for use with movement paths (e.g., Atev et al. 2006, Shao et al.
264 2010).

265 The Fréchet distance metric may be more appropriate as a path similarity index as
266 it was initially designed for comparing polygonal curves. Formally the Fréchet distance
267 for two movement paths M^a and M^b is defined as:

$$268 \quad \delta_F(M^a, M^b) = \inf_{\alpha, \beta} \max_{t \in [0,1]} d(M^a(\alpha(t)), M^b(\beta(s))) \quad [3]$$

269 Where α (resp. β) is an arbitrary continuous non-decreasing function from $[0,1]$ onto
270 $[t_1 \dots t_n]$ (resp. $[s_1 \dots s_n]$) and d is a distance operator (Alt and Godau 1995). In simple
271 terms, the Fréchet distance measures the maximum distance apart of two coinciding
272 movement paths. The Fréchet distance, is best conceptualized using the analogy of a
273 person walking their dog, where no backwards movement is allowed. In the dog walking
274 example, the Fréchet distance is the minimum length of the dog's leash. The discretized
275 form of the Fréchet distance metric (Eiter and Mannila 1994) is useful for its computation
276 with movement data collected by discrete fixes, as described in section 2. In applications
277 involving objects that move with the same temporal granularity this calculation is simply
278 the maximum distance in space between any pair of fixes taken at the same time.
279 However, when object movement is recorded at differing temporal granularities or
280 extents, the value of the Fréchet distance metric is through the use of the scaling
281 functions (α, β) to measure similarity.

282 Vlachos et al. (2002) use longest common subsequences (LCSS), a method taken
283 from time-series analysis, to identify similar movement paths. The LCSS is defined as the
284 number of consecutive fixes from two (or more) paths (M^a, M^b, \dots) that are within d
285 spatial and τ temporal units of each other. This method can be extended to paths that

286 move at a distance, using mapping function $f(M)$ to translate M^b onto a space equivalent
287 to M^a . LCSS is advantageous as it is able to address issues relating movement paths taken
288 at different temporal granularities and/or extents. LCSS is efficient even with paths that
289 contain a significant amount of data noise. When outlying fixes are likely to influence the
290 calculation of other similarity indices LCSS is advantageous as it is insensitive to
291 extreme outliers. The disadvantage of the LCSS method is that it relies on the subjective
292 definition of thresholds – d and τ , and it fails the triangle inequality test (iv. above), and is
293 therefore not a metric distance function.

294 Similarity indices have also been extended to objects moving along a network.
295 For example, Hwang et al. (2005) calculate similarity using points-of-interest, such as
296 major intersections. Movement paths are considered similar if they pass through the same
297 points-of-interest in the same order. This index is not a metric distance function, but
298 moves away from Euclidean based measurements which are inappropriate in a network
299 scenario.

300 Recently, a new similarity method has been proposed by Dodge et al. (2012).
301 Here, a movement path is separated into segments where specific movement parameter
302 patterns (and derivatives of) are observed. In their example, velocity is the parameter of
303 interest, and the metrics deviation from the mean and sinuosity are used to define
304 movement parameter classes. For example, the letters A-D could be used to denote 4
305 unique movement parameter classes, and a path could then be represented as the
306 sequence [ACBCACBDBDA]. To assess the similarity of two paths, a modified version
307 of the edit distance (a string matching algorithm) is computed on the movement
308 parameter class sequences. This method measures similarity in the selected movement

309 parameters, rather than in the space-time geometry of the movement paths. As such, it
 310 may be more appropriate when similarity in various parameters, rather than space-time
 311 geometry is specifically of interest, for instance, in the study of hurricane path dynamics,
 312 as demonstrated by Dodge et al. (2012).

313 When objects interactively move with each other at a distance, they often exhibit
 314 correlated movement. Typically, similarity indices may identify such correlated
 315 movements by mapping the spatial coordinates of one path onto the spatial plane
 316 equivalent to the other. Alternatively, Shirabe (2006) presents a method for computing
 317 the correlation coefficient between two movement paths, each represented as a vector
 318 time-series. Consider a path M with $t = 1 \dots n$ fixes, then for $t = 2 \dots n$, $\mathbf{V} = [M_t - M_{t-1}] =$
 319 $[\mathbf{v}_t]$, is a vector time series of M . Given two two movement paths (M^v, M^w) represented as
 320 vector time-series \mathbf{V} and \mathbf{W} , the correlation coefficient is defined as:

$$321 \quad r(\mathbf{V}, \mathbf{W}) = \frac{\sum_{t=1}^{n-1} (\mathbf{v}_t - \bar{\mathbf{v}}) \cdot (\mathbf{w}_t - \bar{\mathbf{w}})}{\sqrt{\sum_{t=1}^{n-1} |\mathbf{v}_t - \bar{\mathbf{v}}|^2} \sqrt{\sum_{t=1}^{n-1} |\mathbf{w}_t - \bar{\mathbf{w}}|^2}} \quad [4]$$

322 Where $\bar{\mathbf{v}} = \frac{1}{n-1} \sum_{t=1}^{n-1} \mathbf{v}_t$ (resp. $\bar{\mathbf{w}}$) are mean coordinate vectors of (\mathbf{V}, \mathbf{W}) . Note that a
 323 movement path of n fixes is comprised of $n-1$ movement vectors, this distinction we keep
 324 for consistency with other methods. The numerator in [4] is the covariance, which
 325 indicates how the two motions deviate together from their respective means (Shirabe
 326 2006). Geometrically, the dot product in the numerator is the product of vector lengths
 327 multiplied by the cosine of the angle between them, which can be interpreted as the
 328 similarity. The correlation index ranges from -1 to 1, identifying both negatively and
 329 positively correlated movements. Important to note is that this correlation coefficient

330 relies on each movement's deviation from the respective mean, not the raw values of
331 each observed movement. Relating correlations to a global mean can be advantageous in
332 cases where two movements are correlated, but do not move in the same direction. The
333 first drawback of the formulation in [4] is that we are unable to disentangle the effects of
334 correlation in azimuth vs. magnitude of movements. A metric decomposed into each of
335 these components would be advantageous in situations where such distinctions are
336 necessary. A second drawback of equation [4] is that it requires that the fixes from each
337 movement path be taken simultaneously in order to be valid, which is not always
338 realistic. However, Shirabe (2006) does present an extension for modifying [4] to
339 measure movement path correlations at a temporal lag.

340 ***3.4 – Pattern and Cluster Methods***

341 Many applications are interested in identifying broad spatial-temporal patterns
342 from large movement databases (Benkert et al. 2007, Palma et al. 2008, Verhein and
343 Chawla 2008). For example, in the study of tourist behavior, often the goal is to identify
344 places of interest that are frequently visited (e.g., Ahas et al. 2007). Alternatively,
345 studying commuter patterns typically involves the identification of intersections and
346 routes being used by multiple individuals (Verhein and Chawla 2006). In these situations,
347 pattern and cluster methods are employed to identify similar movement behaviors or
348 places of interest.

349 Early work on indexing and querying movement databases coming from the
350 computer and database science literature (e.g., Güting et al. 2000, Pfoser et al. 2000) has
351 been essential to the development of pattern and cluster methods. For instance, many
352 methods for identifying patterns and clusters in large movement databases implement a

353 simple spatial or temporal query (Erwig et al. 1999). Alternatively, pattern or cluster
354 methods may implement one of the aforementioned path similarity indices and perform
355 pair-wise similarity computations over all permutations of stored movement paths. Paths
356 identified as similar based on a query or similarity index may convey some movement
357 pattern, or belong to the same cluster. The use of the term ‘cluster’ comes from methods
358 for statistical analysis of spatial point patterns (Diggle 2003), as many approaches used in
359 point pattern analysis have been adopted for movement data. For example, both Gao et al.
360 (2010) and Güting et al. (2010b) describe methods for performing *k*-nearest neighbor
361 queries in movement databases.

362 For the most part, the identification of patterns and clusters in large movement
363 databases focus on one of space, time, or space-time. Methods that identify spatial
364 clusters look at space first and time second, if at all (e.g., Benkert et al. 2007). The
365 simplest methods for detecting spatial clusters in movement databases generally require
366 that fixes from individual paths be represented as spatial points. Other spatial methods
367 look to define regions of interest (static or dynamic) and identify times at which
368 movement fixes are clustered in these spaces (Giannotti et al. 2007). Alternatively,
369 temporal clusters look at time first and space second, (e.g., D'Auria et al. 2005, Nanni and
370 Pedreschi 2006). Temporal clustering is enhanced (Palma et al. 2008) when movement
371 paths are represented by a sequence of stops (representing activities) and moves
372 (Spaccapietra et al. 2008).

373 Space-time approaches to identifying patterns and clusters strive to consider space
374 and time simultaneously. This is difficult, as previously mentioned, due to scaling
375 differences between space and time. Most space-time approaches fail to properly scale

376 space and time and degenerate to spatial clustering methods linked through time (e.g.,
377 Kalnis et al. 2005). Such methods routinely consider the following problem: given p
378 mobile objects, M^i , $i = 1 \dots p$. Each M^i consists of n fixes taken at coinciding times $t = (1,$
379 $\dots n)$. A set of α ($1 \leq \alpha \leq p$) spatial clusters are identified at each time t (for example with
380 multivariate clustering) using the spatial (x, y) coordinates of $M^i(t)$. In one example,
381 Shoshany et al. (2007) link clusters through time using linear programming. In their
382 example, moving objects M^i can switch between clusters, but all M^i must belong to a
383 cluster, as well clusters can emerge or disappear over time. The appeal of this approach is
384 that linear programming, frequently used in optimization research, can identify flows and
385 trends in movement data clusters.

386 Spatial-temporal association rules (STAR) learning represents an algorithm-based
387 method for discovering spatial-temporal patterns in movement databases (Verhein and
388 Chawla 2006, 2008). The patterns found by STAR methods are able to identify sources,
389 sinks, and thoroughfares in large mobility databases. Verhein and Chawla (2008)
390 demonstrate a STAR-miner software that implements their algorithm, and apply it to a
391 caribou dataset. STAR patterns rely on pre-determined spatial units (termed regions) over
392 which the algorithm is run. Unfortunately, the use of explicit spatial regions in their
393 derivation means that STAR are especially sensitive to changes in the definition of
394 regions (known as the modifiable areal unit problem - Openshaw 1984).

395 Pattern and cluster methods for movement data have also drawn on existing
396 methods from other applications. Shoval and Isaacson (2007) propose sequence
397 alignment methods, originally used to analyze DNA, as a way to identify patterns in
398 human travel behavior. With movement data, sequence alignment methods are able to

399 identify groups of objects that follow a similar sequence of events (e.g., using an event
400 based movement data representation, as in Stewart Hornsby and Cole 2007). Shoval and
401 Isaacson (2007) apply sequence alignment methods to tourist movement data and
402 conclude that sequence alignment methods have potential for identifying patterns of
403 spatial behavior in large movement databases. In another example, Eagle and Pentland
404 (2009) introduce a method for discovering eigenbehaviors in movement databases.
405 Eigenbehaviors represent trends or routines in individual movement data. Principle
406 component analysis is used to identify the eigenbehaviors of each person in their dataset.
407 In their example using the movements of people's daily routines, three trends emerge:
408 workday, weekend, and other behaviors. Increasingly complex questions could be
409 addressed using the eigenbehavior method.

410 ***3.5 – Individual-Group Dynamics***

411 The term individual-group dynamics is used to classify a suite of methods that
412 focus on individual object movement within the context of a larger group. This differs
413 fundamentally from methods for identifying patterns and clusters in movement databases.
414 Most current methods for investigating individual-group dynamics rely on computational
415 algorithms capable of searching movement databases for specific, pre-defined patterns.
416 These algorithms are often computationally demanding and inefficient (Gudmundsson et
417 al. 2007), and thus primarily used only in small, case-study examples.

418 Laube et al. (2004, 2005) provide the most comprehensive examination of
419 individual group-dynamics. Their concept of relative motion (REMO) can be used to
420 detect specific patterns (constancy, concurrence, and trend-setters) in groups of moving
421 objects. Constancy represents when an object moves in the same direction for a number

422 of consecutive fixes. An episode of concurrence occurs when multiple moving objects
423 move in the same direction at the same time. Trend-setters are objects that move in a
424 given direction ahead of a concurrence episode by a group of objects. Trend-setting is
425 identified as the most interesting property, and examined in more detail using the sport of
426 soccer as an example. Players that exhibit trend-setting behavior are able to better
427 anticipate the movement of play. Their concept of trend-setting has been further
428 developed for identifying leaders and followers in groups of moving objects, which is
429 potentially useful for the analysis of wildlife movement data (Andersson et al. 2008).
430 Laube et al. (2005)'s REMO method uses only movement azimuths to determine relative
431 motion. All other movement attributes, such as speed or distance, are ignored in their
432 derivation. Thus, REMO is useful only in situations where a group of objects move with
433 similar speeds and are contained in a relatable geographic space, such as the soccer
434 example. Another disadvantage is that the REMO method relies on the definition of
435 azimuthal breakpoints to define when objects are moving in a similar direction (e.g., East
436 is between 45° and 135°). Due to their discreteness, these breakpoints can lead to
437 misleading interpretations, for example when objects move in similar directions on either
438 side of a breakpoint. Alternatively, Noyon et al. (2007) evaluate the relative movement of
439 objects from the point-of-view of an observer within the system. Using changes in
440 relative inter-object distance and velocity, Noyon et al. (2007) identify relative behavior,
441 for example collision avoidance. Furthermore, Noyon et al. (2007) suggest that such
442 relative movement behavior also include other regions-of-interest such as lines and
443 polygons, which they include in their derivation.

444 Another problem routinely encountered in the study of movement is the detection
445 of flocks and convoys (e.g., groups of individuals that move as a cohesive unit). A flock
446 (see Figure 2a) is defined as a group of at least m moving objects (M) contained within a
447 circle of radius r over a minimum time interval - τ (Gudmundsson and van Kreveld 2006,
448 Benkert et al. 2008). Alternatively, a convoy (see Figure 2b) is defined as a group of at
449 least m moving objects (M) that are *density connected* at a distance d over a minimum
450 time interval - τ (Jeung et al. 2008). Density connected implies that there exists a
451 sequence of segments connecting all points in the convoy, each segment with length $\leq d$.
452 This definition of convoy relaxes the circular requirement of flocks affording flexibility
453 in the shape and extent of convoys that can be identified, for example Canada geese
454 forming their characteristic V-shape. Methods that look at flock/convoy behavior have
455 obvious usefulness in the study of wildlife herds, but also in monitoring crowd dynamics
456 at large events (Benkert et al. 2008). Like space-time clustering, methods describing
457 flocks or convoys build upon Hägerstrand's concept of bundling, identifying areas where
458 objects move coincidentally in space-time. The fundamental difference between the
459 identification of flocks or convoys and space-time cluster methods is that the definition of
460 a flock or convoy explicitly considers the individual in relation to the group in its
461 definition. That is, focus is placed on membership to a given group, with explicit
462 consideration of minimum requirements for flock or convoy behavior (e.g., the
463 parameters m and τ). Space-time cluster methods focus more on identifying broader
464 patterns, typically from large movement databases, and generally rely on pair-wise
465 comparisons of individual movement paths.
466 <approximate location Figure 2>

467 Recently, free space diagrams have been proposed for identifying single-file
 468 motion in movement databases (Buchin et al. 2010). To conceptualize a free space
 469 diagram consider two movement paths (M^a and M^b), over the time intervals m and n
 470 respectively, where the trajectory between fixes is given by some linear or other model
 471 (e.g., Tremblay et al. 2006). The functions φ_a and φ_b give the position of the objects a and
 472 b at time t . The free space diagram for a and b (following Buchin et al. 2010) is given by:

$$473 \quad F_\delta(M^a, M^b) = \{ (t^a, t^b) \in [1, n] \times [1, m] : |\varphi_a(t^a) - \varphi_b(t^b)| \leq \delta \} \quad [5]$$

474 which defines the set of all points in φ_a and φ_b that have a Euclidean distance below some
 475 threshold δ . The map of F_δ describes a two dimensional space where the axes
 476 correspond to the two paths, and the free space is defined as anywhere along the paths
 477 where the distance between the two paths is below the threshold δ . Buchin et al. (2010)
 478 demonstrate a method for interpreting free-space diagrams capable of identifying single-
 479 file movement patterns in groups of moving objects. A criticism of this method is that it
 480 relies on a subjectively defined threshold δ , to constrain the single-file movement
 481 process. Single-file motion has intuitive meaning, but is especially difficult to
 482 conceptualize geometrically. Methods that use Euclidean geometry to measure the spatial
 483 separation between leaders and followers (e.g., Andersson et al. 2008) are inadequate for
 484 identifying single-file movement warranting the free-space diagram approach.

485 **3.6 – Spatial Field Methods**

486 Often it is of interest to represent a movement path (or many movement paths) as
 487 a spatial field in order to identify areas in space (or space-time) that are more or less
 488 frequently visited. Field based representations are especially useful for visualizing large
 489 quantities of movement data when maps become cluttered. As many other spatial datasets

490 are stored as raster fields, a field-based representation of movement allows quantitative
491 map comparisons to be performed in a GIS.

492 Most methods for representing movement data as spatial fields have evolved from
493 those used to analyze spatial point patterns. When spatial point pattern methods are
494 employed the temporal component of movement fixes is ignored. Spatial point pattern
495 methods can be separated into quadrat or density based methods (Diggle 2003). The
496 simplest quadrat methods involve subdividing a study area into a regular grid and
497 determining point densities within each cell (e.g., Dykes and Mountain 2003,
498 Hadjieleftheriou et al. 2003). Cells with high point densities indicate spatial locations of
499 high use. Hengl (2008) proposes a quadrat based space-time density measure based on
500 distance and velocity within each cell [6].

$$501 \quad D_{xyt}(j) = \frac{\hat{d}_j}{\hat{v}_j} \quad [6]$$

502 Here $D_{xyt}(j)$ is the space-time density at cell j , \hat{d}_j is the length of the movement path
503 within cell j , and \hat{v}_j is the average velocity of movement within cell j . For a single
504 moving object the space-time density is simply interpreted as the duration of time the
505 object spends within each cell. If calculated for a movement database of many objects,
506 areas with higher space-time densities represent those where more objects spend more
507 time, the opposite with low values (Hengl et al. 2008). This approach has been extended
508 for three-dimensional visualization, where density is related to the lengths of multiple
509 paths in 3-D voxels defined by two spatial dimensions and a temporal dimension
510 (Demšar and Virrantaus 2010). Voxel densities are visualized in a space-time cube
511 (aquarium), and can be used for exploratory analysis of large movement databases.

512 Density based methods in spatial point pattern analysis stem from bivariate
 513 probability models, where movement fixes represent sampled locations from a two-
 514 dimensional probability density function (Silverman 1986). In the analysis of wildlife,
 515 density based models are frequently used to generate estimates of animal space use (also
 516 discussed in S3.7). Worton (1989) first applied kernel density estimation (KDE) to
 517 wildlife movement data to derive such a surface, termed a *utilization distribution*
 518 (Jennrich and Turner 1969). In movement applications, KDE can be interpreted as the
 519 intensity of space use based upon a collection of fixes. Calculation of KDE requires
 520 selection of a kernel shape and bandwidth parameter, with no consensus on the best way
 521 to do so (Hemson et al. 2005, Kie et al. 2010). Alternatively, Downs (2010) has proposed
 522 time geography’s potential path area (see Figure 1) to replace the kernel shape and
 523 bandwidth parameter, representing a novel approach for integrating temporal constraints
 524 into KDE analysis. Downs (2010) replaces the traditional kernel function with one based
 525 on the potential path area (termed geo-ellipse – G) from time geography [7].

$$526 \quad \hat{f}_i(x) = \frac{1}{(n-1)[(t_E - t_S)v]^2} \sum_{i=1}^{n-1} G \left(\frac{\|x - M_i\| + \|M_j - x\|}{(t_j - t_i)v} \right) \quad [7]$$

527 The numerator in this function sums the distance between a given point x and the object’s
 528 locations (M) at times i and j . The denominator is the maximum distance the object could
 529 have travelled in that time interval given its maximum velocity – v . Others have seen the
 530 need to move away from continuous representations of space, and have developed KDE
 531 for networks (Borruso 2008, Okabe et al. 2009). Such analysis is more appropriate for
 532 depicting the movement of urban travelers as their movement is restricted to travel
 533 networks of roads, paths, and sidewalks.

534 Random walks and diffusion theory have also been used to model movement as a
535 continuous spatial field. Horne et al. (2007) use Brownian bridges to model wildlife
536 movement as a continuous probability surface. Between two consecutive mobility points
537 the probability an object is at a given location at time t is defined using a bivariate normal
538 probability density function. More recently, probabilistic time geography has been
539 proposed (Winter 2009), where a similar probability surface is based on discrete random
540 walks in a cellular automata environment. Winter & Yin (2010) extend on the ideas of
541 Winter (2009) to include directed movements. Random walks are used to derive a
542 probability surface which explicitly considers the time geographic constraints on object
543 movement, using a similarly defined bivariate normal probability surface. Both Winter &
544 Yin (2010) and Horne et al. (2007) discuss the fact that determining movement
545 probabilities based on random walks is limited when objects do not move randomly.
546 Future work looking at probabilistic movement using other movement models (e.g.,
547 correlated random walks or on a network) is thus warranted for moving objects that can
548 be modeled this way. Alternatively, Miller & Bridwell (2009) propose a field-based time
549 geography. Field-based time geography uses movement cost surfaces in the calculation of
550 time geography volumes. Movement possibilities are evaluated in a similar manner to
551 Winter and Yin (2010) but based on an underlying movement cost surface (e.g., as in
552 least-cost path analysis in GIS, Douglas 1994). This approach is advantageous in that it
553 directly considers underlying variables impacting movement, however is limited in that
554 an accurate cost surface must be derived.

555 Brillinger et al. (2001, 2004) provide a unique approach for discovering patterns
556 in movement data. Stochastic differential equations are used to model movement as a

557 Markov process. The drift term in the stochastic movement model can be interpreted as a
558 spatial velocity field and used for exploratory analysis. The spatial velocity field
559 represents a potential function, whereby points of attraction and repulsion can be
560 identified. Methods for statistical inference (e.g., jackknifing) can be used to identify
561 statistically significant movement patterns within this velocity field (Brillinger et al.
562 2002). Brillinger (2007) further applies this approach for analyzing the flow of play in
563 soccer, where the spatial velocity field for ball movement is used to investigate a team's
564 attack formation.

565 ***3.7– Spatial Range Methods***

566 Spatial range can be broadly defined as the area (generally represented as a
567 polygon) containing an object's movement. Measures of spatial range can be useful for
568 examining object mobility and space use. Aspatial metrics, such as net displacement
569 (Turchin 1998), provide no information on the spatial distribution of movement, simply
570 measuring distance, thus spatial measurements are warranted. Furthermore, researchers
571 are often interested in intersections and/or differences in movement ranges (e.g., Righton
572 and Mills 2006). In such cases it is advantageous to represent point/line movement data
573 in an areal format (e.g., as a polygon).

574 The practice of representing movement data using spatial polygons has been
575 developed primarily by wildlife ecologists for studying wildlife home ranges (Burt 1943),
576 however, the concept of home range has also been applied to other subjects (e.g.,
577 children, Andrews 1973). Spatial range methods typically rely on the geometric
578 properties of movement data, for example the calculation of the minimum convex
579 polygon, a common measure of wildlife home range (Laver and Kelly 2008). Other

580 geometric methods include harmonic mean (Dixon and Chapman 1980), Voronoi
581 polygons (Casaer et al. 1999), and characteristic hull (Downs and Horner 2009). It is also
582 common to extract spatial range polygons from spatial field representations of movement
583 (e.g., those from S3.6) by extracting polygon contours based on density. For example,
584 with KDE a 95% volume contour is frequently used to delineate wildlife home range,
585 while a 50% volume contour is used to delineate core habitat areas (Laver and Kelly
586 2008). These spatial range methods ignore temporal information stored in movement data
587 and are likely to contain areas never visited by the object (commission error), and miss
588 actually visited locations (omission error) (Sanderson 1966).

589 Time geography volumes may also be used for generating spatial range estimates.
590 Long & Nelson (2012) propose a spatial range method for wildlife movement data based
591 on time geography's potential path area (Figure 1c). This method is capable of
592 identifying omission and commission errors in other spatial range methods (Long and
593 Nelson 2012). Such time geographic analysis is commonly used to study accessibility in
594 the context of human movement (Kwan 1998). The value of the potential path area as a
595 spatial range method is that it explicitly considers the temporal sequencing of movement
596 data in a time geography context. Spatial range methods that consider the temporal
597 component of movement data are advantageous over purely spatial methods (such as
598 convex polygons) as they consider movement data as a sequence of spatial points taken
599 through time, rather than as an arbitrary collection of spatial points.

600

601 **4 – Discussion**

602 ***4.1 - Time***

603 The first and foremost challenge to the quantitative analysis of movement data is
604 how to effectively characterize time. Despite having well-developed theory and tools for
605 analyzing space, geographers and the GISci community have historically struggled with
606 the temporal dimension (Peuquet 1994). Time is a single, continuous dimension that can
607 be portrayed as either monotonically linear or cyclical (Frank 1998). If time is portrayed
608 as linear, objects are not capable of re-visiting instances in time. If time is portrayed as
609 cyclical, the beginning of a new cycle infers that time is reset to some initial state, thus
610 revisiting is facilitated. For example, consider research on human daily routines; within
611 each day time is treated linearly, but is reset at the beginning of each day signifying the
612 start of a new cycle. Movement data collected over long periods may contain both linear
613 and cyclical temporal patterns, confounding representation and analysis.

614 Theoretical constructs for including time in GIS have long been discussed
615 (Langran and Chrisman 1988, Peuquet 1994) but remain challenging. Some spatial
616 datasets are easily represented at discrete time intervals in a GIS as different layers, for
617 example land cover data in different years. This representation allows for vertical
618 analysis through time using relatively simple map algebra (Mennis et al. 2005). Vertical
619 analysis through time is not straightforward with movement data, as objects move in both
620 space and time and cannot be explicitly linked through the spatial dimension. Others have
621 suggested the notion that geography's fetish for the static (Raper 2002) may lie at the root
622 of the time problem. In practice, researchers have begun to use a 3-D aquarium (drawing
623 on Hägerstrand's ideas) for representing time in GIS, however this is principally a
624 visualization tool (e.g., Kraak 2003, Andrienko and Andrienko 2007, Shaw et al. 2008).
625 Dynamic views (i.e., animations) may overcome the drawbacks of static portrayals of

626 movement, allowing more fluid representations of velocity and acceleration properties
627 (Andrienko et al. 2005). However, dynamic views are also visual-based, and lack
628 potential for developing quantitative analyses.

629 The challenge has been finding appropriate ways to simultaneously represent the
630 different scales of measurement for temporal and spatial attributes associated with
631 movement. Consider that it is common to use measurements of time and space
632 interchangeably in queries associated with movement from everyday life, for example if
633 you were asked the question: how *far* is it from here to the grocery store? You might
634 answer with “about 2 kilometers” or alternatively with “about a 5 minute drive”. Here, a
635 question of spatial distance associated with movement can be equivalently answered
636 using a spatial measurement (2 km) or temporal measurement (5 minutes). This has led to
637 alternative conceptualizations of movement where space and time can be represented
638 using relationships that can scale from spatial to temporal measurements, and vice-versa
639 (Parkes and Thrift 1975). For example, travel can be considered as the *consumption* of
640 physical distance through time (Forer 1998). However in the previous scenario, you may
641 have also answered with “about a 5 minute drive, depending on traffic”. Alternatively,
642 one might add that it depends on mode of transport (e.g., whether you walk or drive).
643 This alternative view demonstrates the non-linear and dynamic relationship that exists
644 between space and time which confounds the direct exchange of measurements of space
645 and time (Forer 1998). With movement data, time is often stored alongside spatial
646 attributes (e.g., $\langle x, y, t \rangle$), which naturally lends itself to Euclidean-type measurements in
647 the space-time aquarium. However, as demonstrated, time is poorly represented by such
648 direct physical measurements, because time cannot be represented as a linear function of

649 space. As there is still no consensus on the best way to represent time with movement
650 data, research on how to effectively characterize space and time in movement data
651 continues to require development.

652 Distance in space is easily computed using Euclidean (or other, such as network)
653 measurements. Differences in time are generally measured using clock times. The
654 conceptualization of a single space-time proximity measure remains one of the biggest
655 hurdles with quantitative analysis of movement data. Moving forward it is imperative to
656 go beyond simple Euclidean based measures, as time and space do not operate on equal
657 scales (Peuquet 2002). The Fréchet distance (Alt and Godau 1995) is an example of a
658 novel method for comparing the similarity of two movement paths that may prove useful
659 in future analyses. Nearest neighbor computations (e.g., Gao et al. 2010), most useful
660 with movement data stored as points, may also provide avenues for exploration.
661 Normalizing different data scales, common to other branches of quantitative analysis
662 such as multivariate cluster analysis (Duda et al. 2001), may be useful for comparing
663 movement processes across scales and relates to work using fractals for describing
664 movement datasets (Dicke and Burrough 1988). Normalization, however, may mask
665 scale specific patterns, and should be done with caution only when scale specific
666 behavior is less-important. Fundamentally, space and time have different dimensions and
667 require special consideration when analyzed together.

668

669 **4.2 – Scale**

670 With any spatial analysis the selection of analysis level (scale) will influence the
671 outcome of quantitative measures and the resulting inferences and conclusions (Dungan

672 et al. 2002). The study of scale and its impacts in spatial analysis remains a key topic in
673 geographic studies. In the analysis of movement data Laube et al. (2007) identify four
674 levels of analysis: *instantaneous*, *interval*, *episodal*, and *global* (Figure 3). The
675 instantaneous (“local”) level represents measures computed at any point along a
676 movement path. Interval (“focal”) level analysis takes the form of a moving temporal
677 window, but may also use a moving spatial window. Episodal (“zonal”) level analysis
678 looks at specific partitions of movement data, often related to some known event. Most
679 common is global level analysis, where a movement dataset is represented as a complete
680 path, from beginning to end, as a single entity. While some methods are specifically
681 designed for a given level of analysis others can be applied to various levels. Methods
682 that can be applied at different analysis levels may not scale from one level to the next,
683 meaning results at a lower level may not sum to the global result, as is the case with some
684 spatially local statistics (termed LISA - Anselin 1995).

685 <approximate location Figure 3>

686 Quantitative methods are also sensitive to changes in the temporal granularity at
687 which movement data is represented (Laube and Purves 2011). Methods for changing
688 granularity can be used when process scale is explicitly known, however this is rarely the
689 case. When movement data are over-sampled (i.e., too fine a granularity) data noise can
690 mask broader-scale process signals. When movement data are under-sampled (i.e., too
691 coarse a granularity) important movement events are missed, leading to incorrect
692 parameter estimates. Some ecologists have suggested that movement data should not be
693 sampled at even time intervals, but rather as a sequence of moves or steps relating to
694 individual behavior (Wiens et al. 1993, Turchin 1998). This aligns with the view of

695 Spaccapietra et al. (2008) that human movement data are best represented as a series of
696 stops (representing activities, as in the event-based model of Stewart Hornsby and Cole
697 2007) and moves. However, many developed methods tend to perform better when
698 implemented with regularly sampled movement data (e.g., Downs et al. 2012). As the
699 toolbox of methods for the quantitative of analysis of movement grows, it will be
700 important to identify at what analysis level(s) and over which temporal granularities
701 various methods are appropriate.

702 As previously identified, and following from Laube et al. (2007) and Laube and
703 Purves (2011), there are two fundamental issues of scale associated with movement
704 analysis, that is, analysis level and temporal granularity. Laube and Purves (2011)
705 suggest a third issue of scale may also exist, in that many approaches for movement
706 analysis are tested only on small, idealized datasets, and do not perform as expected when
707 carried out on larger, real-life datasets. As a result, many existing methods cannot be
708 readily implemented in practical scenarios with large volumes of movement data. We
709 take an alternative view on this issue. Testing of methods with smaller, idealized datasets
710 limits the scope of movement analysis to realistic and manageable problem sets, which
711 are in turn appropriate with *subsets* of a larger movement database. For example, the
712 detection of trend-setters (Laube et al. 2005) is only useful if there is some expectation
713 about where, if observed, this pattern is meaningful. In applied research, one should be
714 able to identify specific scenarios, within a larger movement database, where a given
715 technique is appropriate. Once these specific scenarios are identified, for example using
716 spatial-temporal queries, apply the technique of interest on this subset of the movement
717 database. The result is a multi-tiered analysis, where a specified technique is only

718 performed on smaller, appropriate subsets of the data. The goal being to break down
719 larger movement datasets into pieces resembling the idealized scenarios upon which
720 various techniques are useful.

721 **4.3 – Statistical Significance**

722 Often, it is desirable to examine quantitative problems using a statistical lens, that
723 is, to determine if some pattern is different than an expectation. For those less familiar
724 with statistical inference in GISci, we point the reader to the text by O’Sullivan and
725 Unwin (2010), which provides an introduction to these concepts. Spatial statistics often
726 rely on the concept of complete spatial randomness (CSR) as an *a priori* assumption for
727 assessing the statistical significance of observed spatial patterns (Cressie 1993). With
728 some types of spatial statistics (e.g., join counts, Cliff and Ord 1981) the distributions for
729 computing statistical tests are analytically derived. With other statistics, specifically most
730 spatially local measures, simulation procedures are used to generate test distributions,
731 making these statistics primarily exploratory (Boots 2002).

732 Random walks have been suggested as being to movement data what CSR is to
733 spatial data (Winter and Yin 2010). Two key methodological developments have
734 included random movement in their derivation: Brownian bridge home ranges (Horne et
735 al. 2007) and probabilistic time geography (Winter and Yin 2010). However, these two
736 examples represent essentially the same problem: defining a probability surface for
737 movement between two known locations in space-time. Authors of both methods concede
738 that random movement is inappropriate for modeling objects that move non-randomly,
739 but contend that it represents a necessary starting point.

740 The development of space-time statistics for movement is still in its infancy and
741 lacks clear direction for future research. Some have taken alternative views on this
742 problem, for example treating movement data as a bivariate time series using spatial
743 coordinates as dependent variables (e.g., Jonsen et al. 2003). Others have looked at
744 geographic space first, often ignoring the temporal component altogether (e.g., Casaer et
745 al. 1999). Both approaches are limited as they do not consider movement as a dynamic
746 process that is a function of both space *and* time. To adequately address the process of
747 movement, novel statistical techniques must consider space and time simultaneously in
748 their derivation. This will be challenging however, as inferential statistics are ill-suited to
749 the multidimensional complexity of movement (Holly 1978).

750

751 ***4.4 – Emerging Trends in Quantitative Movement Analysis***

752 Technological advances now facilitate real-time capture and analysis of
753 movement data on both wildlife and humans. In wildlife applications, real-time data
754 acquisition is providing opportunities for conservation and wildlife management. Dettki
755 et al. (2004) implemented a real-time tracking system for moose in Sweden, where data
756 on moose movements could be used to initiate the start-up and shut-down of forestry
757 operations in seasonal moose ranges. This idea relates directly to recent work identifying
758 the importance of *timing* in time geographic measures of space-time accessibility
759 (Neutens et al. 2010, Delafontaine et al. 2011a). As the interface between wildlife and
760 humans narrows, other potential applications exist for real-time tracking. Consider a
761 problematic large carnivore (e.g., lion or bear) residing in a national park. Rather than
762 relocating or exterminating this animal, a real-time tracking system could be used to

763 monitor the animal's movements. Park managers could use this information to improve
764 park safety and minimize human-animal conflicts through trail/site closures and
765 surveillance efforts.

766 Further developments with real-time movement data will involve the creation of
767 increasingly sophisticated models for predicting future movement locations. The space-
768 time cone from time geography (see Figure 1a) provides only the boundary for future
769 movement possibilities (e.g., O'Sullivan et al. 2000), factoring in the uneven distribution
770 of future movement possibilities (e.g., Winter 2009) provides more useful information for
771 prediction. Future movement possibilities can be linked to contextual factors such as
772 obstacles (Prager 2007), underlying movement cost surfaces (Miller and Bridwell 2009),
773 and object kinetics (Kuijpers et al. 2011). Further developments towards probabilistically
774 predicting future movements based on contextual factors will provide researchers and
775 analysts with powerful tools for linking real-time movement data with other data sources.

776 With human movement data a new field that is gaining momentum focuses on
777 leveraging real-time location data in everyday applications: location based services
778 (Raper et al. 2007). Location based services have developed coincidentally with the
779 availability of location-aware devices (e.g., GPS enabled cell-phones and handheld
780 devices), which are now integral to people's daily routines (Kumar and Stokkeland
781 2003). However, given the revealing nature of personal movement data, concerns over
782 the privacy and ownership rights of personal movement information continue to surface
783 (e.g., Dobson and Fischer 2003). With location based services, the fundamental goal is to
784 tailor individual applications, services, and marketing to a user's real-time location
785 (Raper et al. 2007). For example, methods for predicting future movements based on

786 contextual factors, when applied in a real-time application, could provide increased
787 functionality and improve user experiences with location based services. As methods for
788 analyzing real-time movement data emerge, their development in conjunction with
789 applications from location based services should be conducted in order to facilitate their
790 adoption in this field.

791 With the development of technologies for acquiring movement data, the ability to
792 capture finely grained movement data has increased substantially. Opportunities exist for
793 investigating properties of movement previously not feasible with coarser grained
794 movement data. For example, investigating velocities, accelerations, and the role of
795 momentum in moving objects is an area of opportunity. Current research is developing
796 methods for incorporating physical kinetics (based on object velocity and acceleration)
797 into the calculation of time geography volumes, such as those from Figure 1 (Kuijpers et
798 al. 2011). Another avenue for future work is the development of a probabilistic time
799 geographic framework, such as by Winter (2009), that considers the influence of kinetics
800 into the calculation of future movement probabilities.

801 Methods for investigating interactions between individuals in groups of moving
802 objects continue to develop, but remain limited in overall scope and sophistication. Laube
803 et al. (2005)'s relative motion concept can identify trendsetters, but uses only movement
804 azimuth in its derivation. Others have developed other ways to identify specific types of
805 interactions between moving individuals (e.g., Andersson et al. 2008; Buchin et al. 2010).
806 As our ability to characterize these patterns grows, it may be more useful to investigate
807 methods for quantifying the strength of interactions that occur in movement databases.
808 That is, can we measure *how* interactive are the movements of two individuals. The work

809 of Shirabe (2006) provides a necessary starting point for this research which could be
810 further investigated in light of this problem. Further, it may be necessary to examine
811 outside factors influencing the levels of interaction between individuals (e.g., barriers and
812 obstacles represented as lines/polygons, Noyon et al. 2007). Subsequently, how to
813 accommodate other data sources into models for measuring individual level interactions
814 in movement data remains an open research problem.

815 With time geography, Hägerstrand provided a theoretical context for looking at
816 the constraints of object movement. Contemporary geographers continue to expand on
817 time geographic concepts incorporating a range of ideas into time geographic theory
818 (e.g., Winter 2009, Miller and Bridwell 2009, Delafontaine et al. 2011b). As discussed by
819 Lenntorp (1999), Hägerstrand's time geography represents a set of conceptual and
820 methodological building blocks for use in analyzing and understanding movement as a
821 process. As the quantitative toolkit for analyzing movement continues to grow and
822 develop, those methods including theory and ideas from time geography in their
823 derivation will have increased value in a broader range of applications.

824 Other theoretical frameworks have also been successfully implemented in
825 movement research. For example, the idea that movement is motivated by an underlying
826 field (e.g., Brillinger et al. 2001) suggests that forces of attraction and repulsion may
827 influence movements. Such points of attraction, for example in wildlife, may be used to
828 investigate central place foraging theory (Orians and Pearson 1979). Markovian models
829 have also been used to demonstrate how movement operates as a diffusion process (e.g.,
830 Skellum 1951). Diffusion, originally used to describe random dispersal of organisms, can
831 also be related to crowd dynamics in humans (Batty et al. 2003). The use of theoretical

832 constructs in quantitative methods, such as the aforementioned examples, demonstrates
833 thoughtful development of ideas that in the end are easier to interpret for both the reader
834 and analyst.

835 It has been suggested that movement methods must consider the “geography
836 behind trajectories” (Bogorny et al. 2009) in order to understand the geographic
837 processes affecting observed movement patterns. Movement analysis is no longer limited
838 by available data, but rather by the tools required to manage and analyze movement
839 databases in more efficient and sophisticated ways (Miller 2010). Thus, the continued
840 development of methods capable of integrating increasingly large and complex
841 movement databases with available spatial and temporal layers is warranted. With such
842 analysis, the goal is to identify relationships between movement patterns and underlying
843 spatial and/or temporal variables. Data mining work is beginning to enrich movement
844 data with underlying geographic datasets (Alvares et al. 2007, Bogorny et al. 2009).
845 Quantitative methods for movement data must be further developed to consider
846 underlying geographic variables in order for movement to be understood as a function of
847 the environment. Similarly, novel movement datasets are emerging where attribute data
848 are recorded along with spatial and temporal records (e.g., $\langle ID, S, T, A \rangle$, where A
849 represents some attribute data). For example, wildlife tracking systems are being
850 equipped with devices, such as cameras (Hunter et al. 2005), that simultaneously record
851 information alongside movement fixes. The inclusion of attributes with movement fixes
852 can be termed *marked* movement data, comparable to the term marked point pattern in
853 the spatial statistics literature (Cressie 1993). Inclusion of attributes (numerical or
854 categorical) alongside spatial locations in movement data represents an area of

855 opportunity for advanced analysis in the movement-attribute space, as existing methods
856 are not designed for marked movement data.

857

858 **6 – Conclusions**

859 Novel movement datasets are not only becoming readily available they are
860 changing how data on movement processes are captured. Traditionally, movement data
861 have been collected as samples taken at coarse temporal granularities. Coarsely collected
862 movement data represents movement discretely and with considerable uncertainty
863 between sampled points. More recently, movement data are being collected at extremely
864 fine temporal granularities, such as 5 fixes/second with athletes. Finely grained
865 movement data represents a (near) continuous form of movement data which contains
866 minimal uncertainty in space-time location. Not only are existing methods ill-suited for
867 finely grained movement data, but the types of questions being asked must also be
868 revisited to consider that uncertainty between consecutive fixes is negligible.

869 Within GIS data formats, there is a clear lack of appropriate structures for
870 handling movement data. Those interested in purely visualizing movement data have
871 circumvented these problems by generating independent platforms for visualizations
872 (Andrienko et al. 2005). However, the development of quantitative methods is still
873 hindered by difficulties representing the temporal domain within GIS. The development
874 of geospatial data formats exclusively for movement data will invigorate future research
875 into quantitative methods for movement.

876 There is a clear need for novel quantitative methods for extracting information
877 and generating knowledge from ever-expanding movement datasets (Wolfer et al. 2001,

878 Laube et al. 2007). Most existing methods can be classified as data mining algorithms,
879 which are used to identify and categorize trends in movement databases, based on some *a*
880 *priori* notion about movement. Emerging problems investigate more complex patterns
881 and relationships contained in movement datasets, such as the identification of flocking
882 behavior (Benkert et al. 2008). Methods that are able to quantify interactions between
883 individuals (Laube et al. 2005), and with environmental variables (Patterson et al. 2009)
884 in movement databases will be increasingly relevant in more sophisticated movement
885 analyses. Movement models capable of quantifying relationships between moving objects
886 and dynamic features in the environment (e.g., traffic conditions) are justified in order to
887 measure the significance of events or changes on object movement.

888

889 **Acknowledgements**

890 The authors would like to thank the referees for their constructive remarks on earlier
891 versions of this manuscript. Their comments greatly improved the presentation of this
892 work. Support for this work was obtained from the Natural Sciences and Engineering
893 Research Council of Canada, and GEOIDE through the Government of Canada's
894 Networks for Centres of Excellence program.

895

896 **7 - References**

- 897
- 898 Ahas, R., *et al.* 2007. Seasonal tourism spaces in Estonia: Case study with mobile
899 positioning data. *Tourism Management*, 28, 898-910.
- 900 Alt, H. and Godau, M. 1995. Computing the Fréchet distance between two polygonal
901 curves. *International Journal of Computational Geometry & Applications*, 5(1-2),
902 75-91.
- 903 Alvares, L.O., *et al.*, 2007. A model for enriching trajectories with semantic geographical
904 information. *GIS'07: Proceedings of the 15th Annual ACM International*
905 *Symposium on Advances in Geographic Information Systems*, New York, NY, 8.
- 906 Andersson, M., *et al.* 2008. Reporting leaders and followers among trajectories of
907 moving point objects. *Geoinformatica*, 12, 497-528.
- 908 Andrews, H.F. 1973. Home range and urban knowledge of school-age children.
909 *Environment and Behaviour*, 5, 73-86.
- 910 Andrienko, N. and Andrienko, G. 2007. Designing visual analytics methods for massive
911 collections of movement data. *Cartographica*, 42(2), 117-138.
- 912 Andrienko, N., Andrienko, G. and Gatalisky, P., 2005. Impact of data and task
913 characteristics on design of spatio-temporal data visualization tools. *In*: Dykes,
914 J.A., MacEachren, A.M. and Kraak, M.J. eds. *Exploring Geovisualization*. New
915 York: Elsevier, 201-222.
- 916 Andrienko, G. *et al.*, 2010. GeoVA(t) – Geospatial visual analytics: Focus on time.
917 *International Journal of Geographical Information Science*, 24(10), 1453-1457.
- 918 Anselin, L. 1995. Local indicators of spatial association - LISA. *Geographical Analysis*,
919 27(2), 93-115.
- 920 Atev, S., Masoud, O. and Papanikolopoulos, N., 2006. Learning traffic patterns at
921 intersections by spectral clustering of motion trajectories. *IEEE/RSJ International*
922 *Conference on Intelligent Robots and Systems*, Oct. 9-15, Beijing, China, 4851-
923 4856.
- 924 Baer, L.D. and Butler, D.R. 2000. Space-time modeling of grizzly bears. *Geographical*
925 *Review*, 90(2), 206-221.
- 926 Batty, M., Desyllas, J. and Duxbury, E. 2003. The discrete dynamics of small-scale
927 events: agent-based models of mobility in carnivals and street parades.
928 *International Journal of Geographical Information Science*, 17(7), 673-697.
- 929 Benhamou, S. 2004. How to reliably estimate the tortuosity of an animal's path:
930 straightness, sinuosity, or fractal dimension? *Journal of Theoretical Biology*, 229,
931 209-220.
- 932 Benkert, M., *et al.*, 2007. Finding popular places. *In*: Tokuyama, T., ed. *ISAAC 2007*,
933 *LNCS 4835*, 776-787.
- 934 Benkert, M., *et al.* 2008. Reporting flock patterns. *Computational Geometry*, 41, 111-
935 125.
- 936 Bogorny, V., Kuijpers, B. and Alvares, L.O. 2009. ST-DMQL: A semantic trajectory data
937 mining query language. *International Journal of Geographical Information*
938 *Science*, 23(10), 1245-1276.
- 939 Boots, B. 2002. Local measures of spatial association. *Ecoscience*, 9(2), 168-176.

- 940 Borruso, G. 2008. Network density estimation: a GIS approach for analysing point
941 patterns in a network space. *Transactions in GIS*, 12(3), 377-402.
- 942 Brillinger, D.R. 2007. A potential function approach to the flow of play in soccer.
943 *Journal of Quantitative Analysis in Sports*, 3(1), Article 3.
- 944 Brillinger, D.R., *et al.*, 2001. The use of potential functions in modelling animal
945 movement. In: Saleh, M.A.K. ed. *Data Analysis from Statistical Foundations*.
946 Huntington, NY: Nova Science Publishers Ltd., 369-386.
- 947 Brillinger, D.R., *et al.* 2002. Employing stochastic differential equations to model
948 wildlife motion. *Bulletin Brazilian Mathematical Society*, 33(3), 385-408.
- 949 Brillinger, D.R., *et al.* 2004. An exploratory data analysis (EDA) of the paths of moving
950 animals. *Journal of Statistical Planning and Inference*, 122, 43-63.
- 951 Buchin, K., Buchin, M. and Gudmundsson, J. 2010. Constrained free space diagrams: a
952 tool for trajectory analysis. *International Journal of Geographical Information*
953 *Science*, 24(7), 1101-1125.
- 954 Burns, L.D., 1979. *Transportation, Temporal, and Spatial Components of Accessibility*.
955 Lexington, MA: Lexington Books.
- 956 Burt, W.H. 1943. Territoriality and home range concepts as applied to mammals. *Journal*
957 *of Mammalogy*, 24(3), 346-352.
- 958 Calenge, C., Dray, S. and Royer-Carenzi, M. 2009. The concept of animals' trajectories
959 from a data analysis perspective. *Ecological Informatics*, 4, 34-41.
- 960 Casaer, J., *et al.* 1999. Analysing space use patterns by Thiessen polygon and triangulated
961 irregular network interpolation: a non-parametric method for processing
962 telemetric animal fixes. *International Journal of Geographical Information*
963 *Science*, 13(5), 499-511.
- 964 Chrisman, N.R., 1998. Beyond the snapshot: Changing the approach to change, error, and
965 process. In: Egenhofer, M.J. and Golledge, R. eds. *Spatial and Temporal*
966 *Reasoning in Geographic Information Systems*. New York, NY: Oxford
967 University Press, 85-93.
- 968 Claussen, D.L., Finkler, M.S. and Smith, M.M. 1997. Thread trailing of turtles: methods
969 for evaluating spatial movements and pathway structure. *Canadian Journal of*
970 *Zoology*, 75, 2120-2128.
- 971 Cliff, A.D. and Ord, J.K., 1973. *Spatial Autocorrelation*. London: Pion Limited.
- 972 Cliff, A.D. and Ord, J.K., 1981. *Spatial Processes: Models and Applications*. London:
973 Pion.
- 974 Cressie, N., 1993. *Statistics for Spatial Data*. New York: John Wiley & Sons Inc.
- 975 D'Auria, M., Nanni, M. and Pedreschi, D., 2005. Time-focused density-based clustering
976 of trajectories of moving objects. *Proceeding of the Workshop on Mining Spatio-*
977 *temporal Data (MSTD-2005)*, Porto, 14.
- 978 Delafontaine, M., *et al.*, 2011a. The impact of opening hours on the equity of individual
979 space-time accessibility. *Computers, Environment and Urban Systems*, 35(4),
980 pp.276-288.
- 981 Delafontaine, M., Neutens, T. and Van De Weghe, N., 2011b. Modelling potential
982 movement in constrained travel environments using rough space – time prisms.
983 *International Journal of Geographical Information Science*, 25(9), 1389-1411.

- 984 Demšar, U. and Verrantaus, K. 2010. Space -time density of trajectories: exploring spatio-
 985 temporal patterns in movement data. *International Journal of Geographical*
 986 *Information Science*, 24(10), 1527-1542.
- 987 Dettki, H., Ericsson, G. and Edenius, L. 2004. Real-time moose tracking: An internet
 988 based mapping application using GPS/GSM-collars in Sweden. *Alces*, 40, 13-21.
- 989 Dicke, M. and Burrough, P.A. 1988. Using fractal dimensions for characterizing
 990 tortuosity of animal trails. *Physiological Entomology*, 13, 393-398.
- 991 Diggle, P.J., 2003. *Statistical Analysis of Spatial Point Patterns*. 2nd ed. New York, NY:
 992 Arnold.
- 993 Dixon, K.R. and Chapman, J.A. 1980. Harmonic mean measure of animal activity areas.
 994 *Ecology*, 61(5), 1040-1044.
- 995 Dobson, J.E., and Fisher, P., 2003. Geoslavery. *IEEE Technology and Society Magazine*,
 996 22(1), 47-52.
- 997 Dodge, S., Laube, P. and Weibel, R., 2012. Movement similarity assessment using
 998 symbolic representation of trajectories. *International Journal of Geographical*
 999 *Information Science*. DOI:10.1080/13658816.2011.630003, 26p.
- 1000 Dodge, S., Weibel, R. and Lautenschutz, A.-K. 2008. Towards a taxonomy of movement
 1001 patterns. *Information Visualization*, 7, 240-252.
- 1002 Douglas, D.H. 1994. Least-cost path in GIS using an accumulated cost surface and
 1003 slopelines. *Cartographica*, 31(3), 37-51.
- 1004 Downs, J.A., 2010. Time-geographic density estimation for moving point objects. *In*:
 1005 Fabrikant, S.I., *et al.*, eds. *GIScience 2010, LNCS 6292*, Sep. 14-17, Zurich,
 1006 Switzerland, 16-26.
- 1007 Downs, J.A. and Horner, M.A. 2009. A characteristic-hull based method for home range
 1008 estimation. *Transactions in GIS*, 13(5-6), 527-537.
- 1009 Downs, J.A., Horner, M.W. and Tucker, A.D., 2012. Time-geographic density estimation
 1010 for home range analysis. *Annals of GIS*, 17(3), 163-171.
- 1011 Drewe, P., 2005. What about time in urban planning & design in the ICT age? *CORP*
 1012 *2005 & Geomultimedia05*, Feb. 22-25, Vienna, Austria, 13-37.
- 1013 Duda, R.O., Hart, P.E. and Stork, D.G., 2001. *Pattern Classification*. New York: John
 1014 Wiley & Sons.
- 1015 Dungan, J.L., *et al.* 2002. A balanced view of scale in spatial statistical analysis.
 1016 *Ecography*, 25(5), 626-640.
- 1017 Dykes, J., MacEachren, A.M. & Kraak, M.J. eds., 2005. Exploring Geovisualization. 1st
 1018 ed., San Diego, CA: Elsevier.
- 1019 Dykes, J.A. and Mountain, D.M. 2003. Seeking structure in records of spatio-temporal
 1020 behaviour: visualization issues, efforts and applications. *Computational Statistics*
 1021 *& Data Analysis*, 43, 581-603.
- 1022 Eagle, N. and Pentland, A.S. 2009. Eigenbehaviors: identifying structure in routine.
 1023 *Behavioral Ecology and Sociobiology*, 63, 1057-1066.
- 1024 Eiter, T. and Mannila, H., 1994. *Computing discrete Fréchet distance*. Technical
 1025 University of Wien, CD-TR 94/64.
- 1026 Erwig, M., *et al.* 1999. Spatio-temporal data types: An approach to modeling and
 1027 querying moving objects in databases. *Geoinformatica*, 3(3), 269-296.
- 1028 Forer, P., 1998. Geometric approaches to the nexus of time, space, and microprocess:
 1029 Implementing a practical model for mundane socio-spatial systems. *In*:

- 1030 Egenhofer, M.J. and Golledge, R.G. eds. *Spatial and Temporal Reasoning in*
1031 *Geographic Information Systems*. New York: Oxford University Press, 171-190.
- 1032 Frank, A.U., 1998. Different types of "times" in GIS. *In: Egenhofer, M.J. and Golledge,*
1033 *R.G. eds. Spatial and Temporal Reasoning in Geographic Information Systems.*
1034 *New York: Oxford University Press, 40-62.*
- 1035 Gao, Y., *et al.* 2010. Algorithms for constrained k -nearest neighbor queries over moving
1036 object trajectories. *Geoinformatica*, 14, 241-276.
- 1037 Giannotti, F., *et al.*, 2007. Trajectory pattern mining. *Knowledge Discovery and Data*
1038 *Mining*, August 12-15, San Jose, CA, USA, 330-339.
- 1039 Goodchild, M.F. 1992. Geographical information science. *International Journal of*
1040 *Geographical Information Systems*, 6(1), 31-45.
- 1041 Gudmundsson, J. and van Kreveld, M., 2006. Computing longest duration flocks in
1042 trajectory data. *ACM-GIS'06, Nov. 10-11*, Arlington, VA, USA, 8.
- 1043 Gudmundsson, J., Van Kreveld, M. and Speckmann, B. 2007. Efficient detection of
1044 patterns in 2D trajectories of moving points. *Geoinformatica*, 11, 195-215.
- 1045 Güting, R.H., Behr, T. and Christian, D., 2010a. SECONDO: A platform for moving
1046 objects database research and for publishing and integrating research
1047 implementations. *Bulletin of the IEEE Computer Society Technical Committee on*
1048 *Data Engineering*, 33(2), 56-63.
- 1049 Güting, R.H., Behr, T. and Xu, J. 2010b. Efficient k -nearest neighbor search on moving
1050 object trajectories. *The VLDB Journal*, 19, 687-714.
- 1051 Güting, R.H., *et al.* 2000. A foundation for representing and querying moving objects.
1052 *ACM Transactions on Database Systems*, 25(1), 1-42.
- 1053 Güting, R.H. and Mamoulis, N., 2011. Special issue on data management for mobile
1054 services. *The VLDB Journal*, 20(5), 641-642.
- 1055 Güting, R.H. and Schneider, M., 2005. *Moving Objects Databases*. New York, NY:
1056 Elsevier.
- 1057 Hadjieleftheriou, M., *et al.* 2003. On-line discovery of dense areas in spatio-temporal
1058 databases. *LNCS 2750*, 306-325.
- 1059 Hägerstrand, T. 1970. What about people in regional science? *Papers of the Regional*
1060 *Science Association*, 24, 7-21.
- 1061 Hallin, P.O. 1991. New paths for time-geography? *Geografiska Annaler*, 73(3), 199-207.
- 1062 Hemson, G., *et al.* 2005. Are kernels the mustard? Data from global positioning system
1063 (GPS) collars suggests problems for kernel home-range analyses with least-
1064 squares cross-validation. *Journal of Animal Ecology*, 74, 455-463.
- 1065 Hengl, T., *et al.*, 2008. Geostatistical analysis of GPS trajectory data: space-time
1066 densities. *Proceedings of the 8th International Symposium on Spatial Accuracy*
1067 *Assessment in Natural Resources and Environmental Sciences*, Shanghai, June
1068 25-27, 17-24.
- 1069 Holly, B.P., 1978. The problem of scale in time-space research. *In: Carlstein, T., Parkes,*
1070 *D.N. and Thrift, N. eds. Time and Regional Dynamics*. New York: John Wiley &
1071 *sons, 5-18.*
- 1072 Horne, J.S., *et al.* 2007. Analyzing animal movement using Brownian bridges. *Ecology*,
1073 88(9), 2354-2363.
- 1074 Hornsby, K. and Egenhofer, M.J. 2002. Modeling moving objects over multiple
1075 granularities. *Annals of Mathematics and Artificial Intelligence*, 36, 177-194.

- 1076 Hunter, A., El-Sheimy, N. and Stenhouse, G.B. 2005. Up close and grizzly: GPS/camera
1077 collar captures bear doings. *GPS World*, Feb. 2005, 24-31.
- 1078 Huttenlocher, D.P., Klanderman, G.A. and Rucklidge, W.J. 1993. Comparing images
1079 using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and*
1080 *Machine Intelligence*, 15(9), 850-863.
- 1081 Hwang, J.-R., Kang, H.-Y. and Li, K.-J., 2005. Spatio-temporal similarity analysis
1082 between trajectories on road networks. In: Akoka, J., et al., eds. *Perspectives in*
1083 *Conceptual Modeling, ER Workshops, LNCS 3770*, 280-289.
- 1084 Jennrich, R.I. and Turner, F.B. 1969. Measurement of non-circular home range. *Journal*
1085 *of Theoretical Biology*, 22, 227-237.
- 1086 Jensen, C.S., Lu, H. and Yang, B., 2010. Indoor - A new data management frontier.
1087 *Bulletin of the IEEE Computer Society Technical Committee on Data*
1088 *Engineering*, 33(2), 12-17.
- 1089 Jeung, H., et al., 2008. Discovery of convoys in trajectory databases. *Proceedings VLDB*,
1090 Auckland, New Zealand, Aug. 23-28, 1068-1080.
- 1091 Johnston, R.J. 1997. W(h)ither spatial science and spatial analysis. *Futures*, 29(4/5), 323-
1092 336.
- 1093 Jonsen, I.D., Myers, R.A. and Mills Flemming, J. 2003. Meta-analysis of animal
1094 movement using state-space models. *Ecology*, 84(11), 3055-3063.
- 1095 Kalnis, P., Mamoulis, N. and Bakiras, S., 2005. On discovering moving clusters in spatio-
1096 temporal data. In: Medeiros, C.B., Egenhofer, M.J. and Bertino, E., eds. *9th*
1097 *International Symposium on Advances in Spatial and Temporal Databases, LNCS*
1098 *3633*, 364-381.
- 1099 Kareiva, P.M. and Shigesada, N. 1983. Analyzing insect movement as a correlated
1100 random walk. *Oecologia*, 56, 234-238.
- 1101 Kie, J.G., et al. 2010. The home-range concept: are traditional estimators still relevant
1102 with modern telemetry technology? *Philosophical Transactions of the Royal*
1103 *Society B*, 365, 2221-2231.
- 1104 Kondo, K. and Kitamura, R. 1987. Time-space constraints and the formation of trip
1105 chains. *Regional Science and Urban Economics*, 17, 49-65.
- 1106 Kraak, M.J., 2003. The space-time cube revisited from a geovisualization perspective.
1107 *Proceedings 21st International Cartographic Conference*, August 10-16, Durban,
1108 South Africa, 1988-1995.
- 1109 Kuijpers, B., Miller, H.J. and Othman, W., 2011. Kinetic space-time prisms. In *19th ACM*
1110 *SIGSPATIAL International Conference on Advances in Geographic Information*
1111 *Systems*. New York, NY: Association of Computing Machinery, 162-170.
- 1112 Kumar, S. and Stokkeland, J. 2003. Evolution of GPS technology and its subsequent use
1113 in commercial markets. *International Journal of Mobile Communications*, 1(1/2),
1114 180-193.
- 1115 Kwan, M.P. 1998. Space-time and integral measures of individual accessibility: A
1116 comparative analysis using a point-based framework. *Geographical Analysis*,
1117 30(3), 191-216.
- 1118 Kwan, M.P. 2004. GIS methods in time-geographic research: Geocomputation and
1119 geovisulization of human activity patterns. *Geografiska Annaler*, 86 B(4), 267-
1120 280.

- 1121 Langran, G. and Chrisman, N.R. 1988. A framework for temporal geographic
1122 information. *Cartographica*, 25(3), 1-14.
- 1123 Laube, P., *et al.* 2007. Movement beyond the snapshot - Dynamic analysis of geospatial
1124 lifelines. *Computers, Environment and Urban Systems*, 31, 481-501.
- 1125 Laube, P., Imfeld, S. and Weibel, R. 2005. Discovering relative motion patterns in groups
1126 of moving point objects. *International Journal of Geographical Information
1127 Science*, 19(6), 639-668.
- 1128 Laube, P. and Purves, R.S. 2011. How fast is a cow? Cross-scale analysis of movement
1129 data. *Transactions in GIS*, 15(3), 401-418.
- 1130 Laube, P., van Kreveld, M. and Imfeld, S., 2004. Finding REMO - Detecting relative
1131 motion patterns in geospatial lifelines. In: Fisher, P.F. ed. *Developments in
1132 Spatial Data Handling, Proceedings of the 11th International Symposium on
1133 Spatial Data Handling*. Berlin: Springer, 201-215.
- 1134 Laver, P.N. and Kelly, M.J. 2008. A critical review of home range studies. *Journal of
1135 Wildlife Management*, 72(1), 290-298.
- 1136 Lenntorp, B., 1976. *Paths in Space-Time Environments: a Time Geographic Study of
1137 Movement Possibilities of Individuals*. Stockholm: Gleerup.
- 1138 Lenntorp, B. 1999. Time-geography - at the end of its beginning. *GeoJournal*, 48, 155-
1139 158.
- 1140 Long, J.A. and Nelson, T.A. 2012. Time geography and wildlife home range delineation.
1141 *Journal of Wildlife Management*, 76(2), 407-413.
- 1142 Mark, D.M., 1998. Geospatial lifelines. *Integrating Spatial and Temporal Databases*,
1143 98471, Dagstuhl Seminars, 12.
- 1144 Mennis, J., Viger, R. and Tomlin, C.D. 2005. Cubic map algebra functions for spatio-
1145 temporal analysis. *Cartography and Geographic Information Science*, 32(1), 17-
1146 32.
- 1147 Miller, H.J. 1991. Modelling accessibility using space-time prism concepts within
1148 geographical information systems. *International Journal of Geographical
1149 Information Systems*, 5(3), 287-301.
- 1150 Miller, H.J. 2003. What about people in geographic information science? *Computers,
1151 Environment and Urban Systems*, 17, 447-453.
- 1152 Miller, H.J. 2005. A measurement theory for time geography. *Geographical Analysis*, 37,
1153 17-45.
- 1154 Miller, H.J. 2010. The data avalanche is here. Shouldn't we be digging? *Journal of
1155 Regional Science*, 50(1), 181-201.
- 1156 Miller, H.J. and Bridwell, S.A. 2009. A field-based theory for time geography. *Annals of
1157 the Association of American Geographers*, 99(1), 49-75.
- 1158 Miller, H.J. and Wu, Y.H. 2000. GIS Software for measuring space-time accessibility in
1159 transportation planning and analysis. *Geoinformatica*, 4(2), 141-159.
- 1160 Nanni, M. and Pedreschi, D. 2006. Time-focused clustering of trajectories of moving
1161 objects. *Journal of Intelligent Information Systems*, 27, 267-289.
- 1162 Nathan, R., *et al.* 2008. A movement ecology paradigm for unifying organismal
1163 movement research. *Proceedings of the National Academy of Science*, 105(49),
1164 19052-19059.
- 1165 Neutens, T., *et al.* 2007. Human interaction spaces under uncertainty. *Transportation
1166 Research Record*, 2021, 28-35.

- 1167 Neutens, T. *et al.*, 2010. Evaluating the temporal organization of public service provision
1168 using space-time accessibility analysis. *Urban Geography*, 31(8), 1039-1064.
- 1169 Noyon, V., Claramunt, C. and Devoegele, T. 2007. A relative representation of trajectories
1170 in geographical spaces. *Geoinformatica*, 11, 479-496.
- 1171 O'Sullivan, D., Morrison, A. and Shearer, J. 2000. Using desktop GIS for the
1172 investigation of accessibility by public transport: an isochrone approach.
1173 *International Journal of Geographical Information Science*, 14(1), 85-104.
- 1174 O'Sullivan, D. and Unwin, D.J., 2010. *Geographic Information Analysis*. 2nd ed.
1175 Hoboken, NJ: John Wiley & Sons.
- 1176 Okabe, A., Satoh, T. and Sugihara, K. 2009. A kernel density estimation method for
1177 networks, its computational method and a GIS-based tool. *International Journal*
1178 *of Geographical Information Science*, 23(1), 7-32.
- 1179 Openshaw, S., 1984. *The Modifiable Areal Unit Problem*. Norwich, England: GeoBooks.
- 1180 Orians, G.H. and Pearson, N.E., 1979. On the theory of central place foraging. *In*: Horn,
1181 D.J., Stairs, G.R. and Mitchell, R.D. eds. *Analysis of Ecological Systems*.
1182 Columbus, OH: Ohio State University Press, 155-178.
- 1183 Palma, A.T., *et al.*, 2008. A clustering-based approach for discovering interesting places
1184 in trajectories. *Proceedings of the 2008 ACM Symposium on Applied Computing*,
1185 Fortaleza, Ceara, Brazil, 6.
- 1186 Parkes, D.N. and Thrift, N. 1975. Timing space and spacing time. *Environment and*
1187 *Planning A*, 7, 651-670.
- 1188 Patterson, T.A., *et al.* 2009. Classifying movement behaviour in relation to environmental
1189 conditions using hidden Markov models. *Journal of Animal Ecology*, 78, 1113-
1190 1123.
- 1191 Peuquet, D.J. 1994. It's about time: a conceptual framework for the representation of
1192 temporal dynamics in geographic information systems. *Annals of the Association*
1193 *of American Geographers*, 84(3), 441-461.
- 1194 Peuquet, D.J., 2002. *Representations of Space and Time*. New York, NY: The Guilford
1195 Press.
- 1196 Pfoser, D., Jensen, C.S. and Theodoridis, Y., 2000. Novel approaches to the indexing of
1197 moving object trajectories. *26th Conference on Very Large Databases*, Cairo,
1198 Egypt, Sep. 10-14, 12p.
- 1199 Prager, S.D., 2007. Environmental contextualization of uncertainty for moving objects.
1200 *Computers, Environment and Urban Systems*, 31(3), 303-316.
- 1201 Pred, A., ed., 1981. *Space and Time in Geography*. Gleerup: The Royal University of
1202 Lund.
- 1203 Pultar, E., *et al.* 2010. EDGIS: a dynamic GIS based on space time points. *International*
1204 *Journal of Geographical Information Science*, 24(3), 329-346.
- 1205 Raper, J., 2002. The dimensions of GIScience. *Second International Conference of*
1206 *Geographic Information Science, GIScience 2002.*, Boulder, CO, Last retrieved:
1207 May 2011, <[http://www soi.city.ac.uk/~raper/research/GIScience2002-OHs-](http://www soi.city.ac.uk/~raper/research/GIScience2002-OHs-pub.ppt)
1208 [pub.ppt](http://www soi.city.ac.uk/~raper/research/GIScience2002-OHs-pub.ppt)>.
- 1209 Raper, J., *et al.* 2007. A critical evaluation of location based services and their potential.
1210 *Journal of Location Based Services*, 1(1), 5-45.
- 1211 Righton, D. and Mills, C. 2006. Application of GIS to investigate the use of space in
1212 coral reef fish: a comparison of territorial behaviour in two Red Sea

- 1213 butterflyfishes. *International Journal of Geographical Information Science*, 20(2),
1214 215-232.
- 1215 Rose, C. 1977. Reflections on the notion of time incorporated in Hägerstrand's time-
1216 geographic model of society. *Tijdschrift voor Economische en Sociale Geografie*,
1217 68(1), 43-50.
- 1218 Sanderson, G.C. 1966. The study of mammal movements: A review. *Journal of Wildlife*
1219 *Management*, 30(1), 215-235.
- 1220 Shao, F., Cai, S. and Gu, J., 2010. A modified Hausdorff distance based algorithm for 2-
1221 dimensional spatial trajectory matching. *The 5th International Conference on*
1222 *Computer Science & Education*, Aug. 24-27, Hefei, China, 166-172.
- 1223 Shaw, S.L., Yu, H. and Bombom, L.S. 2008. A space-time GIS approach to exploring
1224 large individual-based spatiotemporal datasets. *Transactions in GIS*, 12(4), 425-
1225 441.
- 1226 Sheppard, E. 2001. Quantitative geography: representations, practices, and possibilities.
1227 *Environment and Planning D: Society and Space*, 19, 535-554.
- 1228 Shirabe, T., 2006. Correlation analysis of discrete motions. In: Raubal, M., et al. eds.
1229 *GIScience 2006. LNCS, vol. 4197*. Berlin: Springer-Verlag, 370-382.
- 1230 Shoshany, M., Even-Paz, A. and Bekhor, S. 2007. Evolution of clusters in dynamic point
1231 patterns: with a case study of Ants' simulation. *International Journal of*
1232 *Geographical Information Science*, 21(7), 777-797.
- 1233 Shoal, N. and Isaacson, M. 2007. Sequence alignment as a method for human activity
1234 analysis in space and time. *Annals of the Association of American Geographers*,
1235 97(2), 282-297.
- 1236 Silverman, B.W., 1986. *Density Estimation for Statistics and Data Analysis*. New York:
1237 Chapman and Hall.
- 1238 Sinha, G. and Mark, D.M. 2005. Measuring similarity between geospatial lifelines in
1239 studies of environmental health. *Journal of Geographical Systems*, 7, 115-136.
- 1240 Skellum, J.G. 1951. Random dispersal in theoretical populations. *Biometrika*, 38, 196-
1241 218.
- 1242 Spaccapietra, S., et al. 2008. A conceptual view on trajectories. *Data & Knowledge*
1243 *Engineering*, 65, 126-146.
- 1244 Stewart Hornsby, K. and Cole, S. 2007. Modeling moving geospatial objects from an
1245 event-based perspective. *Transactions in GIS*, 11(4), 555-573.
- 1246 Thomas, J.J. and Cook, K.A., 2005. *Illuminating the Path: The Research and*
1247 *Development Agenda for Visual Analytics*. National Visualization and Analytics
1248 Center. 184 p.
- 1249 Tremblay, Y., et al. 2006. Interpolation of animal tracking data in a fluid environment.
1250 *The Journal of Experimental Biology*, 209, 128-140.
- 1251 Turchin, P., 1998. *Quantitative Analysis of Movement: Measuring and Modelling*
1252 *Population Redistribution in Animals and Plants*. Sunderland, MA: Sinauer.
- 1253 Verhein, F. and Chawla, S., 2006. Mining spatio-temporal association rules, sources,
1254 sinks, stationary regions and thoroughfares in object mobility databases.
1255 *Proceedings of the 11th International Conference on Database Systems for*
1256 *Advanced Applications (DASFAA)*, 187-201.
- 1257 Verhein, F. and Chawla, S. 2008. Mining spatio-temporal patterns in object mobility
1258 databases. *Data Mining & Knowledge Discovery*, 16, 5-38.

- 1259 Viswanathan, G.M., *et al.* 1996. Lévy flight search patterns of wandering albatrosses.
1260 *Nature*, 381, 413-415.
- 1261 Vlachos, M., Gunopulos, D. and Kollios, G., 2002. Robust similarity measures for mobile
1262 object trajectories. *5th International Workshop on Mobility in Databases and*
1263 *Distributed Systems (MDDS)*, Aix-en-Provence, France, 721-726.
- 1264 Wells, L.L., 1981. Real-time missile tracking with GPS. *National Aerospace Meeting*,
1265 Apr. 8-10, Washington, DC, 56-61.
- 1266 Wiens, J.A., Crist, T.O. and Milne, B.T. 1993. On quantifying insect movements.
1267 *Environmental Entomology*, 22(4), 709-715.
- 1268 Winter, S., 2009. Towards a probabilistic time geography. *In: Mokbel, M., Scheuermann,*
1269 *P. and Aref, W.G., eds. Proceedings of the 17th ACM SIGSPATIAL International*
1270 *Conference on Advances in Geographic Information Systems*, Seattle, WA, 528-
1271 531.
- 1272 Winter, S. and Yin, Z.C. 2010. Directed movements in probabilistic time geography.
1273 *International Journal of Geographical Information Science*, 24(9), 1349-1365.
- 1274 Wolfer, D.P., *et al.* 2001. Extended analysis of path data from mutant mice using the
1275 public domain software Wintrack. *Physiology & Behavior*, 73, 745-753.
- 1276 Worton, B.J. 1989. Kernel methods for estimating the utilization distribution in home-
1277 range studies. *Ecology*, 70(1), 164-168.
- 1278 Yanagisawa, Y., Akahani, J.I. and Satoh, T., 2003. Shape-based similarity query for
1279 trajectory of mobile objects. *In: Chen, M.-S., et al. eds. Mobile Data*
1280 *Management, LNCS 2574*. Berlin: Springer-Verlag, 63-77.
- 1281 Yu, B. and Kim, S.H., 2006. Interpolating and using most likely trajectories in moving-
1282 objects databases. *In: Bressan, S., Kung, J. and Wagner, R., eds. DEXA 2006,*
1283 *LNCS 4080*, 718-727.
- 1284 Zhang, Z., Huang, K. and Tan, T., 2006. Comparison of similarity measures for trajectory
1285 clustering in outdoor surveillance scenes. *IEEE International Conference on*
1286 *Pattern Recognition*, 1135-1138.
- 1287
1288

Table 1: Terms used synonymously for describing movement data.

Description	Term	Synonymous terms (with selected references)
A single record of object movement (of the form <ID, S, T>).	Movement Fix (M_i)	point, observation, relocation
A sequence of ordered records in time depicting the movement of a single object.	Movement Path (M^a)	space-time path (Hägerstrand 1970), trip-chain (Kondo and Kitamura 1987), geospatial lifeline (Mark 1998), trajectory, trace, track
A collection of records depicting the movements of many objects or the same object at different occasions, potentially including attribute information.	Movement Database	moving objects database (Güting and Schneider 2005)

Table 2: Parameters extractable from movement data sorted by dimension. After Table 1 from Dodge et al. (2008).

	Primitive	Primary Derivatives	Secondary Derivatives
Spatial (x, y)	Position	Distance	Spatial distribution
		Direction	Change of direction
		Spatial extent	Sinuosity
Temporal (t)	Instance	Duration	Temporal distribution
	Interval	Travel time	Change of duration
Spatio-temporal (x, y, t)	—	Speed	Acceleration
		Velocity	Approaching rate

Figure Captions

Figure 1: Volumes used in Hägerstrand's time geography: a) space-time cone, b) space-time prism, c) potential path area, and d) path bundling.

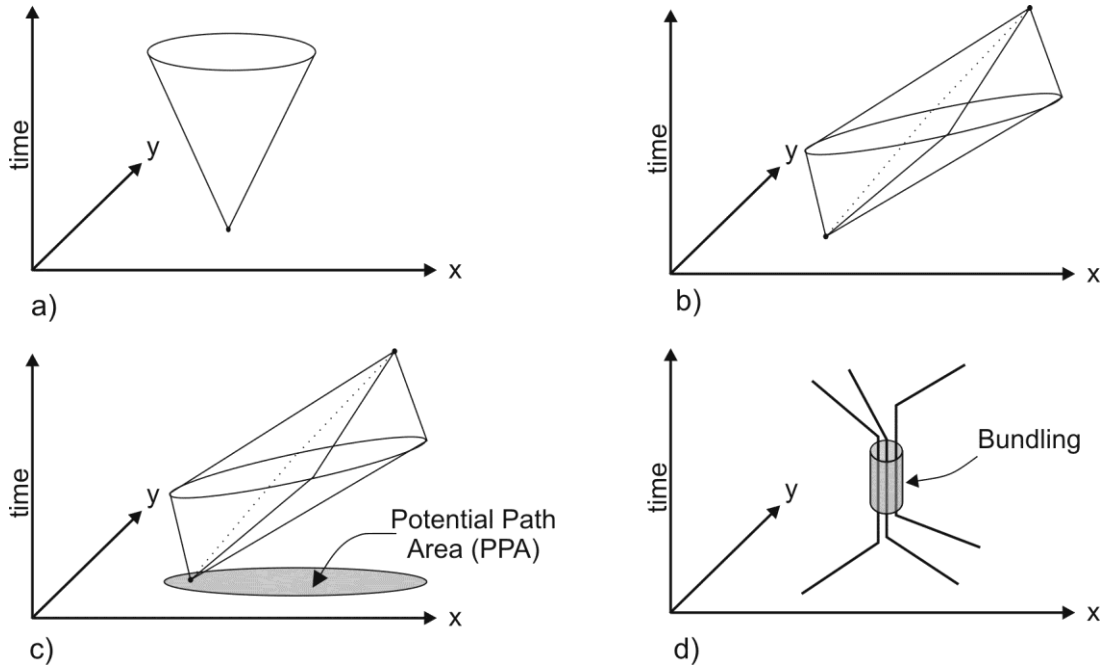


Figure 2: Comparison between definitions of a) flocks, and b) convoys. A flock requires objects be contained in a circle of radius r , while a convoy is defined as those objects that are *density connected* at distance d . Both methods require that objects be included in the group over a minimum time interval τ .

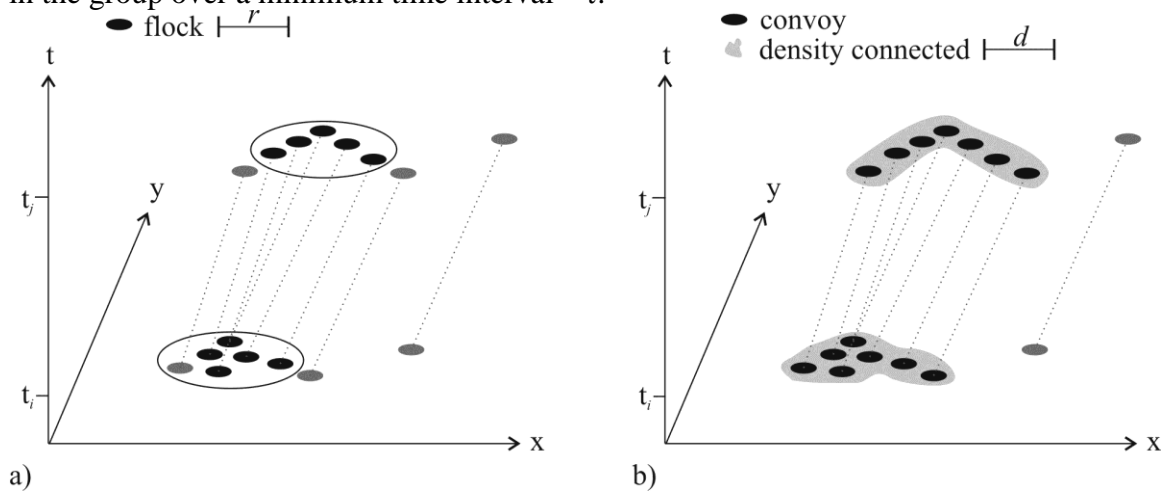


Figure 3: Four analysis levels for movement data: *instantaneous*, *interval*, *episodal*, and *global*. After Figure 2 from Laube et al. (2007).

