



HAL
open science

A framework for pre-processing individual location telemetry data for freshwater fish in a river section

Dominique Lamonica, Hilaire Drouineau, Hervé Capra, Hervé Pella, Anthony Maire

► **To cite this version:**

Dominique Lamonica, Hilaire Drouineau, Hervé Capra, Hervé Pella, Anthony Maire. A framework for pre-processing individual location telemetry data for freshwater fish in a river section. *Ecological Modelling*, 2020, 431, pp.109190. 10.1016/j.ecolmodel.2020.109190 . hal-03128885

HAL Id: hal-03128885

<https://hal.science/hal-03128885>

Submitted on 12 Sep 2023

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

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

A framework for pre-processing individual location telemetry data for freshwater fish in a river section

Dominique Lamonica^{*a,b}, Hilaire Drouineau^c, Hervé Capra^a, Hervé Pella^a,
Anthony Maire^d

^aINRAE, RiverLy, HYNES (Irstea-EDF R&D), F-69625, Villeurbanne, France

^bInstitute of Landscape and Plant Ecology, University of Hohenheim, Stuttgart, Germany

^cINRAE, EABX, HYNES (Irstea-EDF R&D), F-33612, Cestas, France

^dEDF R&D, LNHE (Laboratoire National d'Hydraulique et Environnement), HYNES
(Irstea - EDF R&D), 6 quai Watier, 78401 Chatou Cedex, France

Abstract

Animal movement study often relies on individual tracking. The data scale (in time and space) varies according to the species, the environment where individuals live, or the exogenous processes that drive movement. To explore freshwater fish movement in rivers, fine-scale data are needed. Also, in rivers, recorded telemetry frequently shows missing data and location errors. The irregular time-steps, huge amount of data, environmental complexity (river section) and how fish move in such anisotropic environments undermine the use of statistical frameworks such as state-space models. To deal with these specificities, data pre-treatment can be required. We propose a generic method of telemetry data pre-processing, which can be transposed to other datasets. This framework includes interpolation to handle trajectories at fine time scales and performs data analysis within a state-space model. [We combined analyses on observed and simulated data at various interpola-](#)

*Corresponding author. Tel.: +33472208732.

Email address: dominique.lamonica@gmail.com (Anthony Maire)

tion time-steps to choose the one that best preserves the general movement while reducing the total amount of data required. First, we directly compared raw and interpolated data, and the results of parameter inference of a simple state-space model using the interpolated data. The state-space model infers behavioural state based on speed and turning angle between successive locations in animal trajectories. We also included two additional variables computed from raw data: a quantitative indicator of the correspondence between the interpolated trajectory and the raw data, and the variance of turning angles of raw data within the interpolation time-step. We were finally able to determine the most appropriate time-step to obtain locations that were regularly spaced in time and to reduce the amount of data while maintaining the precision of the raw data. Computational time was reduced 12-fold by using a 30-second time-step to interpolate data simulated at 3-second intervals. The inclusion of the two variables derived from raw data compensated for the loss of information in interpolated trajectories and allowed more efficient discrimination between behaviours.

Keywords: Animal location data, Movement model, State-space model, Switching behaviour, Bayesian inference, Parameter estimation

1. Introduction

2 Movement is a key issue in animal ecology and has been the focus of in-
3 creasing research, especially in aquatic ecology, in both marine and freshwater
4 environments (Giuggioli and Bartumeus 2010, Lennox et al. 2017, Nathan et al. 2008).
5 Movement influences many processes, at individual, population and commu-
6 nity levels: habitat selection (Block et al. 2011, Capra et al. 2017), migra-

7 tion (Bultel et al. 2014, Drouineau et al. 2017, Tétard et al. 2016, Tétard et al. 2019),
8 trophic dynamics (Lima 2002), spread of disease (Carraro et al. 2017, Jonsen et al. 2001,
9 Pinder et al. 2005) and adaptability to climate change and extreme events
10 (Boucek et al. 2017). The study of animal movement often relies on indi-
11 vidual tracking. Recent technical progress has revolutionized such studies:
12 the development of high-frequency tags enables high-frequency data col-
13 lection at fine spatial resolution (Cagnacci et al. 2010, Hussey et al. 2015,
14 Lennox et al. 2017). Tracking data at high temporal resolution are inval-
15 able for species that are rarely static although staying in the same area.
16 Location per hour, half-day or day provides information on individual move-
17 ment within an animal’s living-range (e.g., within a region, or along a migra-
18 tion route) over a long period of time rather than small quick displacements
19 within or between local habitats (Capra et al. 2018, Donaldson et al. 2014).
20 In contrast, tracking several locations per minute allow displacements to be
21 described at fine scale, enabling precise investigations of suitable migratory
22 conditions, habitat selection or behaviour choices at individual level (e.g.
23 (Capra et al. 2017, Cooke et al. 2004, Tétard et al. 2019)). Telemetry data
24 at a very fine time scale (of a few seconds) are being increasingly collected
25 worldwide (Cooke et al. 2013, Hussey et al. 2015, Lennox et al. 2017). Po-
26 sitions are generally estimated by triangulation, either by satellite or by
27 multiple fixed receivers.

28 Statistical modelling frameworks have been developed to analyse aquatic
29 telemetry data (Whoriskey et al. 2019) by assessing and correcting posi-
30 tion errors (Bergé et al. 2012, Roy et al. 2014), dealing with missing data
31 (Woillez et al. 2016), inferring behaviour (Dorazio and Price 2019, Thiebault et al. 2018,

32 Vermard et al. 2010), and investigating the influence of the environment on
33 individual movements (Bestley et al. 2013, Drouineau et al. 2017, Patterson et al. 2009).
34 State-space modelling is one of the useful existing statistical frameworks
35 for animal movement analysis (Dorazio and Price 2019, Jonsen et al. 2003,
36 Joo et al. 2013, Patterson et al. 2008).

37 Most animal movement studies have focused on large terrestrial or marine
38 species, e.g. (Andersen et al. 2017, Bailey et al. 2008, Franke et al. 2006,
39 Hedger et al. 2008, McClintock et al. 2012) which move in a wide and open
40 environment, few have dealt with freshwater organisms. Telemetry data for
41 anisotropic, irregular environments, such as large rivers, show some speci-
42 ficities (Cooke et al. 2013). Firstly, triangulation is made difficult by ground
43 irregularities and the presence of vegetation, causing frequent signal loss. Sec-
44 ondly, the precision of triangulation varies in space (Bergé et al. 2012). And
45 thirdly, the anisotropic closed conditions of hydrographic networks are spe-
46 cific limitations on animal movement (Quaglietta and Porto 2019, Sutherland et al. 2015)
47 and complicate the trajectory analysis. Fish navigation is known to be influ-
48 enced by physical cues such as current fields and physical obstacles (e.g., [dams](#)
49 or river banks) (Goodwin et al. 2014) leading to highly orientated navigation
50 which does not fulfil the isotropy assumption usually applied in open environ-
51 ments. The second major issue is the temporal and spatial scales at which fish
52 movement is characterised. Telemetry studies of freshwater fish generally fo-
53 cus on seasonal movement, e.g. (Dorazio and Price 2019, Fraley et al. 2016,
54 Koehn and Nicol 2016, Muhlfeld 2012), and modelling frameworks are rarely
55 applied to analysing individual movements at fine scales. However, such
56 fine-scale data and corresponding analytical methods are critical to study-

57 ing small-scale foraging movements of fishes within their home range, which
58 (Dingle 1996) called "station keeping" movement, in contrast to migration
59 over large distances.

60 The combination of environmental limitations, which entail irregularities in
61 signal recording, and fine spatial and temporal scale has three main conse-
62 quences. Firstly, irregular time-steps have to be handled. Secondly, analysing
63 the huge amount of data generated requires great computer power, which
64 may lead to a trade-off between reducing computing time and not degrading
65 fine-scale data quality. And thirdly, the consideration of entire individual
66 trajectories from the beginning to the end of the recording period is not al-
67 ways necessary; a certain number of trajectories spaced in time can also be
68 used to explore individual behaviour.

69 To deal with irregular time-steps, an appropriately longer time-step can be
70 chosen, coupled with interpolation, if necessary, to deal with any missing
71 data. Increasing the time-step generally reduces the rate of missing data
72 while also reducing the amount of data to be analysed. However, it impairs
73 overall precision compared to the raw dataset, and may impact ecological in-
74 terpretations based on these trajectories. In this paper, we propose a generic
75 method of data processing to accurately infer individual behaviours from
76 trajectories at fine temporal and spatial scales. Analysis was performed at
77 different time-steps and compared so as to select the most appropriate one:
78 *i.e.*, the one that preserved the general movement while efficiently discrimi-
79 nating between behaviours (here, slow and fast movements). To this end, we
80 directly compared the raw and interpolated data, and compared the results of
81 the parameters inference of a simple state-space model with the interpolated

82 data. We did not deal with location error, as our aim was to infer indi-
83 vidual behaviour based on observed movement rather than reconstruct the
84 exact individual trajectories. Moreover, usual correction methods are based
85 on an isotropy assumption (*i.e.*, that there is no favoured direction within
86 the space, and animal navigation consequently depends only on behaviour
87 and not on environmental characteristics such as flow fields around physical
88 obstacles), which is not fulfilled here.

89 To illustrate this approach, we used telemetry data for individual fish col-
90 lected in the Rhône River [at 3-second intervals](#) (Capra et al. 2017). The
91 initial objective of the study was to infer the relationships between fish be-
92 haviour and hydraulic conditions through analysis of fish movements. In this
93 context, we used a state-space model to discriminate fish behaviours. For
94 that, raw fish location data (hereafter referred to as "raw data") must be
95 pre-processed. While maintaining the precision of the raw data, gaps be-
96 tween locations must be dealt with so as to obtain regularly time-spaced
97 data.

98 **2. Material and methods**

99 *2.1. Case study*

100 Bergé et al. (Bergé et al. 2012) collected telemetry data on freshwater
101 fish in the Rhône River using the HTI ([https://www.innovasea.com/fish-](https://www.innovasea.com/fish-tracking/)
102 [tracking/](https://www.innovasea.com/fish-tracking/)) acoustic fixed telemetry system. Our system includes a set of
103 pre-positioned hydrophones used to detect ultrasounds emitted by acoustic
104 tags (frequency of 307 KHz). Tags signals that allow the identification of
105 the tag, and precise positioning of the tag through a triangulation process

106 provided that the signal is detected by at least 3 hydrophones (hydrophones
107 were all connected to a single controller and synchronized with UTC time
108 to improve the triangulation). Further details are provided in Berg et al.
109 (Bergé et al. 2012).

110 Locations of 94 individuals of various species were tracked for 3 months at
111 a time interval of 3 seconds. The dataset suffered from the usual defects
112 affecting tracking data. Firstly, individuals were not systematically located
113 every 3 seconds during the 3 months of the experiment. There were two
114 types of gap in the data: large gaps, in which the individual signal was lost
115 for several minutes to several days, and small gaps where the individual signal
116 was lost for a few 3-second periods (*i.e.*, 3 seconds to a few minutes). A second
117 defect was specific to the triangulation process of the HTI system: some
118 successive locations form artefactual "star-shaped" trajectories (Appendix
119 A, Figure A1). Though this type of star-shaped pattern is specific to the
120 triangulation process, it is more generally one of the types of location error
121 besetting most tracking studies. The fine time-scale data may also incur
122 a specific problem: when an individual is static or moving very slowly, its
123 successive locations can be tracked as being as distant as when it is moving
124 faster, due to triangulation error.

125 2.2. Proposed data pre-processing

126 Three pre-processing steps were performed on the raw data. The first was
127 the choice of a time-step p . The second dealt with the large gaps between
128 certain successive locations. For this, we considered using a threshold s equal
129 to twice the time-step p , in order to split the trajectory in two while avoiding
130 interpolation in-between points that were too far apart in time. If the dura-

131 tion between two successive locations $(x, y)_i$ and $(x, y)_{i+1}$ (x and y denoting
132 the spatial coordinates and i the [index](#) of the location) is greater than s ,
133 then these two locations are assumed to belong to two distinct independent
134 trajectories. The raw data for a given trajectory are denoted hereafter as the
135 "support of the trajectory". The third step consisted in the linear interpo-
136 lation of new locations within each trajectory according to the time-step p
137 to deal with missing data within each trajectory. Interpolations were carried
138 out using the *move* R package (Kranstauber and Smolla 2008).

139 We tested 6 interpolation time-steps (with the associated s threshold): 30, 60,
140 120, 180, 240, and 300 seconds, denoted by p_k , $k \in [1, 6]$. In order to test the
141 different means of pre-processing the raw data, 16 of the 731 trajectories of
142 individual European catfish (*Silurus glanis*) were selected. European catfish
143 were chosen because this species included well-tracked individuals presenting
144 numerous locations, and thus longer trajectories than for other species.

145 The complete trajectories can comprise three general patterns: (i) [travelling](#)
146 ([T](#)), in which the individual moves over a long distance, where start and end
147 locations are distant from one another; (ii) [stationary](#) ([S](#)), where locations
148 are concentrated within a short perimeter (potentially "star-shaped") with
149 start and end locations nearby; and (iii) a mixed pattern ([B](#), for "bi-type")
150 where a stationary move follows a long move or vice versa. Four trajecto-
151 ries of each for the stationary and travelling types, and 8 trajectories for
152 the mixed type were selected for the test dataset. Two examples of each
153 trajectory type at each time-step are shown in Figure 1.

154 *2.3. Validation criteria*

155 To compare the interpolated trajectories using the different time-steps p_k ,
 156 several criteria were considered.

157 *2.3.1. Computed variables*

158 Two variables were computed to quantify the consistency between the
 159 raw data and each interpolated trajectory. They showed different patterns
 160 according to the type of trajectory, and were computed to see how they
 161 varied according to time-step. The two additional variables, denoted by U_1
 162 and U_2 , concerned speed and turning angles respectively. For a trajectory j
 163 associated with time-step p_k , let $(x, y)_{n_j \times p_k}$, $n_j \in [0, N_j]$ (with N_j the total
 164 number of locations in the trajectory j) be the n_j^{th} location in the trajectory.
 165 Considering the support of the trajectory j , let $i \in [0, I_{n_j}]$ be an index for
 166 raw data locations, denoted by $(x, y)_{i_{n_j}}$, between time $n_j \times p_k$ and $(n_j + 1) \times p_k$
 167 of the trajectory locations and $\theta_{i_{n_j}}$ the angle between the locations $(x, y)_{i_{n_j}}$,
 168 $(x, y)_{(i+1)_{n_j}}$ and $(x, y)_{(i+2)_{n_j}}$.
 169 For a time-step p_k and a total time $n_j \times p_k$ of the trajectory j , $U_{1_{k, n_j}}$ and
 170 $U_{2_{k, n_j}}$, are defined as follows:

$$U_{1_{(k, n_j)}} = \frac{d((x, y)_{n_j \times p_k}, (x, y)_{(n_j + 1) \times p_k})}{\sum_{i=0}^{I_{n_j} - 1} d((x, y)_{i_{n_j}}, (x, y)_{i_{n_j + 1}})} \quad (1)$$

171 where $d(X_1, X_2)$ represents the covered distance between locations X_1 and
 172 X_2 . U_1 is thus the ratio between the covered distance according to the
 173 interpolated data and the covered distance according to the raw data, within
 174 the interpolation time-step. U_1 represents a quantitative indicator of the

175 correspondence between the interpolated trajectory and the raw data: the
176 higher U_1 , the better the fit.

$$U_{2(k,n_j)} = \text{var}(\theta_{i_{n_j}}) \quad (2)$$

177 where $i \in [0, I_n]$ and $\text{var}(X)$ denotes the variance of X . U_2 is thus the
178 variance of turning angles of raw data within the interpolation time-step. If
179 the variance is small, the trajectory is rectilinear within the interpolated time-
180 step and little information is lost. However, a large variance may indicate a
181 star-shaped pattern, especially if associated with a low U_1 .

182 2.3.2. State-space modelling

183 We developed a state-space model based on (Morales et al. 2004) to dis-
184 criminate between the different individual behaviours in the 16 selected tra-
185 jectories. Model parameters were estimated independently for each time-step,
186 so as to assess variations in parameter estimates according to the time-step.
187 Results for each time-step were also compared with and without adding U_1
188 and U_2 , to see whether behaviour discrimination was improved by including
189 these two variables.

190 *Model definition.* Model states correspond to the succession of fish behaviours
191 at each time increment, with two possible behaviours: "Resting" (denoted by
192 R), which corresponds to slow or erratic movements, and "Moving" (denoted
193 by M), which corresponds to fast, oriented movements. We assumed con-
194 stant behaviour switching probabilities between successive time increments.
195 The observation model links the state at time t to corresponding movement
196 variables (*i.e.*, speed between two locations and turning angles between two

197 moves) (Morales et al. 2004). Low mean speed and high turning angles vari-
 198 ance are taken to characterise "Resting" behaviour, whereas high speed and
 199 mean turning angles around 0° are taken to characterise "Moving" behaviour.
 200 "Resting" behaviour is expected to predominate in "Stationary" trajectories,
 201 "Moving" behaviour in "Travelling" trajectories. The model is written as fol-
 202 lows:

203 Transition matrix

$$M_q = \begin{pmatrix} q_{R \rightarrow R} & 1 - q_{R \rightarrow R} \\ 1 - q_{M \rightarrow M} & q_{M \rightarrow M} \end{pmatrix} \quad (3)$$

204 State equation

$$z_t \sim \mathcal{B}(M_q[z_{t-1}]) \quad (4)$$

205 Observation model

$$\begin{aligned} y_{v_t} &\sim \mathcal{G}(a[z_t], \lambda[z_t]) \\ y_{\phi_t} &\sim \mathcal{WC}(b[z_t], \rho[z_t]) \end{aligned} \quad (5)$$

206 with $q_{R \rightarrow R}$ and $q_{M \rightarrow M}$ being the probability of maintaining resting, or
 207 moving, behaviour following resting, or moving, behaviour, and z_t being the
 208 behaviour at time t (R or M). Concerning the data, y_{v_t} is the observed speed
 209 between $t - 1$ and t , and y_{ϕ_t} is the observed turning angle between $t - 2$ and
 210 $t - 1$ and $t - 1$ and t . Concerning the parameters, $a[z_t]$ and $\lambda[z_t]$ describe the
 211 speed for behaviour z at time t and $b[z_t]$, and $\rho[z_t]$ describes the turning angle
 212 for behaviour z at time t . \mathcal{B} , \mathcal{G} and \mathcal{WC} represent the Bernoulli [distribution](#),
 213 the Gamma [distribution](#) and the Wrapped Cauchy [distribution](#), respectively.
 214 Given that U_1 and U_2 allow integration of information derived from the raw
 215 data independently of the interpolation, two models were tested to determine

216 whether taking into account of U_1 and U_2 improved the model’s behaviour
 217 discrimination. The first (hereafter, ”Model 1”) included only variables from
 218 interpolated data (observed speed and observed turning angles). In the sec-
 219 ond (hereafter, ”Model 2”), U_1 and U_2 were added to the variables, the
 220 observation model becoming as follows:

$$\begin{aligned}
 y_{v_t} &\sim \mathcal{G}(a[z_t], \lambda[z_t]) \\
 y_{\phi_t} &\sim \mathcal{WC}(b[z_t], \rho[z_t]) \\
 U_{1_t} &\sim \text{Beta}(\alpha[z_t], \beta[z_t]) \\
 U_{2_t} &\sim \mathcal{N}(\mu[z_t], \sigma[z_t])
 \end{aligned}
 \tag{6}$$

221 with $\alpha[z_t]$ and $\beta[z_t]$ being the parameters describing U_1 for behaviour z at
 222 time t , $\mu[z_t]$ and $\sigma[z_t]$ being the parameters describing U_2 for behaviour z
 223 at time t . *Beta* and \mathcal{N} represent for the Beta [distribution](#) and the normal
 224 [distribution](#), respectively.

225 *Computation.* Bayesian inference was used to fit the model to the data: *i.e.*,
 226 the trajectories obtained for each time-step p_k . A single model was fitted for
 227 all selected trajectories at once. Prior distributions were defined summaris-
 228 ing all available information on each parameter (Appendix B, Table B1).
 229 Markov Chain Monte Carlo (MCMC) computations were performed using
 230 JAGS software and the *rjags* R package (Plummer 2009, R Core team 2018).
 231 A total of 10,000 iterations were performed as a burn-in phase, and infer-
 232 ence was based on 20,000 additional iterations for each of the three inde-
 233 pendent chains ([with different initiations](#)). The Gelman and Rubin tests
 234 (Gelman and Rubin 1992) were used to check the convergence of the esti-
 235 mation process. The computation times for the different time-steps p_k were
 236 compared. The estimated behaviour at each time-step was recorded from the

237 MCMC iterations. For each location, the credibility of resting and moving
238 behaviours was calculated, as the mean of the behaviours estimated for all
239 the MCMC iterations. We also calculated the mean duration of maintain-
240 ing each behaviour (*i.e.*, the expected value of the geometric distribution
241 with switching probability, multiplied by the time-step), with the medians
242 and 95% credibility intervals of the posterior distributions for both switching
243 probabilities.

244 Twelve fits of the models were performed for all 16 selected trajectories: one
245 with Model 1 and one with Model 2, for each of the 6 tested time-steps. We
246 also performed 2 additional fits using all the individual trajectories: one with
247 Model 1 and one with Model 2, for the time-step which appeared to be the
248 most appropriate after checking the various criteria. As convergence might
249 not be reached for a simple model with two behaviours (Morales et al. 2004),
250 we also tested similar models (without and with U_1 and U_2) including a third
251 intermediate behaviour.

252 *Simulated data analysis.* A Markov chain of 10,000 behaviours was simu-
253 lated with fixed transition probabilities ($q_{R \rightarrow R} = 0.99$ and $q_{M \rightarrow M} = 0.97$)
254 using Model 1. Each point in the chain stands for a theoretical location of
255 the individual and, for each point, a speed and a relative angle was derived
256 from the distributions of speeds and relative angles of each behaviour (with
257 the following fixed parameter values for speed and turning angles distribu-
258 tions: $a_1 = 0.1$, $a_2 = 1.5$, $\lambda_1 = 10.5$, $\lambda_2 = 5$, $b_1 = 3.15$, $b_2 = 0$, $\rho_1 = 0.3$
259 and $\rho_2 = 0.9$). Then, for each chain point, a theoretical location was cal-
260 culated. Thus, we obtained a series of 10,000 locations associated with a
261 behaviour. The theoretical time-step of this trajectory was 3 seconds. The

262 pre-processing described in the Section 2.2 was applied: for the 6 tested time-
263 steps, the simulated trajectory was interpolated, speeds and turning angles
264 were calculated, as well as the two additional variables U_1 and U_2 . Then, we
265 estimated the parameters of Model 1 (with speeds and turning angles only)
266 and Model 2 (with speeds, turning angles and the two additional variables
267 U_1 and U_2) for each time-step and for the raw trajectory.
268 Computation time was recorded and, when the model converged, the good-
269 ness of fit between the simulated and inferred behaviours was computed.
270 For the 3-second trajectory (raw simulated data), inferred and simulated
271 behaviours were compared directly. For the interpolated trajectories, first
272 we computed the mean of the simulated behaviours between two locations
273 within the time-step, then the root mean squared error for all the behaviours
274 of the trajectory were calculated, comparing for each location the mean simu-
275 lated behaviour and the inferred one. To compare results between time-steps,
276 the root mean squared error was divided by the number of locations in the
277 interpolated trajectory.

278 3. Results

279 3.1. Influence of the time-step between interpolated locations on behaviour 280 inference

281 3.1.1. General correspondence of interpolated trajectories and raw data

282 Superimposition of raw data and interpolated trajectories is (for conve-
283 nience) shown for 6 typical trajectories out of the 16 studied trajectories in
284 Figure 1. For bi-type 1 and the two stationary trajectories, the interpolated
285 trajectories did not show any major differences in overall movement between

286 time-steps. However, the duration of the trajectories increased with the time-
287 step, which was expected because of the increasing threshold separating two
288 distinct trajectories, whereas the number of locations decreased. On the con-
289 trary, for bi-type 2 and the two travelling trajectories with the most changes
290 in direction, the interpolated trajectories did not show any major differences
291 for time-steps below 60 seconds, but diverged from the raw data for the four
292 longer time-steps (120, 180, 240 and 300 seconds).

293 *3.1.2. Additional variables*

294 The ratio between covered distance from interpolated and raw data,
295 namely U_1 , showed a similar pattern for all types of trajectory and for all
296 time-steps (Figure 2a). As expected, U_1 decreased with increasing time-step,
297 due to shorter covered distance in interpolated data. In the 5 time-steps
298 longer than 30 seconds, the difference in U_1 between the three types of tra-
299 jectories was greater than in the first time-step (30 seconds). Likewise, the
300 variance of turning angles within the interpolation time-step, namely U_2 ,
301 showed a pattern similar to U_1 (Figure 2b).

302 *3.1.3. Modelling results*

303 *Parameter estimation.* The MCMC algorithm consistently converged accord-
304 ing to Gelman and Rubin diagnostics for each simulation performed with the
305 16 selected trajectories. For all parameters, posterior distributions were nar-
306 rower than prior distributions. The narrowness of the posterior distributions
307 suggests that sufficient information was available from the data to accurately
308 estimate the model parameters. Overall, there were progressive differences
309 between parameter estimates with increasing time-step (Figure 3). However,

310 the posterior distributions of the first two time-steps (30 and 60 seconds)
311 differed strongly from those of the other four time-steps, all of which were
312 quite similar.

313 The mean of speed distributions of resting behaviour a_R decreased with in-
314 creasing time-step, because of decreasing estimated travel distance in the
315 interpolated trajectory. Mean turning angle distributions b_R and b_M were
316 constant for all time-steps. The modes of U_1 (the ratio between the covered
317 distance from the interpolated and raw data, within the interpolation time-
318 step) distributions α_R and α_M decreased with increasing time-step, because
319 the longer the time-step, the lower the fit between interpolated trajectory
320 and raw data. Mean U_2 for moving behaviour μ_M increased with increasing
321 time-step, where the turning angles were more acute (*e.g.*, travelling trajec-
322 tories in Figure 1). Mean U_2 for the resting behaviour μ_R was similar for all
323 time-steps.

324 Concerning the transition probability of maintaining resting behaviour $q_{R \rightarrow R}$,
325 posterior distributions were similar for the four longer time-steps (120 to
326 300 seconds), with a median value around 0.975. For the second tested
327 time-step (60 seconds), the median of the posterior distribution was lower
328 (around 0.95), and much lower for the shortest time-step of 30 seconds
329 (around 0.88). Concerning the transition probability of maintaining mov-
330 ing behaviour $q_{M \rightarrow M}$, posterior distributions were similar for the two longest
331 time-steps of 240 and 300 seconds, with a median value around 0.90, and for
332 the 30, 120 and 180 second time-steps, with a median value around 0.87. For
333 the 60 second time-step, the median value of the posterior distribution was
334 higher (around 0.93).

335 *Inferred behaviours.* Behaviours were highly discriminated for all tested time-
336 steps, although discrimination was a bit lower for the shortest time-step, of
337 30 seconds (Table 1). Using Model 2, inferred behaviours were similar for
338 all time-steps for the two travelling trajectories and the second stationary
339 trajectory (Figure 4). Stationary trajectory 1 showed a higher proportion of
340 resting behaviour for the four longest time-steps (120 to 300 seconds) (Figure
341 4); for the two shortest time-steps, some movements were associated with
342 moving behaviour, or else were not discriminated (Figure 5). Similarly, the
343 stationary phase at the end of the bi-type 1 trajectory was more associated
344 with moving behaviour (or else not discriminated) for the two shortest time-
345 steps (Figure 5). Concerning the bi-type 2 trajectory, a higher proportion
346 of resting behaviour was inferred for the four longest time-steps (Figure 4),
347 due to the increased number of locations in the stationary phase at the end
348 of the trajectory (Figure 5).

349 *Simulated data analysis.* For the three longest time-steps (180, 240 and 360
350 seconds), the models (Model 1 and Model 2) failed to converge and dis-
351 criminate between behaviours. For raw data, 99% of the inferred behaviours
352 matched the simulated ones (RMSE of 0.01) (Appendix D Table D1). For the
353 three shortest tested time-steps (30, 60 and 120 seconds), RMSE increased
354 with time-step. For the two shortest time-steps (30 and 60 seconds), RMSE
355 was lower on Model 2 than on Model 1. Part of the trajectory, with true and
356 inferred behaviours for the three shortest tested time-steps (30, 60 and 120
357 seconds), is shown in Appendix D, Figure D1.

358 *3.2. Behaviour discrimination with the state-space model*

359 *3.2.1. Interest of the two additional variables*

360 For stationary trajectories ("S"), U_1 was lower than for travelling trajec-
361 tories ("T"), and showed average values for bi-type trajectories ("B") (Figure
362 2 (a)). This was due to the higher rate of star-shaped trajectories when the
363 individual was stationary: in star-shaped trajectories, distance covered was
364 shorter on interpolated than raw data, due to the triangulation artefact.
365 U_2 was lower for travelling than stationary trajectories and intermediate for
366 bi-type trajectories (Figure 2 (b)). The differences shown by U_1 and U_2 be-
367 tween the different types of trajectory provide information for discriminating
368 "resting" versus "moving" behaviour related to "stationary" and "travelling"
369 trajectory types, respectively. Behaviours were slightly better discriminated
370 (*i.e.*, were in most cases inferred as being either resting or moving), with the
371 addition of U_1 and U_2 (Table 1).

372 Finally, we fitted the model on all the 731 available trajectories, except for
373 those comprising fewer than 5 locations. At first, neither of the two models
374 (Model 1 and Model 2) was able to converge. We therefore added a third
375 behaviour in the models. This third behaviour was an intermediate between
376 the resting and moving behaviours, with mean turning angle close to that
377 of the resting behaviour, and mean speed between the mean speeds of the
378 other two behaviours. With the third behaviour, Model 2 converged whereas
379 Model 1 still failed to converge.

380 *3.2.2. Fit between inferred behaviours and observed trajectories*

381 Overall, with Model 2, inferred behaviours were consistent with observed
382 trajectories: moving behaviours were mostly inferred in travelling trajec-

383 ries, and resting behaviours in stationary trajectories (Figure 4). For the
384 bi-type 2 trajectory, the travelling (starting) and stationary (ending) phases
385 mostly corresponded to moving and resting behaviours, respectively, as ex-
386 pected. Some movements in stationary phases were, however, associated with
387 moving behaviour, or with resting behaviour with low certainty, for the two
388 shortest time-steps. Increasing time-step increased mean duration for each
389 behaviour (Table 2), and also increased the difference in duration between
390 the behaviours: the longer the time-step, the longer the resting duration,
391 compared to the moving duration.

392 4. Discussion

393 4.1. Influence of time-step on behaviour discrimination and computation 394 time

395 A wide range of durations between two consecutive locations (*i.e.*, time-
396 steps) are used in animal movement studies: for example from 60 seconds
397 (eels, (Bassett and Montgomery 2011)) to several minutes (caribou, 15 min-
398 utes (Andersen et al. 2017)), one hour (sea lions, (Breed et al. 2012)), or
399 several hours (turtles, 6 hours (Bailey et al. 2008), wolves, 12 hours (Franke et al. 2006)).
400 Time-steps should be adapted to the species, its travel mode and the question
401 being addressed. For resident fish moving in small areas, time-steps shorter
402 than several minutes are necessary to explore the different behaviours. We
403 observed progressive but not major differences between variables (*i.e.*, the
404 movement variables mean speed and turning angles (Appendix A, Figure
405 A2 (a) and (b)) and the additional variables U_1 and U_2) used to evalu-
406 ate the tested interpolation time-steps. As of the second tested time-step

407 (60 seconds), behaviours were better discriminated. However, with time-
408 steps longer than 120 seconds, interpolation generated trajectories that were
409 very remote from the raw data, considerably degrading location information.
410 This can cause problems: for instance, when the individual often switches
411 behaviour or when spatial data such as environmental parameters are to
412 be included in the analysis. Likewise, the differences in mean duration per
413 behaviour according to time-step showed that switching phases may be over-
414 looked if a long time-step is used and, consequently, information on species
415 ecology can be lost due to unreliable interpolation. Furthermore, model fit
416 to the data (speed and turning angles; data not shown) was better for the
417 two shortest time-steps. This is supported by the analysis of simulated data,
418 which gave satisfactory behaviour predictions for the two shortest time-steps
419 (30 and 60 seconds), poorer prediction for 120 seconds, and no discrimina-
420 tion between behaviours for the three longest time-steps (180, 240, and 360
421 seconds).

422 Overall, computation time (Appendix B, Table B2) was reasonable. There
423 were no large differences between the time-steps from 60 to 240 seconds.
424 However, computation time was almost twice as long for the shortest (30
425 seconds) as for the longest (300 seconds) time-step. For simulated data,
426 computation time was almost 12-fold longer for raw data (3 seconds) than
427 for the shortest time-step (30 seconds) and 2.5-fold longer for the 30-second
428 than the 60-second time-step (Appendix D, Table D2). Such a difference can
429 be critical when analysing all trajectories of all individuals together, even
430 though data generated by interpolation made the model inference feasible,
431 in terms of computation time, for all tested time-steps.

432 *4.2. Model results and behaviour discrimination*

433 The results concerning the movement descriptors (*i.e.*, speed and turning
434 angle distributions) in resting behaviour should be regarded with caution.
435 Most artefactual trajectories that are "star-shaped" are stationary phases
436 which mainly include resting behaviour. Nevertheless, interpolation with
437 all tested time-steps reduced that artefact, enabling discrimination between
438 moving and resting behaviours. The high estimated transition probabilities
439 $q_{R \rightarrow R}$ and $q_{M \rightarrow M}$ (above 0.80) imply that behaviours presented long dura-
440 tion, and are clearly influenced by our choices of "exemplary" trajectories
441 displaying contrasted movement phases.

442 In order to check whether behaviours were adequately discriminated accord-
443 ing to the tested time-steps, the selected trajectories had to be composed of
444 phases that clearly represent the two main behaviours. As a consequence,
445 some other patterns of movement were firstly discarded. When all trajec-
446 tories were included (even excluding those with fewer than 5 locations), the
447 simple model with two behaviours did not converge, partly due to the pres-
448 ence of these other movement patterns. In addition, we assumed a constant
449 transition matrix, which is too simplistic since fish behaviour is influenced by
450 environmental factors. This might also prevent the model from converging.
451 Further validation with all data is needed, but a simple model tested on a
452 set of sample trajectories is still useful to determine the pre-processing to be
453 performed on the data.

454 *Additional variables.* Focusing on the 16 selected trajectories, including the
455 two variables U_1 and U_2 in the model improved discrimination of behaviours.
456 The simulated data analysis showed that including the additional variables

457 reduced prediction error for the two shortest time-steps (30 and 60 seconds).
458 Thus, it appears highly profitable to combine these kinds of variable, cal-
459 culated from raw data, with a short time-step to compensate the deficit
460 in behaviour discrimination. Furthermore, using all trajectories (except for
461 those with fewer than 5 locations) confirmed the benefit of including the
462 additional variables U_1 and U_2 for a given interpolation time-step (here, 60
463 seconds). Other additional variables could have been considered such as
464 specific indicators related to the known behaviours of the studied species.

465 *4.3. Implications of the interpolation and trajectory cutting processes for be-*
466 *haviour discrimination*

467 Tools from earlier studies of state-space models (*e.g.*, (Johnson et al. 2008,
468 Jonsen et al. 2003, Vermard et al. 2010)) attempted to deal with irregular
469 time-steps, location errors or the reconstruction of entire trajectories. Such
470 tools could have been appropriate to process the present data, but the prime
471 issue was computing time, and increasing the time-step solved this while also
472 dealing with most of the numerous small gaps in location. Combined with
473 the use of the two additional variables U_1 and U_2 , the proposed processing
474 reduced computing time, dealt with irregularly time-spaced locations and
475 preserved the information provided by the initial 3-second- time-step of the
476 raw data. For the present study, several smaller unconnected trajectories
477 provided enough information, and we did not seek to determine individual
478 locations and behaviours when the signal had been lost for a long period of
479 time (several hours). To achieve convergence with the 3-behaviour model,
480 the shortest trajectories (fewers than 5 points) were deleted from the dataset.
481 This necessity might be due to a bias in estimating transition probabilities

482 for these very short trajectories. An initial behaviour (namely z_{init} , see Ap-
483 pendix C, Model code) was mandatory to initiate the Markov chain of be-
484 haviours across time: z_{init} was derived from a categorical [distribution](#) with
485 the 3 equal probabilities. When short trajectories are numerous, the esti-
486 mated transition probability from the initial behaviour to the first behaviour
487 of the trajectory has a major weight, although it is only an "artefact" of the
488 modelling procedure.

489 As short trajectories may bias the estimation of transition probabilities, a
490 more suitable dataset would favour long trajectories (*i.e.*, with the maximum
491 number of points). For this, two possibilities emerged. Firstly, long time-
492 steps could be used, so that the raw dataset is less divided: trajectories are
493 longer in time, and also likely made up of numerous points. However, tra-
494 jectories shorter in time than twice the chosen time-step were deleted due to
495 interpolation, because at least two movements are needed to calculate speed
496 and turning angle. These trajectories would be preserved with a smaller
497 time-step. Secondly, a short time-step could be preferred, in which case tra-
498 jectories are denser in points, but the cutting process of the raw data to
499 obtain trajectories excludes more points than with a long time-step, because
500 the s threshold is lower.

501 The cutting process and the chosen s threshold are thus appear key points
502 in data pre-processing. The present cutting method has two main disad-
503 vantages. [Firstly, the different time-steps were difficult to compare, as the](#)
504 [number of locations differed between the tested time-steps.](#) Secondly, with a
505 small threshold, trajectories are liable to be small, which could lead to over-
506 looking some switches in behaviour. One solution could be to use a higher

507 threshold, chosen on the basis of the histogram of trajectory durations, com-
508 bined with a small time-step. But this solution would lead to interpolating
509 data on time ranges for which no observed locations are available, which
510 is not desirable. An auto-correlogram of covered distances in the raw data
511 could also give indications on where to cut the raw data to obtain trajec-
512 tories. Missing locations could also be considered as missing data, instead of
513 interpolating them. In the present case, this was not possible because us-
514 ing a Wrapped Cauchy distribution for turning angles within JAGS required
515 using observed turning angles as input rather than observed variable (see An-
516 nex C Model code and BUGS trick in (Morales et al. 2004), Supplementary
517 Information).

518 *4.4. Outline*

519 The present study developed a generic method of data pre-processing to
520 handle trajectories at fine time scales and infer behaviours based on teleme-
521 try data, which could be transposed to other datasets. Data pre-processing
522 is an essential step in trajectory analysis, although rarely highlighted. In-
523 creasing time-steps allowed efficient discrimination between behaviours, with
524 locations regularly spaced in time and a smaller amount of data to process.
525 At the same time, the additional variables computed from the raw data com-
526 pensated for the loss of information in interpolated trajectories resulting from
527 the increased time-step. The time-step should be adapted according to the
528 ecology and habitat preference of the studied species. The raw data cutting
529 process should be explored to optimise trajectory length while preserving
530 small time-steps and correspondence between trajectories and raw data.
531 We demonstrated the possibility of discriminating behaviours for the whole

532 dataset of a given individual, using a state-space model. This opens up in-
533 teresting perspectives. Individual variation in movement or behavioural pa-
534 rameters could be quantified using the trajectories of several individuals, by
535 hierarchical modelling (Jonsen 2016) using a 3-behaviour state-space model
536 including additional variables computed from the raw data. The reduction in
537 computation time is then particularly valuable when all trajectories of several
538 individuals are included. Datasets for several individuals of various species
539 have more sources of variability. Hierarchical modelling is then necessary
540 to deal with individual and species variability. Environmental variability re-
541 quires explicit modelling of the link between environmental variables, such
542 as hydraulic or thermal parameters, and the transition matrix. **Considering**
543 **that i) behavioural state and habitat selection are linked, and ii) movement**
544 **and trajectory changes are behavioural state proxy, movement analysis is a**
545 **way of investigating dynamical selection of habitat (short term, below several**
546 **minutes) in a highly contrasted and variable (below the hour) environment.**
547 **Therefore,** dynamically favourable habitats (according to dynamic hydraulic
548 conditions) should be mapped, providing quantitative information to evalu-
549 ate the impact of events such as thermal discharge or dam functioning.

550 **5. Acknowledgements**

551 We thank the Agence de l'Eau Rhône-Méditerranée-Corse, Electricité de
552 France (EDF-DTG and EDF R&D), the European Union/FEDER and the
553 Aquitaine Region for their financial support. This study is part of a part-
554 nership research program in hydrobiology (HYNES) between EDF R&D and
555 IRSTEA.

556 **6. Figures titles and captions**

557 *6.1. Figure 1*

558 *Title.* Six trajectories and raw data superimposed, for all the tested time-
559 steps.

560 *Caption.* Gray triangles are the raw data and black dots are the interpolated
561 locations (linked by black lines) for each time-step (*i.e.*, the trajectories).
562 These trajectories were chosen because they are characteristic of a stationary
563 behaviour (Stationary 1 and 2), of an active travelling behaviour (Travelling
564 1 and 2), or of an alternation between these two behaviours (Bi-type 1 and
565 2).

566 *6.2. Figure 2*

567 *Title.* Distributions of the additional variables U_1 (a) and U_2 (b), for all the
568 tested time-steps.

569 *Caption.* U_1 is the ratio between covered distances according to interpolated
570 data and covered distances according to raw data. U_2 is the variance of
571 turning angles of raw data within the interpolation time-step.

572 *6.3. Figure 3*

573 *Title.* Posterior distributions of the parameters, for all the tested time-steps,
574 from Model 2.

575 *6.4. Figure 4*

576 *Title.* Percentages of each inferred behaviour for six trajectories, for all the
577 tested time-steps (computed from Model 2).

578 *Caption.* Resting and Moving correspond to locations where resting and
579 moving behaviour respectively was inferred with probability > 0.8 (*i.e.*, for
580 80% of the MCMC iterations the estimated behaviour was resting/moving).
581 Hesitating corresponds to locations where behaviours were inferred with
582 probability < 0.8 .

583 *6.5. Figure 5*

584 *Title.* Six trajectories with inferred behaviour, for all the tested time-steps
585 (computed from Model 2).

586 *Caption.* These trajectories were chosen because they are characteristic of a
587 stationary behaviour (Stationary 1 and 2), of an active travelling behaviour
588 (Travelling 1 and 2), or of an alternation between these two behaviours (Bi-
589 type 1 and 2).

590 **7. Appendix A: Figures**

591 *7.1. Figure A1*

592 *Title.* Example of a star-shaped trajectory.

593 *7.2. Figure A2 (a) and (b)*

594 *Title.* Distributions of movement variables (speed and turning angles), for
595 all the tested time-steps.

596 *7.3. Figure A3*

597 *Title.* The 16 selected trajectories with inferred behaviour (red: moving,
598 blue: resting), for the 60-second time-step (computed from Model 2).

599 **8. Appendix B: Table of model parameters**

600 **9. Appendix C: JAGS code for Model 2**

601 **10. Appendix D: Simulated data analysis**

602 *10.1. Tables 1 and 2, Appendix D*

603 *10.2. Figure D1*

604 *Title.* Portion of the simulated trajectory with true (triangles) and predicted
605 (circles) behaviours for the three shortest tested time-steps (30, 60 and 120
606 seconds, from top to bottom), with Model 1 (left side) and Model 2 (right
607 side). Red stands for moving behaviour, and blue for resting.

608 **11. References**

609 [Andersen et al. 2017] K. H. Andersen, A. Nielsen, U. H. Thygesen, H. H.
610 Hinrichsen, and S. Neuenfeldt. Using the particle filter to geolocate
611 Atlantic cod (*Gadus morhua*) in the Baltic Sea, with special emphasis
612 on determining uncertainty. *Canadian Journal of Fisheries and Aquatic*
613 *Sciences*, 64(4):618–627, 2007.

614 [Bailey et al. 2008] H. Bailey, G. Shillinger, D. Palacios, S. Bograd,
615 J. Spotila, F. Paladino, and B. Block. Identifying and comparing phases
616 of movement by leatherback turtles using state-space models. *Journal*
617 *of Experimental Marine Biology and Ecology*, 356(1-2):128–135, 2008.

618 [Bassett and Montgomery 2011] D. Bassett and J. Montgomery. Home range
619 use and movement patterns of the yellow moray eel *Gymnothorax pras-*
620 *inus*. *Journal of Fish Biology*, 79(2):520–525, 2011.

- 621 [Bergé et al. 2012] J. Bergé, H. Capra, H. Pella, T. Steig, M. Ovidio, E. Bul-
622 tel, and N. Lamouroux. Probability of detection and positioning error of
623 a hydro acoustic telemetry system in a fast-flowing river: Intrinsic and
624 environmental determinants. *Fisheries Research*, 125-126:1–13, 2012.
- 625 [Bestley et al. 2013] S. Bestley, I. D. Jonsen, M. A. Hindell, C. Guinet, and
626 J. B. Charrassin. Integrative modelling of animal movement: incorpor-
627 ating in situ habitat and behavioural information for a migratory ma-
628 rine predator. *Proceedings of The Royal Society B: Biological sciences*,
629 280(1750):2012–2262, 2013.
- 630 [Block et al. 2011] B. A. Block, I. D. Jonsen, S. J. Jorgensen, a. J. Winship,
631 S. a. Shaffer, S. J. Bograd, E. L. Hazen, D. G. Foley, G. a. Breed, A.-L.
632 Harrison, J. E. Ganong, A. Swithenbank, M. Castleton, H. Dewar, B. R.
633 Mate, G. L. Shillinger, K. M. Schaefer, S. R. Benson, M. J. Weise, R. W.
634 Henry, and D. P. Costa. Tracking apex marine predator movements in
635 a dynamic ocean. *Nature*, 475(7354):86–90, 2011.
- 636 [Boucek et al. 2017] R. E. Boucek, M. R. Heithaus, R. Santos, P. Stevens,
637 and J. S. Rehage. Can animal habitat use patterns influence their vul-
638 nerability to extreme climate events? An estuarine sportfish case study.
639 *Global Change Biology*, (April):1–13, 2017.
- 640 [Breed et al. 2012] G. A. Breed, D. P. Costa, I. D. Jonsen, P. W. Robinson,
641 and J. Mills-Flemming. State-space methods for more completely cap-
642 turing behavioral dynamics from animal tracks. *Ecological Modelling*,
643 235-236:49–58, 2012.

- 644 [Bultel et al. 2014] E. Bultel, E. Lasne, A. Acou, J. Guillaudeau, C. Bertier,
645 and E. Feunteun. Migration behaviour of silver eels (*Anguilla anguilla*)
646 in a large estuary of Western Europe inferred from acoustic telemetry.
647 *Estuarine, Coastal and Shelf Science*, 137(1):23–31, 2014.
- 648 [Cagnacci et al. 2010] F. Cagnacci, L. Boitani, R. A. Powell, and M. S.
649 Boyce. Animal ecology meets GPS-based radiotelemetry: a perfect
650 storm of opportunities and challenges. *Philosophical transactions of the*
651 *Royal Society of London. Series B, Biological sciences*, 365(1550):2157–
652 62, 2010.
- 653 [Capra et al. 2017] H. Capra, L. Plichard, J. Bergé, H. Pella, M. Ovidio,
654 E. McNeil, and N. Lamouroux. Fish habitat selection in a large hy-
655 dropeaking river: Strong individual and temporal variations revealed by
656 telemetry. *Science of the Total Environment*, 578:109–120, 2017.
- 657 [Capra et al. 2018] H. Capra H., H. Pella, and M. Ovidio. Individual move-
658 ments, home ranges and habitat use by native rheophilic cyprinids and
659 non-native catfish in a large regulated river. *Fisheries Management and*
660 *Ecology*, 25:136–149, 2018.
- 661 [Carraro et al. 2017] L. Carraro, L. Mari, M. Gatto, A. Rinaldo, and E.
662 Bertuzzo. Spread of proliferative kidney disease in fish along stream
663 networks: A spatial metacommunity framework. *Freshwater Biology*,
664 (March), 2017.
- 665 [Cooke et al. 2004] S. J. Cooke, C. M. Bunt, and J. F. Schreer. Understand-
666 ing fish behavior, distribution, and survival in thermal effluents using

- 667 fixed telemetry arrays: A case study of smallmouth bass in a discharge
668 canal during winter. *Environmental Management*, 33:140–150, 2004.
- 669 [Cooke et al. 2013] S. J. Cooke, J. D. Midwood, J. D. Thiem, P. Klimley,
670 M. C. Lucas, E. B. Thorstad, J. Eiler, C. Holbrook, and B. C. Ebner.
671 Tracking animals in freshwater with electronic tags: past, present and
672 future. *Animal Biotelemetry*, 1(5):1–19, 2013.
- 673 [Dingle 1996] H. Dingle. *Migration: the Biology of Life on the Move*. New
674 York: Oxford University Press.
- 675 [Donaldson et al. 2014] M. R. Donaldson, S. G. Hinch, C. D. Suski, A. T.
676 Fisk, M. R. Heupel, and S. J. Cooke. Making connections in aquatic
677 ecosystems with acoustic telemetry monitoring. *Frontiers in Ecology
678 and the Environment*, 12: 565–573, 2014.
- 679 [Dorazio and Price 2019] R. M. Dorazio and M. Price. State-space models
680 to infer movements and behavior of fish detected in a spatial array of
681 acoustic receivers. *Canadian Journal of Fisheries and Aquatic Sciences*,
682 76(4):543–550, 2019.
- 683 [Drouineau et al. 2017] H. Drouineau, F. Bau, A. Alric, N. Deligne, P.
684 Gomes, and P. Sagnes. Silver eel downstream migration in fragmented
685 rivers: use of a Bayesian model to track movements triggering and du-
686 ration. *Aquatic Living Resources*, 30:1–9, 2017.
- 687 [Fraley et al. 2016] K. M. Fraley, J. A. Falke, R. Y., and S. Ivey. Seasonal
688 Movements and Habitat Use of Potamodromous Rainbow Trout Across

- 689 a Complex Alaska Riverscape. *Transactions of the American Fisheries*
690 *Society*, 145(5):1077–1092, 2016.
- 691 [Franke et al. 2006] A. Franke, T. Caelli, G. Kuzyk, and R. J. Hudson. Pre-
692 diction of wolf (*Canis lupus*) kill-sites using hidden Markov models. *Eco-*
693 *logical Modelling*, 197(1-2):237–246, 2006.
- 694 [Gelman and Rubin 1992] A. Gelman and D. B. Rubin. Inference from iter-
695 ative simulation using multiple sequences. *Statistical science*, 7(4):457–
696 511, 1992.
- 697 [Giuggioli and Bartumeus 2010] L. Giuggioli and F. Bartumeus. Ani-
698 mal movement, search strategies and behavioural ecology: a cross-
699 disciplinary way forward. *Journal of Animal Ecology*, 79: 906–909, 2010.
- 700 [Goodwin et al. 2014] R. A. Goodwin, M. Politano, J. W. Garvin, J. M.
701 Nestler, D. Hay, J. J. Anderson, L. J. Weber, E. Dimperio, D. L. Smith,
702 and M. Timko. Fish navigation of large dams emerges from their mod-
703 ulation of flow field experience. *Proceedings of the National Academy of*
704 *Sciences of the United States of America*, 111(14):5277–82, 2014.
- 705 [Hedger et al. 2008] R. D. Hedger, F. Martin, J. J. Dodson, D. Hatin, F.
706 Caron, and F. G. Whoriskey. The optimized interpolation of fish posi-
707 tions and speeds in an array of fixed acoustic receivers. *ICES Journal*
708 *of Marine Science*, 65(7):1248–1259, 2008.
- 709 [Hussey et al. 2015] N.E. Hussey, S. T. Kessel, K. Aarestrup, S. J. Cooke,
710 P. D. Cowley, A. T. Fisk, R. G. Harcourt, K. N. Holland, S. J. Iverson,
711 J. F. Kocik, J. E. M. Flemming, and F. G. Whoriskey. Aquatic animal

- 712 telemetry: A panoramic window into the underwater world. *Science*,
713 348(6240):1255642, 2015.
- 714 [Johnson et al. 2008] D. S. Johnson, J. M. London, M. A. L., and J. W.
715 Durban. Continuous-Time Correlated Random Walk Model for Animal
716 Telemetry Data. *Ecology*, 89(5):1208–1215, 2008.
- 717 [Jonsen et al. 2001] I. D. Jonsen, R. S. Bouchier, and J. Roland. The in-
718 fluence of matrix habitat on *Aphthona* flea beetle immigration to leafy
719 spurge patches. *Oecologia*, 127(2):287–294, 2001.
- 720 [Jonsen et al. 2003] I. D. Jonsen, R. A. Myers, and J. M. Flemming.
721 Meta-analysis of animal movement using state-space models. *Ecology*,
722 84(11):3055–3063, 2003.
- 723 [Jonsen 2016] I. D. Jonsen. Joint estimation over multiple individuals im-
724 proves behavioural state inference from animal movement data. *Sci Rep*,
725 6:20625, 2016.
- 726 [Joo et al. 2013] R. Joo, S. Bertrand, J. Tam, and R. Fablet. Hidden Markov
727 Models: The Best Models for Forager Movements? *PLoS ONE*, 8(8),
728 2013.
- 729 [Koehn and Nicol 2016] J. D. Koehn and S. J. Nicol. Comparative move-
730 ments of four large fish species in a lowland river. *Journal of Fish*
731 *Biology*, 88(4):1350–1368, 2016.
- 732 [Kranstauber and Smolla 2008] B. Kranstauber and M. Smolla. Visualizing
733 and Analyzing Animal Track Data. Rpackage version 2.1.0, 2008.

- 734 [Lima 2002] S. L. Lima. Putting predators back into behavioral predator-
735 prey interactions. *Trends in Ecology & Evolution*, 17(2):70–75, 2002.
- 736 [Lennox et al. 2017] R. J. Lennox, K. Aarestrup, S. J. Cooke, P. D. Cow-
737 ley, Z. D. Deng, A. T. Fisk, R. G. Harcourt, M. Heupel, S. G. Hinch,
738 K. N. Holland, N. E. Hussey, S. J. Iverson, S. T. Kessel, J. F. Kocik,
739 M. C. Lucas, J. M. Flemming, V. M. Nguyen, M. J.W. Stokesbury, S.
740 Vagle, D. L. VanderZwaag, F. G. Whoriskey, and N. Young. Envision-
741 ing the Future of Aquatic Animal Tracking: Technology, Science, and
742 Application. *BioScience*, 67:884–896, 2017.
- 743 [Mcclintock et al. 2012] B. T. Mcclintock, R. King, J. Matthiopoulos, B. J.
744 Mcconnell, and J. M. Morales. A General Modeling Framework for
745 Animal Movement in Discrete Time Using Multi-State Random Walks.
746 *Ecological Monographs*, 82(3):335–3349, 2012.
- 747 [Morales et al. 2004] J. M. Morales, D. T. Haydon, J. Frair, K. E. Holsinger,
748 and J. M. Fryxell. Extracting More Out of Relocation Data : Building
749 Movement Models As Mixtures of Random Walks. *Ecology*, 85(9):2436–
750 2445, 2004.
- 751 [Muhlfeld 2012] C. C. Muhlfeld, J. J. Giersch, and B. Marotz. Seasonal
752 movements of non-native lake trout in a connected lake and river system.
753 *Fisheries Management and Ecology*, 19(3):224–232, 2012.
- 754 [Nathan et al. 2008] R. Nathan, W. M. Getz, E. Revilla, M. Holyoak, R.
755 Kadmon, D. Saltz, and P. E. Smouse. A movement ecology paradigm

756 for unifying organismal movement research. *Pnas*, 105(49):19052–19059,
757 2008.

758 [Patterson et al. 2008] T. A. Patterson, L. Thomas, C. Wilcox, O.
759 Ovaskainen, and J. Matthiopoulos. State-space models of individual
760 animal movement. *Trends in Ecology and Evolution*, 23(2):87–94, 2008.

761 [Patterson et al. 2009] T. A. Patterson, M. Basson, M. V. Bravington, and
762 J. S. Gunn. Classifying movement behaviour in relation to environmental
763 conditions using hidden Markov models. *Journal of Animal Ecology*,
764 78(6):1113–1123, 2009.

765 [Pinder et al. 2005] A. C. Pinder, R. E. Gozlan, and J. R. Britton. Dispersal
766 of the invasive topmouth gudgeon, *Pseudorasbora parva* in the UK: A
767 vector for an emergent infectious disease. *Fisheries Management and*
768 *Ecology*, 12(6):411–414, 2005.

769 [Plummer 2009] M. Plummer. rjags: Bayesian graphical models using mcmc.
770 Rpackage version 1.0.3-12, 2009.

771 [Quaglietta and Porto 2019] L. Quaglietta and M. Porto. SiMRiv: An R
772 package for mechanistic simulation of individual, spatially-explicit mul-
773 tistate movements in rivers, heterogeneous and homogeneous spaces in-
774 corporating landscape bias. *Movement Ecology*, 7(1):1–9, 2019.

775 [R Core team 2018] R Core Team. *R: A Language and Environment for*
776 *Statistical Computing*. R Foundation for Statistical Computing, Vienna,
777 Austria, 2018.

- 778 [Roy et al. 2014] R. Roy, J. Beguin, C. Argillier, L. Tissot, F. Smith, S.
779 Smedbol S. and E. De Oliveira. Testing the VEMCO Positioning System:
780 spatial distribution of the probability of location and the positioning
781 error in a reservoir. *Animal Biotelemetry*, 2:1, 2014.
- 782 [Sutherland et al. 2015] C. Sutherland, A. K. Fuller, and J. A. Royle. Mod-
783 elling non-Euclidean movement and landscape connectivity in highly
784 structured ecological networks. *Methods in Ecology and Evolution*,
785 6(2):169–177, 2015.
- 786 [Tétard et al. 2016] S. Tétard, E. Feunteun, E. Bultel, R. Gadais, M. L.
787 Bégout, T. Trancart, and E. Lasne. Poor oxic conditions in a large
788 estuary reduce connectivity from marine to freshwater habitats of a di-
789 adromous fish. *Estuarine, Coastal and Shelf Science*, 169:216–226, 2016.
- 790 [Tétard et al. 2019] S. Tétard, A. Maire, M. Lemaire, E. De Oliveira, P. Mar-
791 tin, and D. Courret. Behaviour of Atlantic salmon smolts approaching a
792 bypass under light and dark conditions: importance of fish development.
793 *Ecological Engineering*, 131:39–52, 2019.
- 794 [Thiebault et al. 2018] A. Thiebault, L. Dubroca, R. Mullers, Y. Tremblay,
795 and P. Pistorius. "m2b" package in R: deriving multiple variables
796 from movement data to predict behavioural states with random forests..
797 *Methods in Ecology and Evolution*, 9:1548–1555, 2018.
- 798 [Vermard et al. 2010] Y. Vermard, E. Rivot, S. Mahévas, P. Marchal, and
799 D. Gascuel. Identifying fishing trip behaviour and estimating fishing

800 effort from VMS data using Bayesian Hidden Markov Models. *Ecological*
801 *Modelling*, 221(15):1757–1769, 2010.

802 [Whoriskey et al. 2019] K. Whoriskey, E. G. Martins, M. Auger-Méthé,
803 L. F.G. Gutowsky, R. J. Lennox, S. J. Cooke, M. Power, and J. Mills
804 Flemming. Current and emerging statistical techniques for aquatic
805 telemetry data: A guide to analysing spatially discrete animal detec-
806 tions. *Methods in Ecology and Evolution*, (February):1–14, 2019.

807 [Wuillez et al. 2016] M. Woillez, R. Fablet, T.-T. Ngo, M. Lalire, P. Lazure,
808 and H. de Pontual. A HMM-based model to geolocate pelagic fish from
809 high-resolution individual temperature and depth histories: European
810 sea bass as a case study. *Ecological Modelling*, 321:10–22, 2016.

Table 1: Percentage of locations where moving or resting behaviour was inferred with probability > 0.8 .

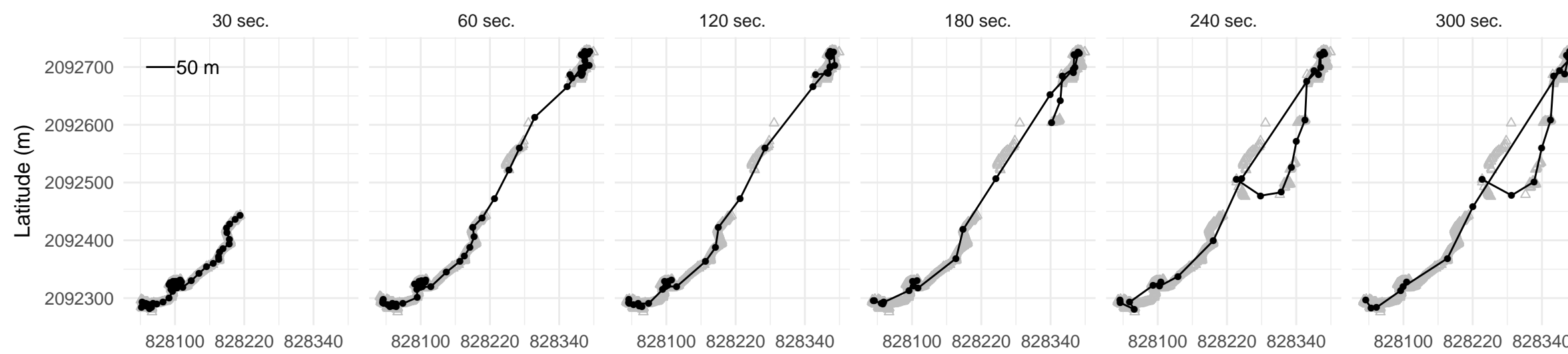
| Time-step (sec.) | Model 1, without U_1 and U_2 (%) | Model 2, with U_1 and U_2 (%) |
|------------------|--------------------------------------|-----------------------------------|
| 30 | 83 | 91 |
| 60 | 92 | 94 |
| 120 | 96 | 98 |
| 180 | 92 | 99 |
| 240 | 92 | 99 |
| 300 | 91 | 99 |

Table 2: Mean duration (minutes) in resting or moving behaviour according to the tested time-steps (computed from Model 2).

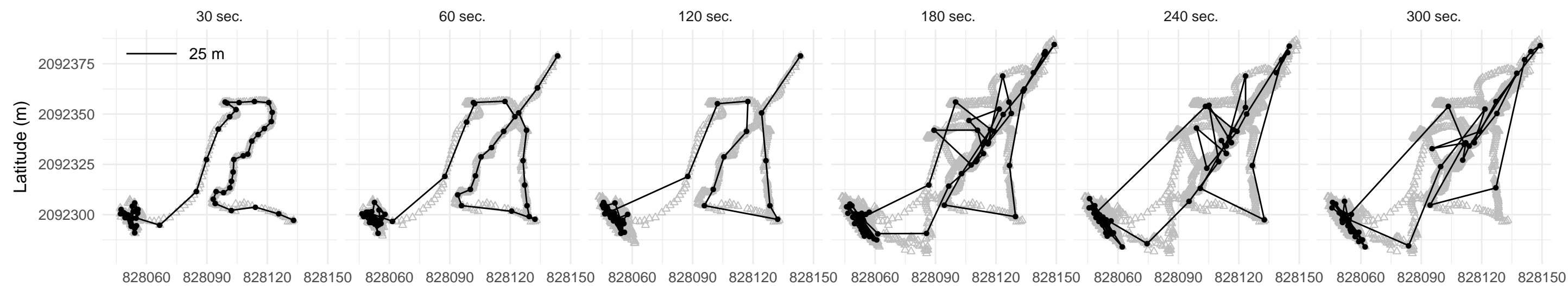
| Time-step (sec.) | Resting | Moving |
|------------------|----------------|----------------|
| 30 | 4.1 [3.2, 5.5] | 3.7 [2.9, 4.9] |
| 60 | 21 [14, 33] | 12 [8.8, 18] |
| 120 | 65 [43, 106] | 15 [11, 22] |
| 180 | 118 [74, 208] | 24 [17, 36] |
| 240 | 229 [125, 499] | 38 [25, 61] |
| 300 | 234 [129, 486] | 45 [30, 73] |

The mean duration (median and 95% credibility interval) in resting, or moving, behaviour was computed as the expected value (*i.e.*, the number of independent trials to get the first success) from the geometric distribution of probability equal to $1 - q_{R \rightarrow R}$, or $1 - q_{M \rightarrow M}$, multiplied by the time-step.

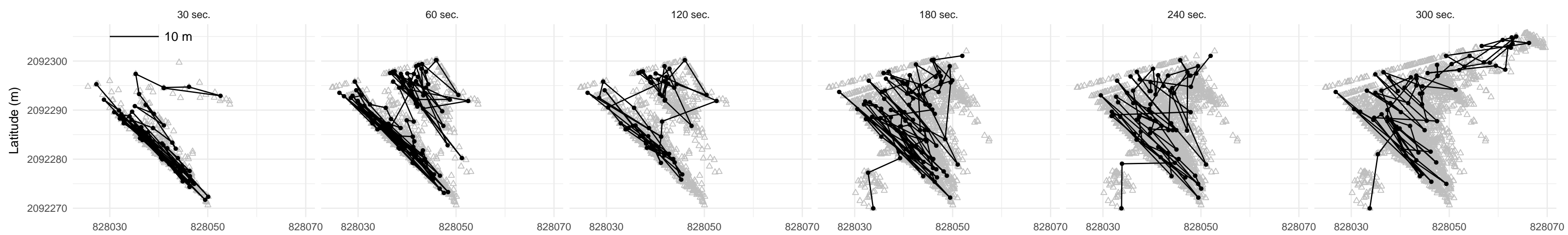
Bi-type 1



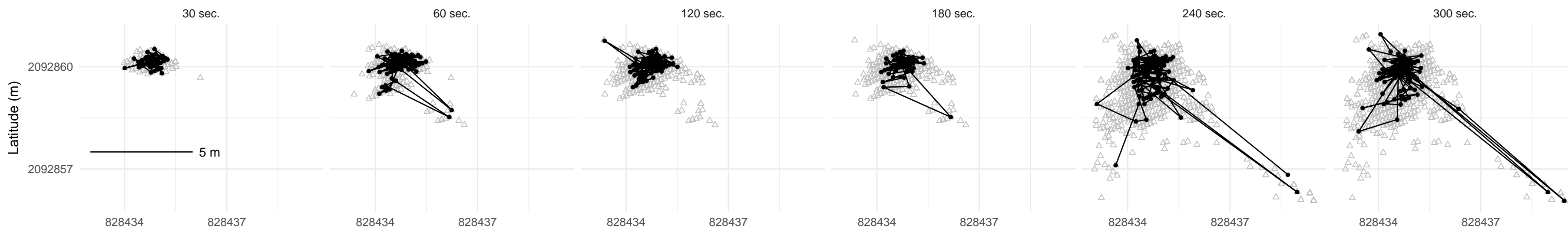
Bi-type 2



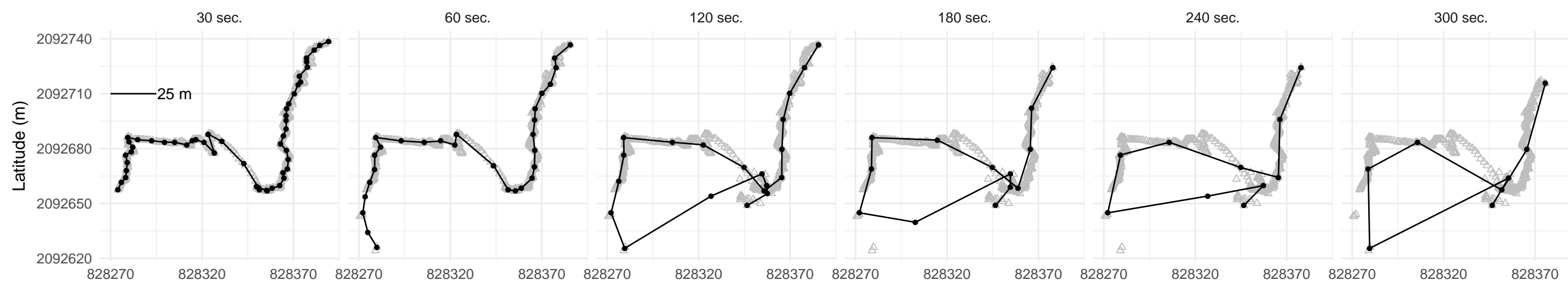
Stationary 1



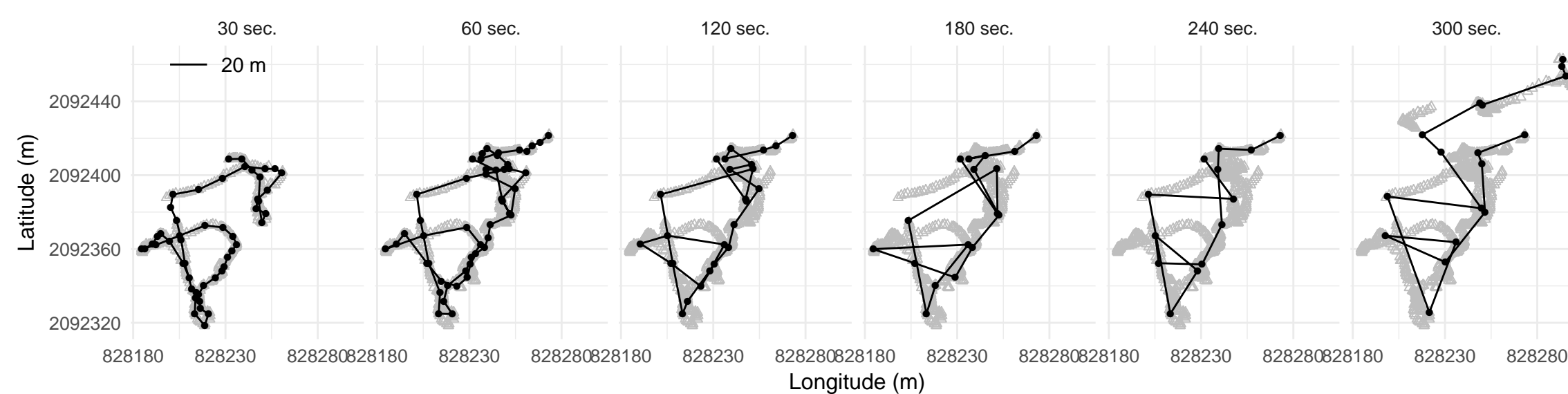
Stationary 2

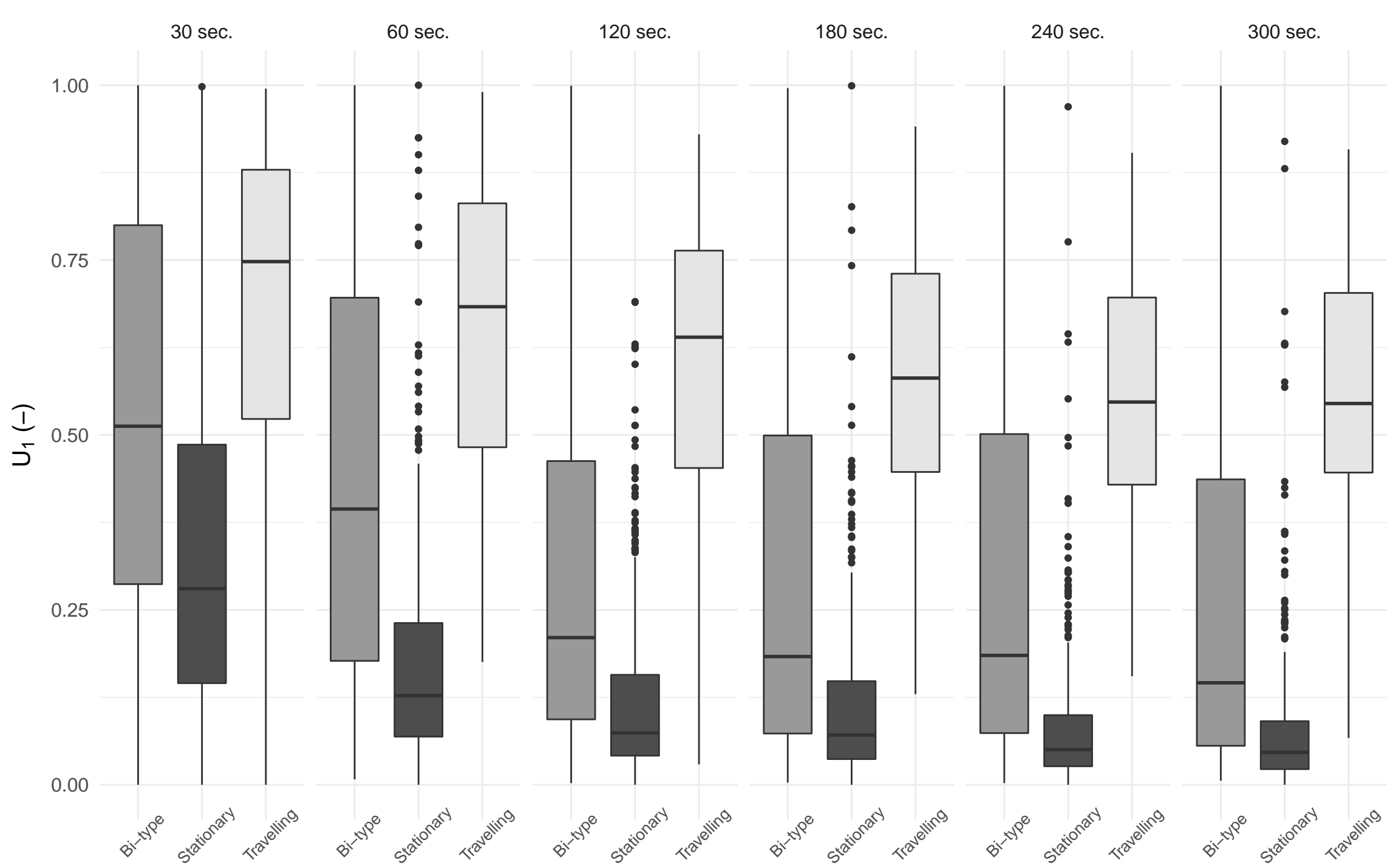


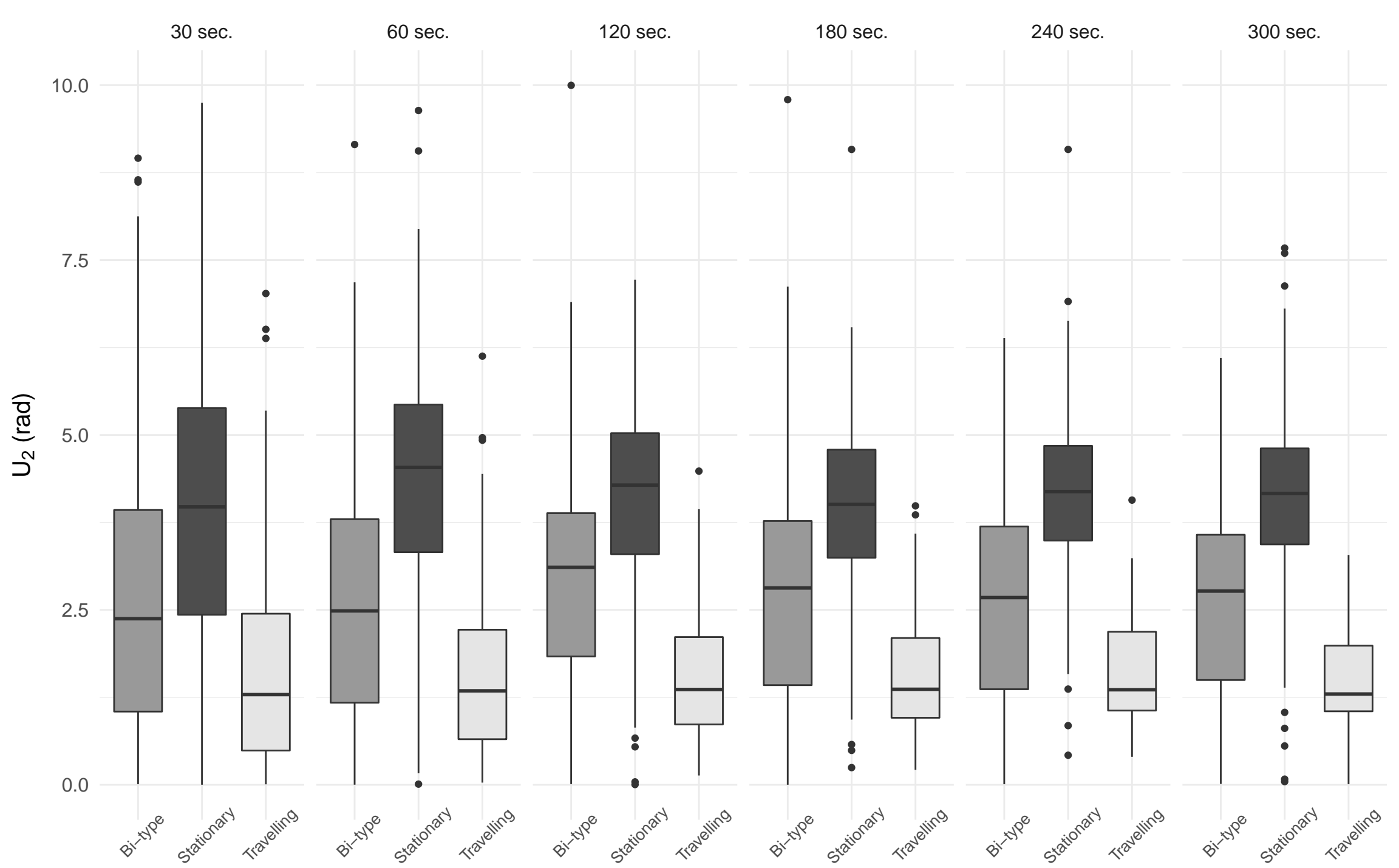
Travelling 1

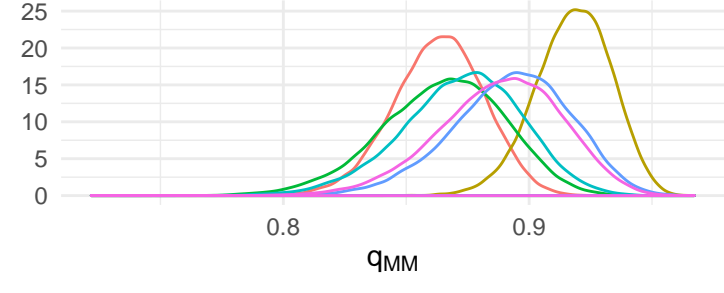
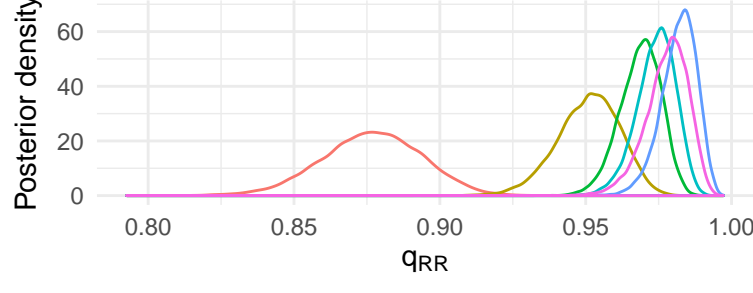
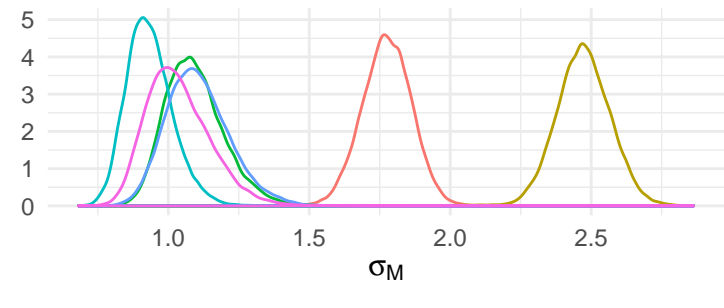
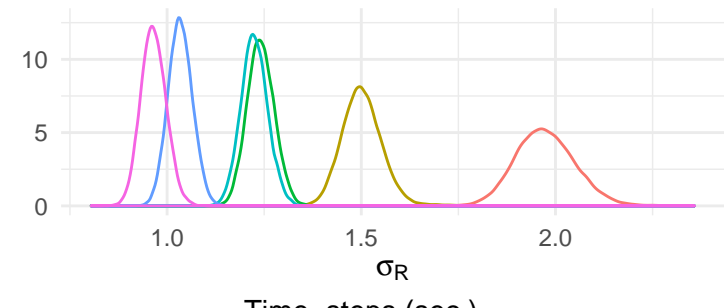
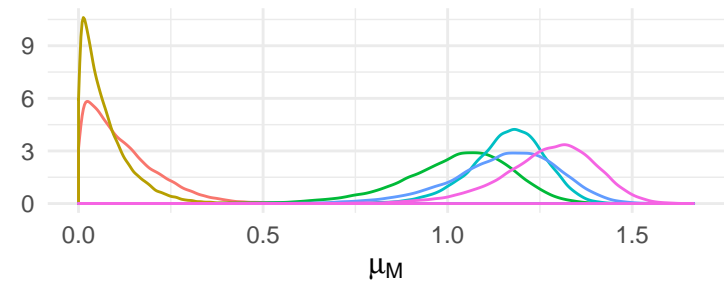
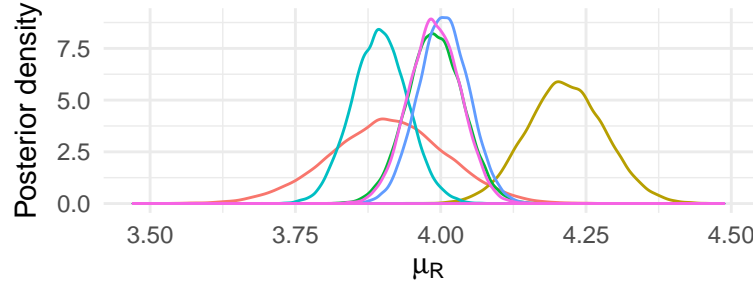
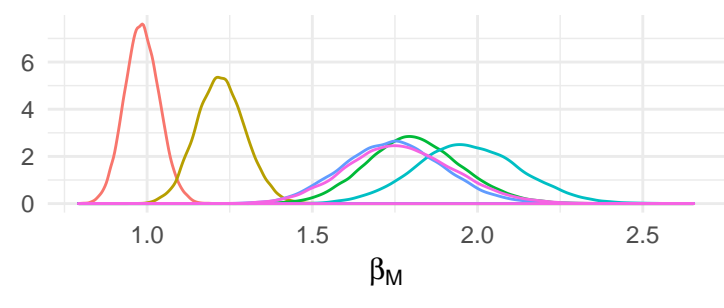
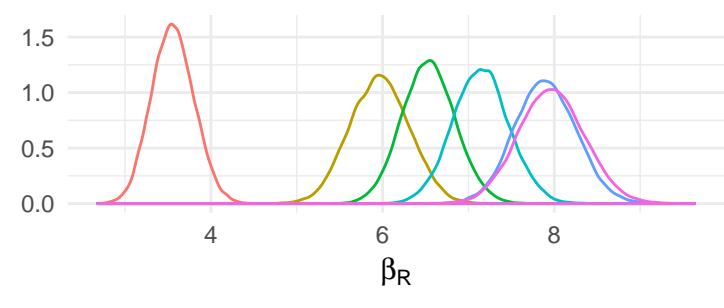
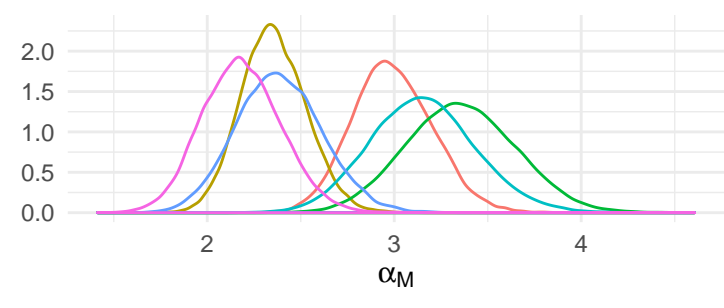
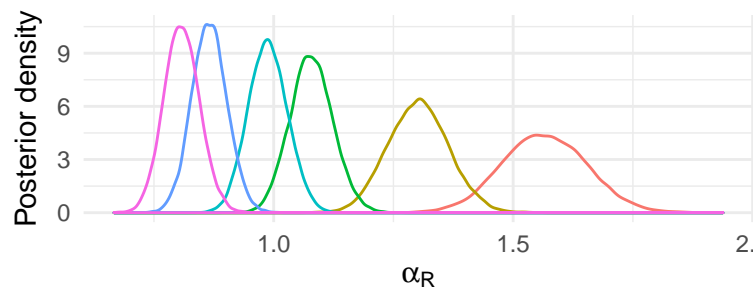
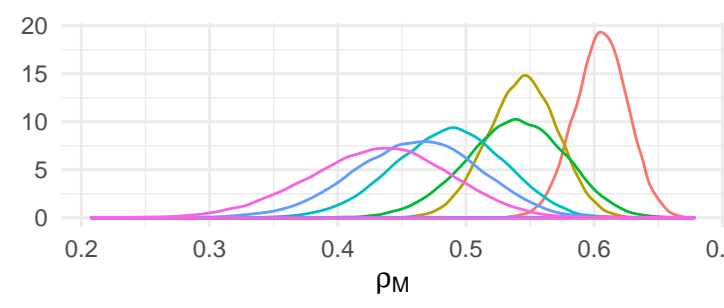
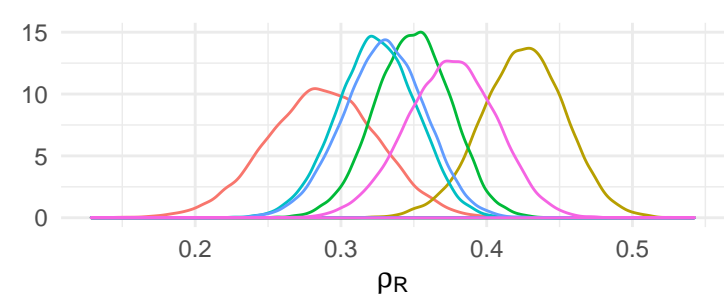
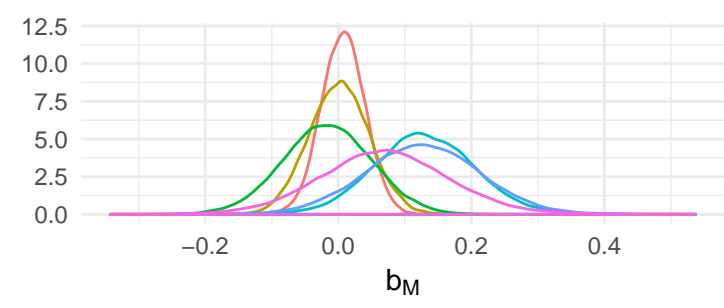
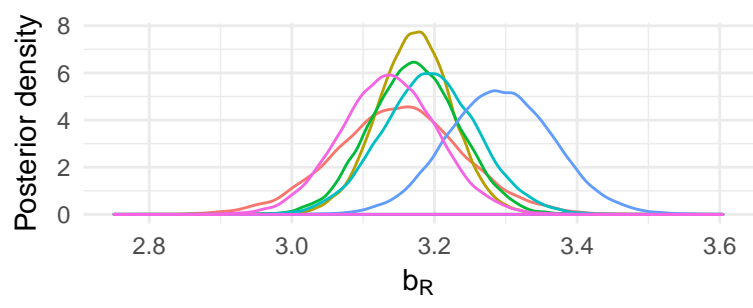
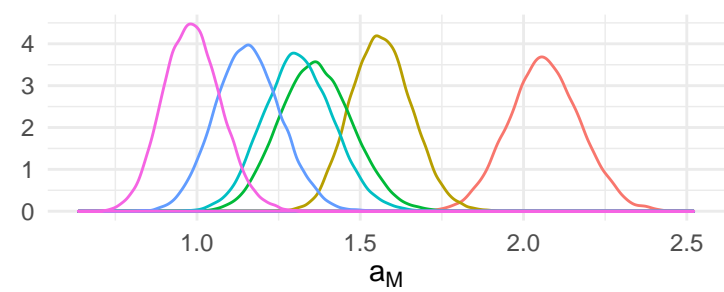
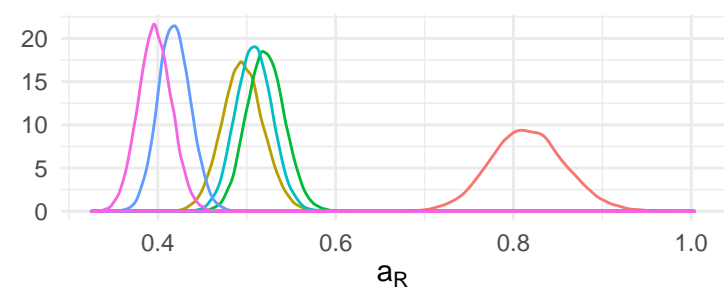
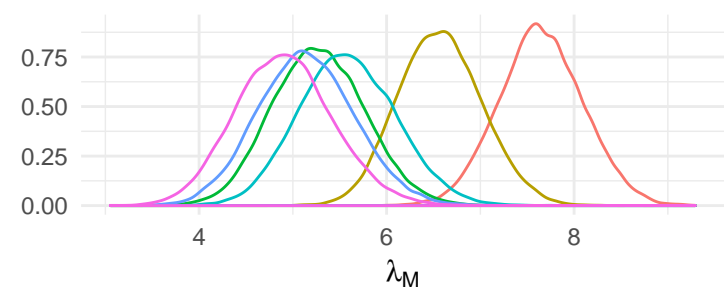
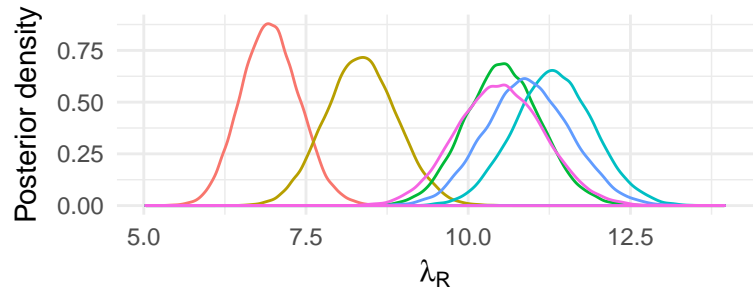


Travelling 2

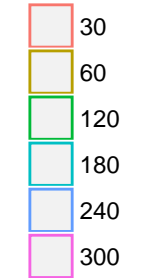


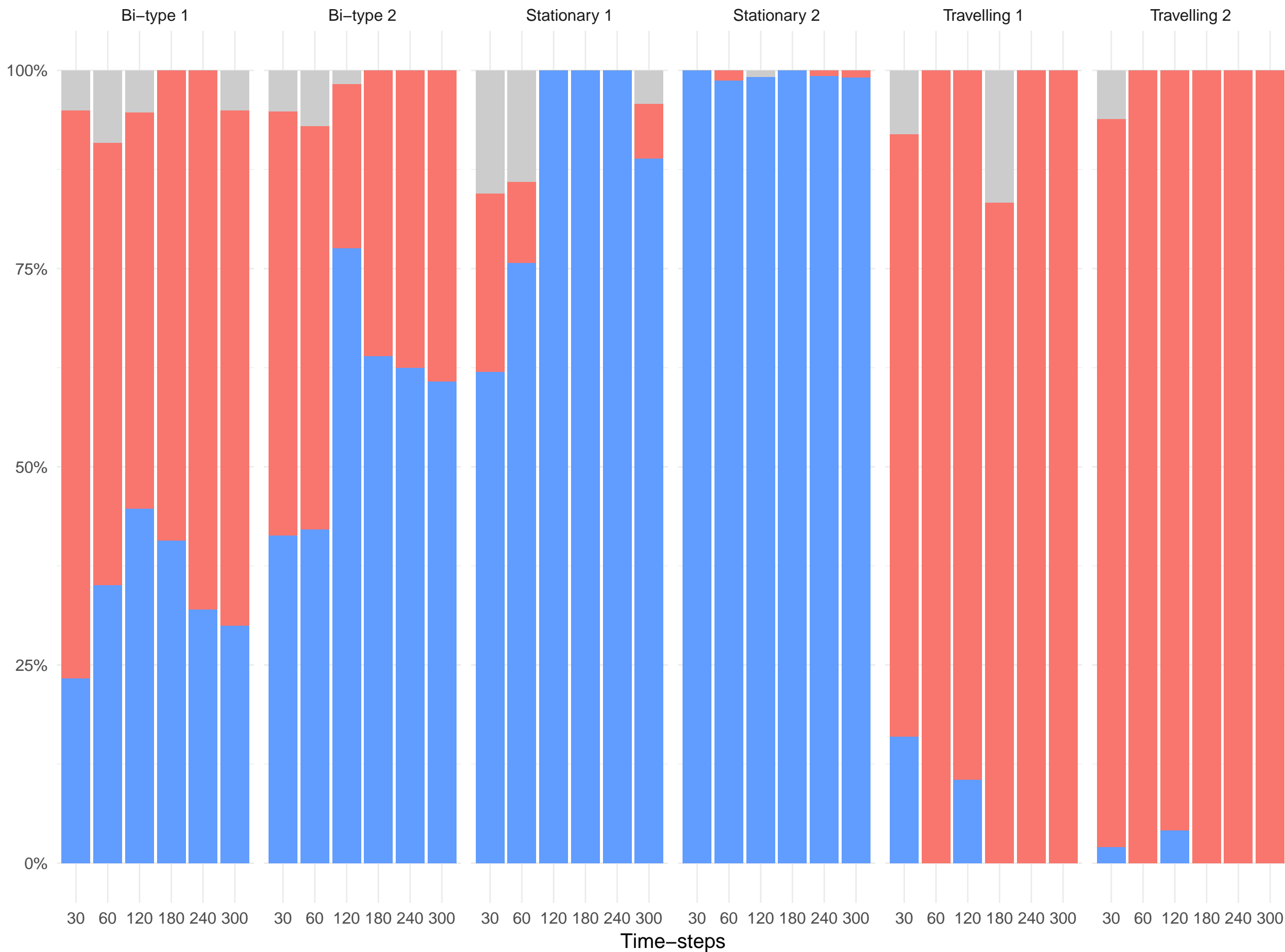






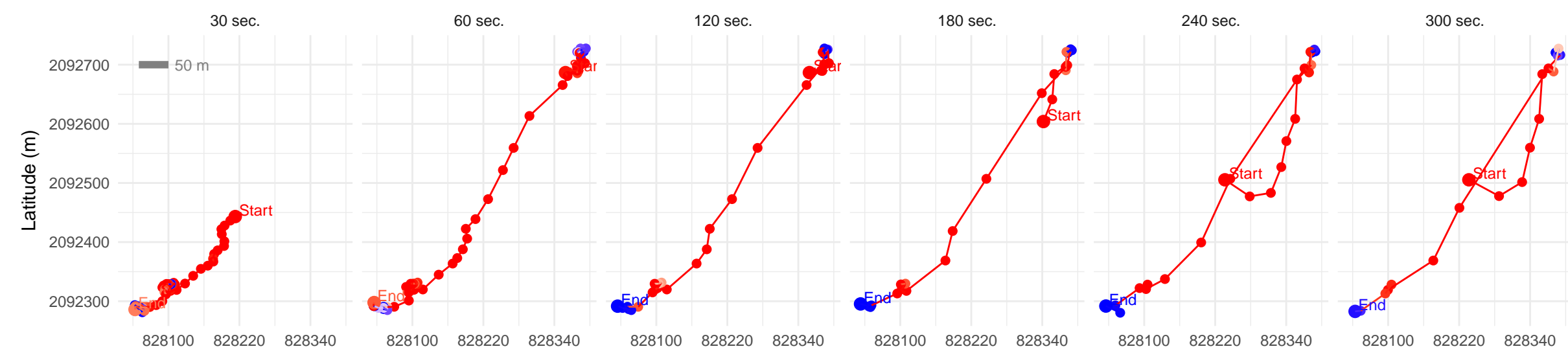
Time-steps (sec.)



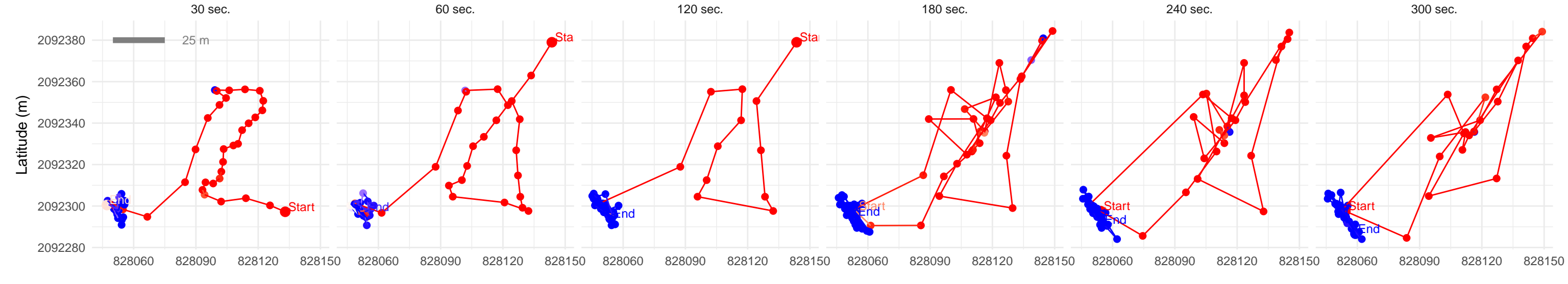


Behaviours
 Hesitating
 Moving
 Resting

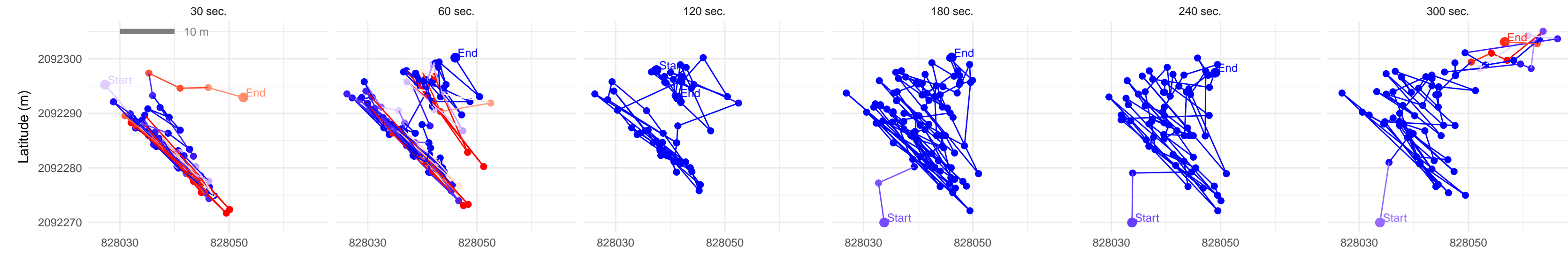
Bi-type 1



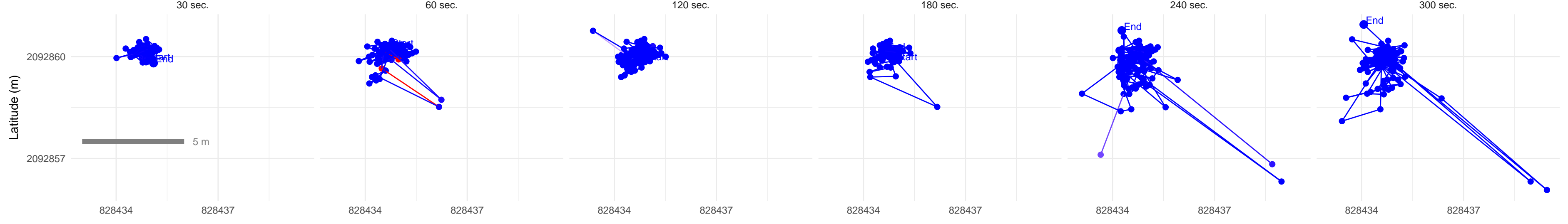
Bi-type 2



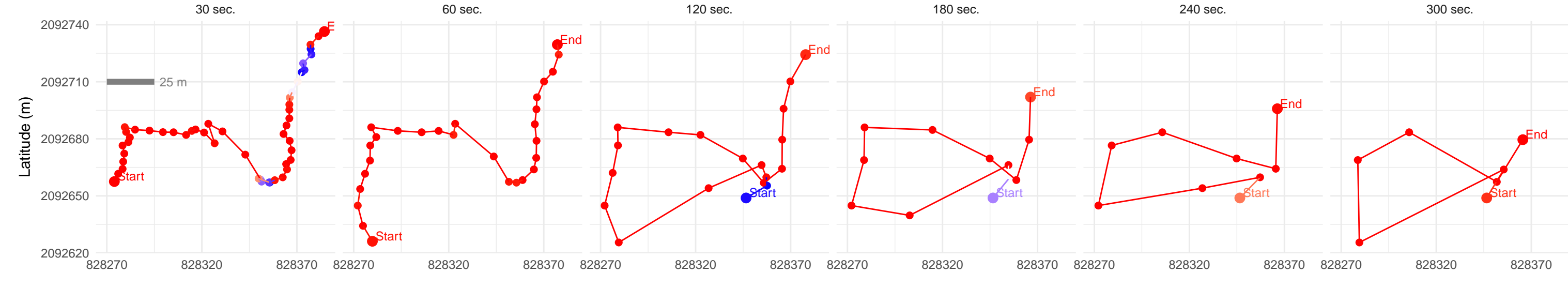
Stationary 1



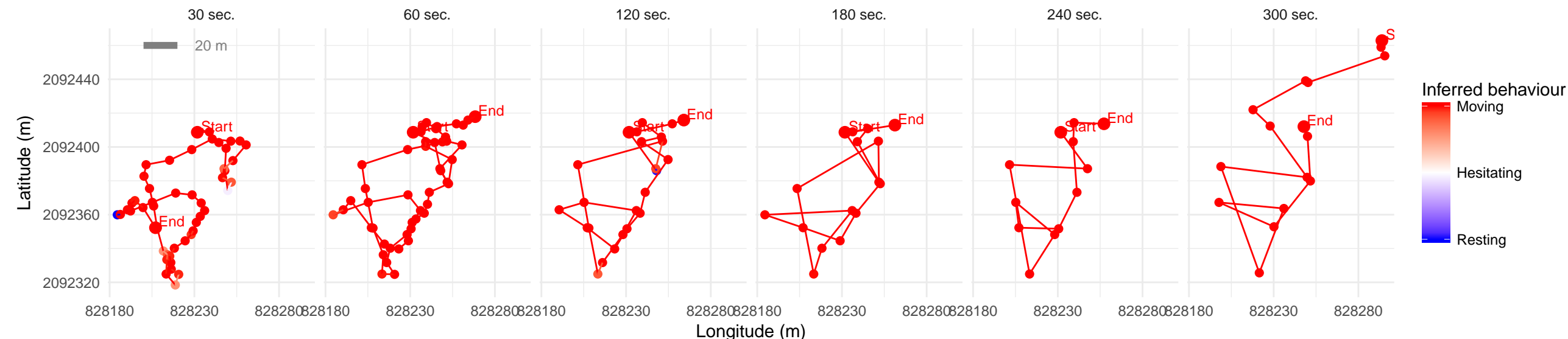
Stationary 2

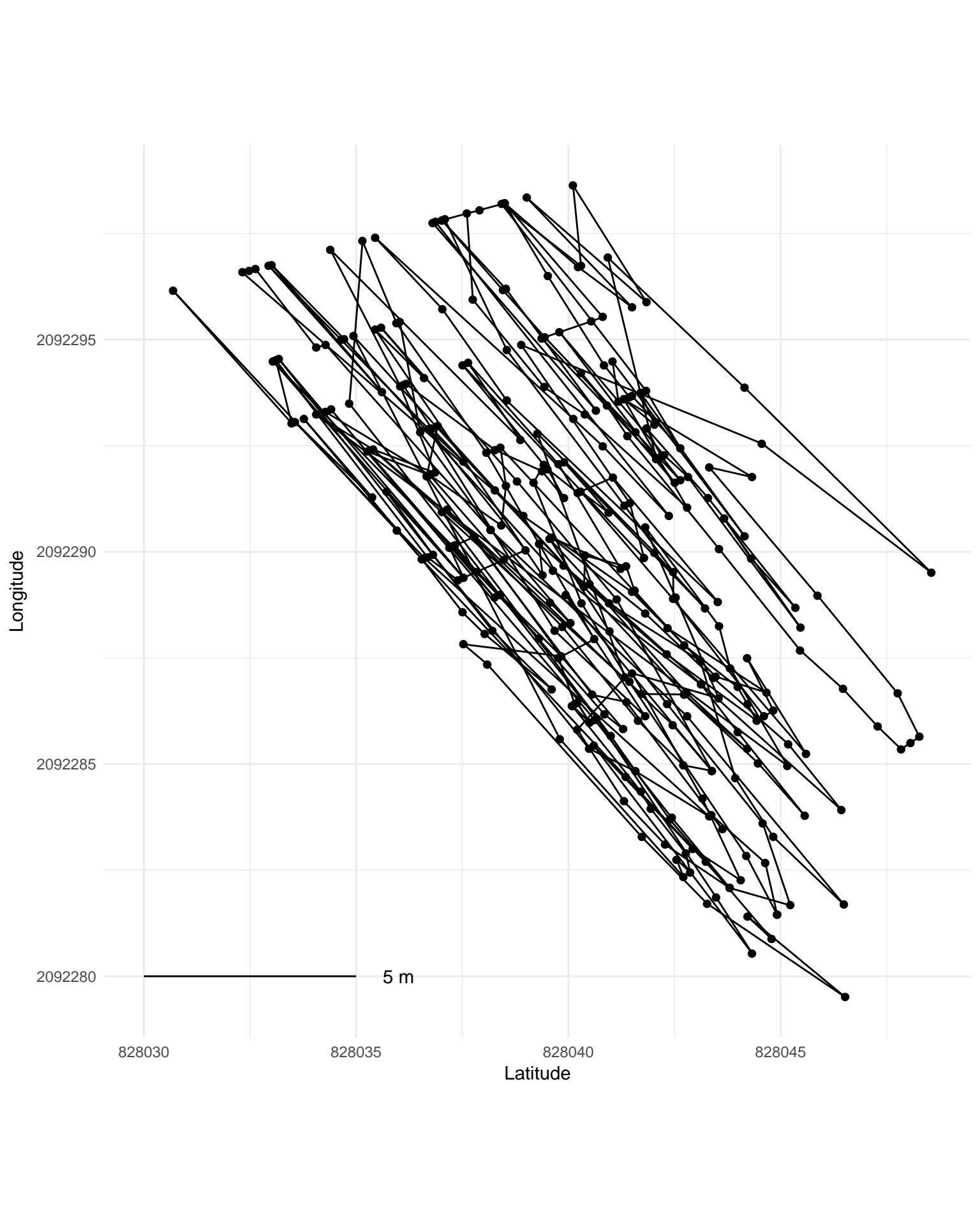


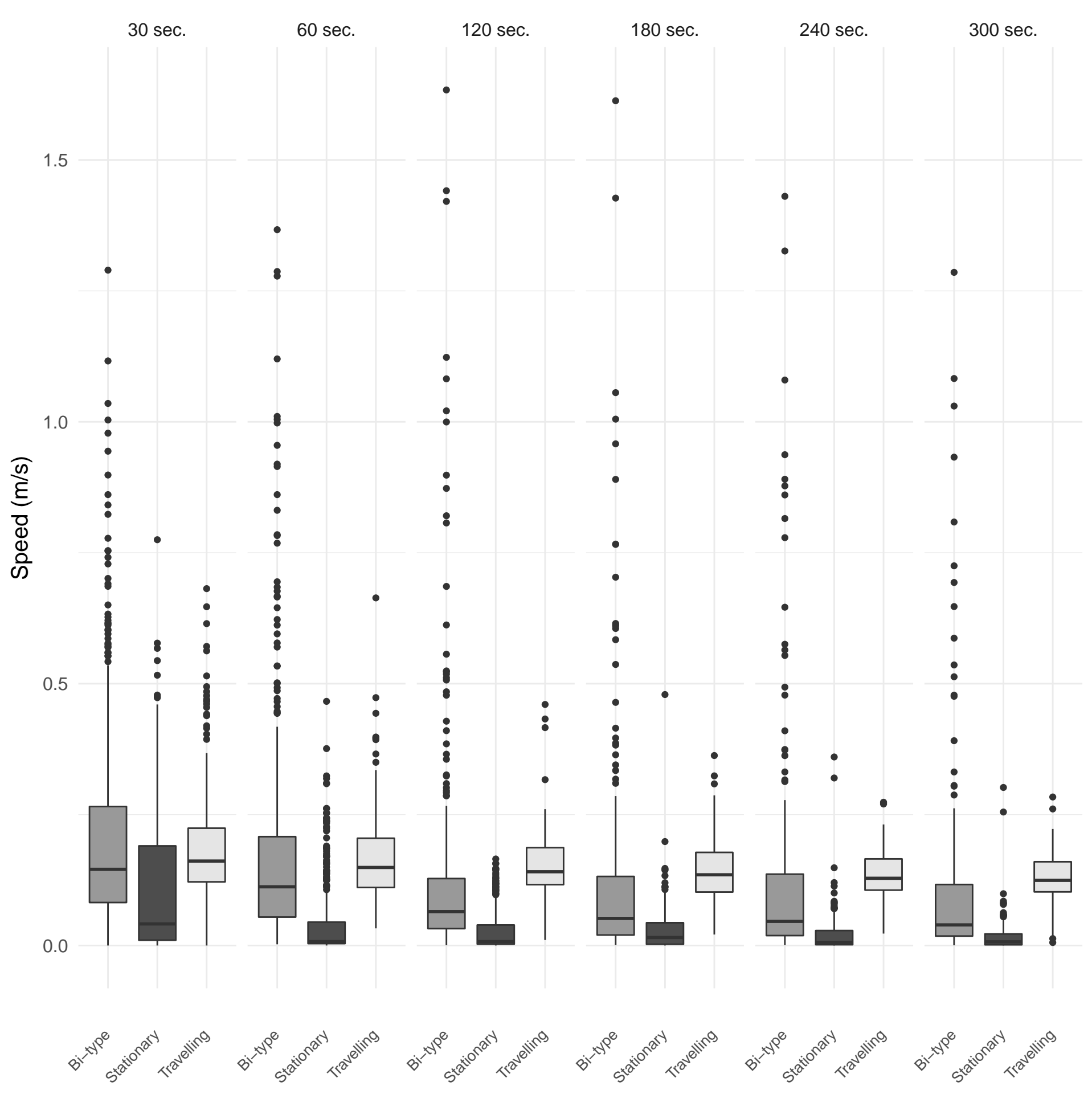
Travelling 1

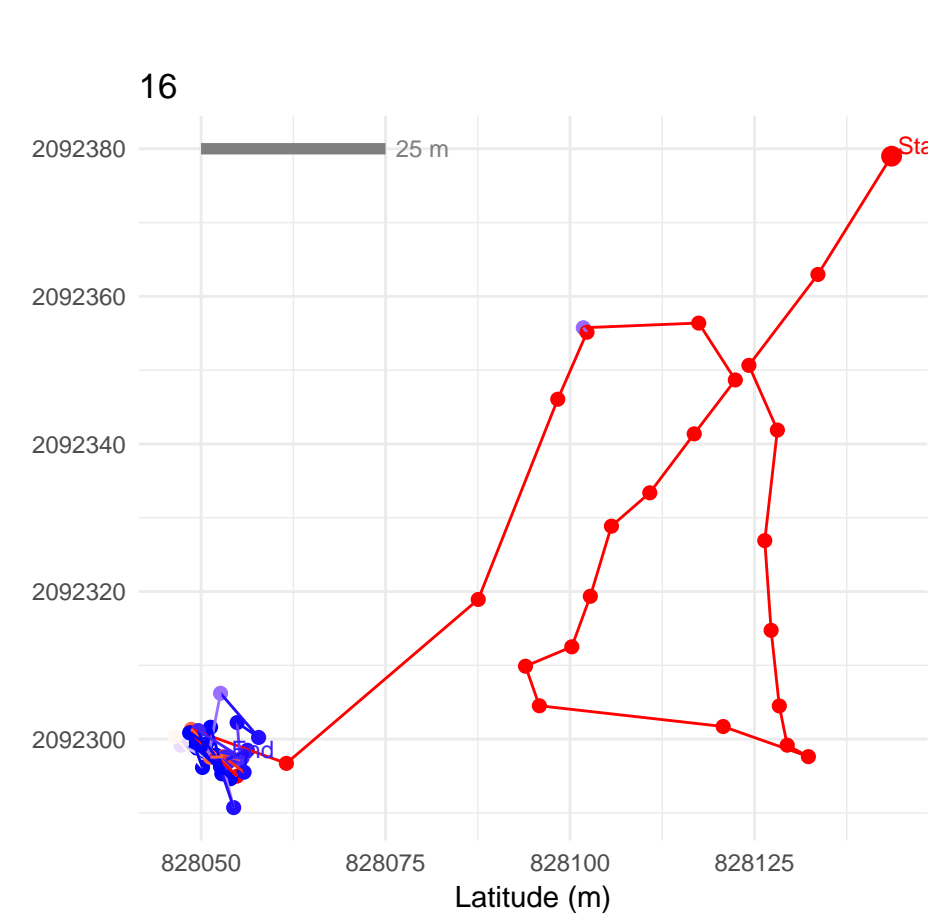
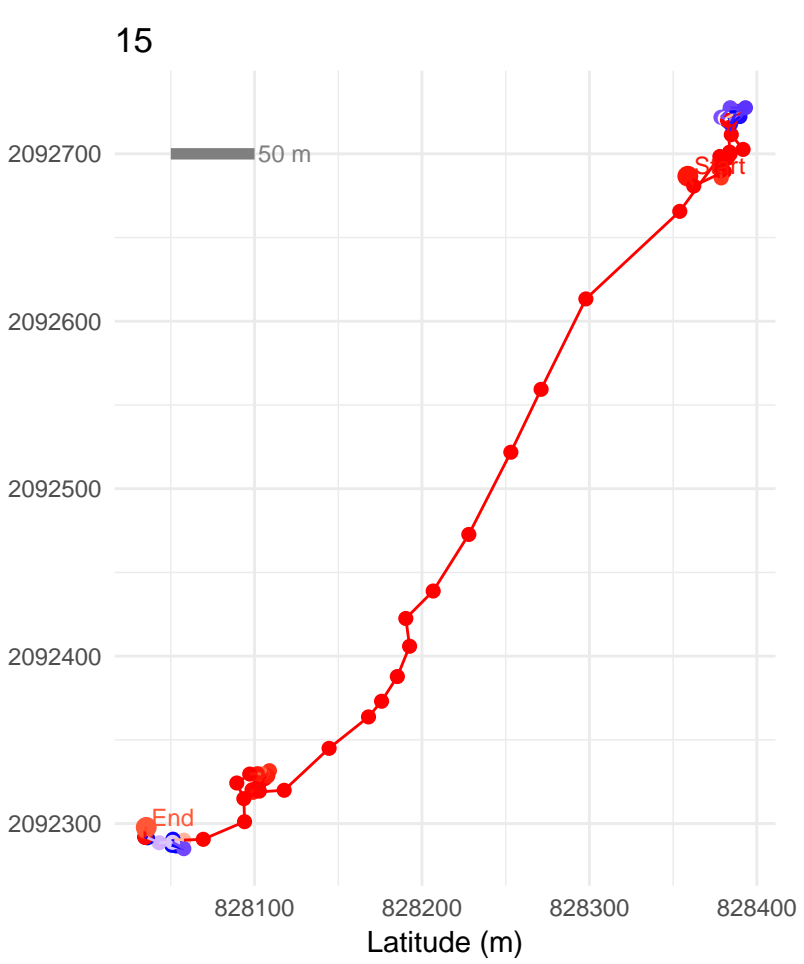
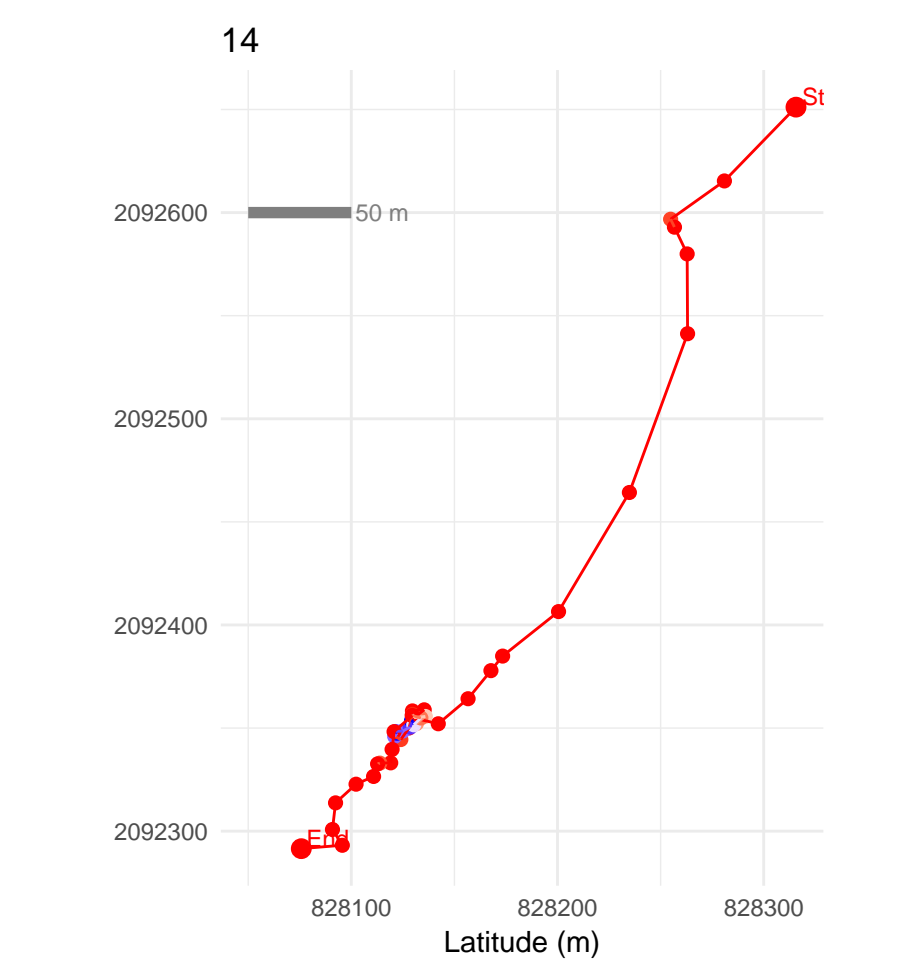
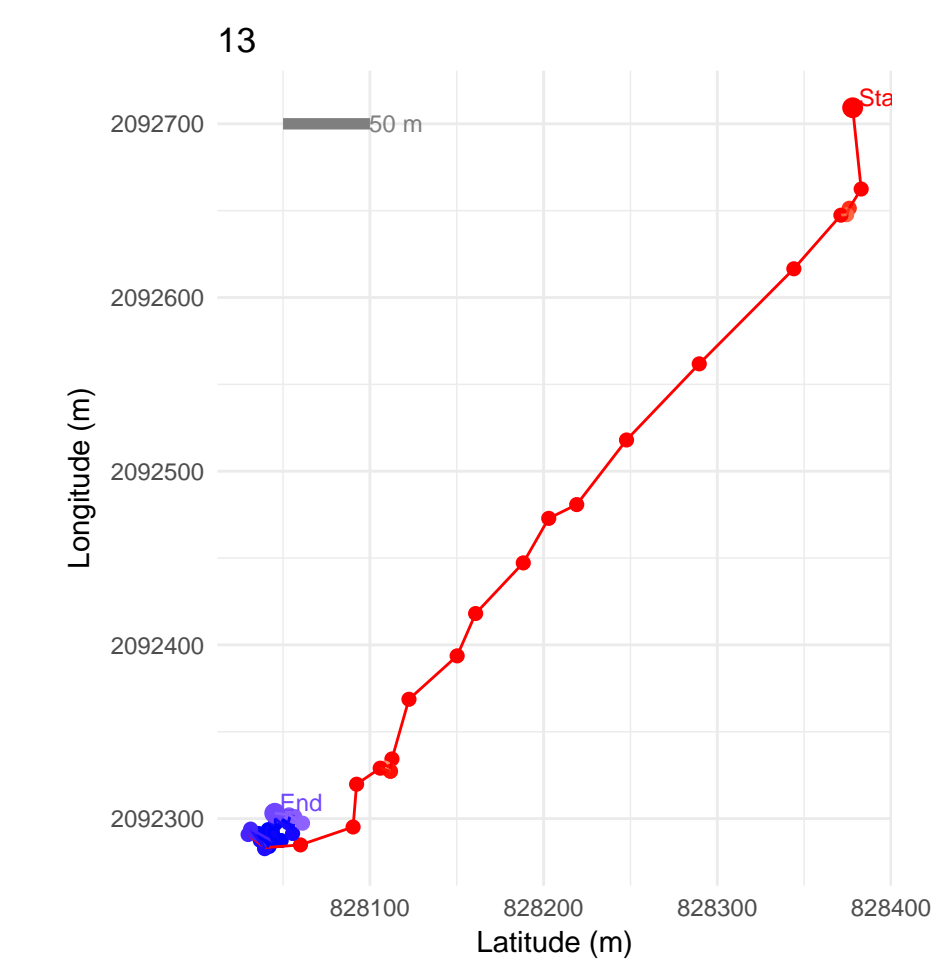
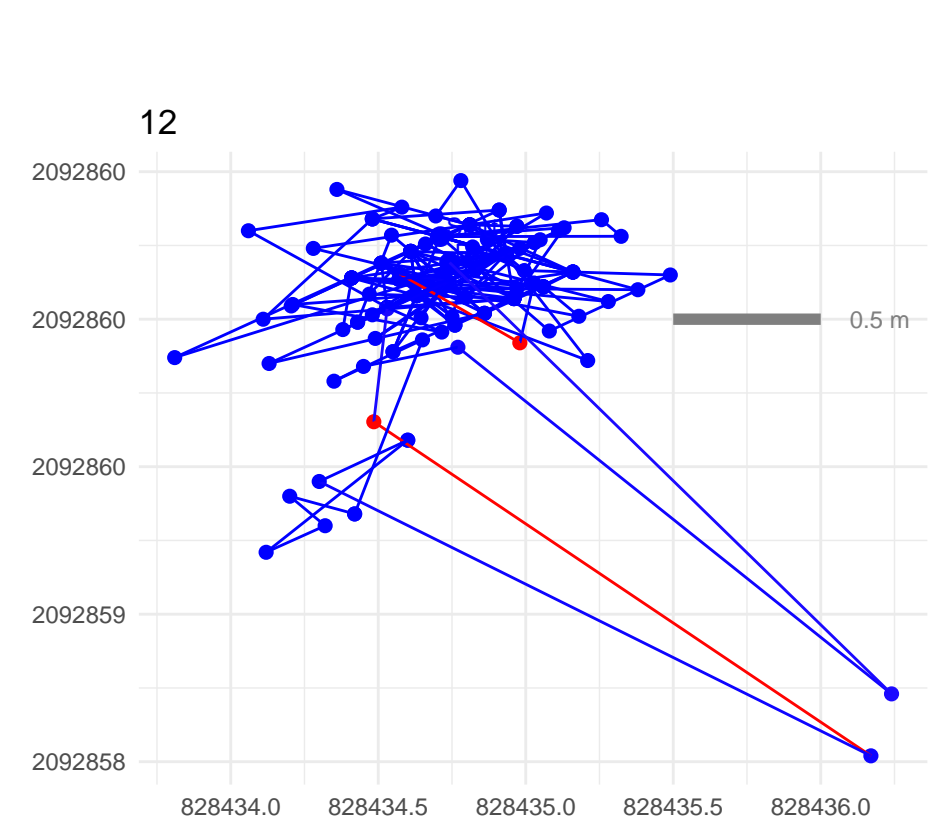
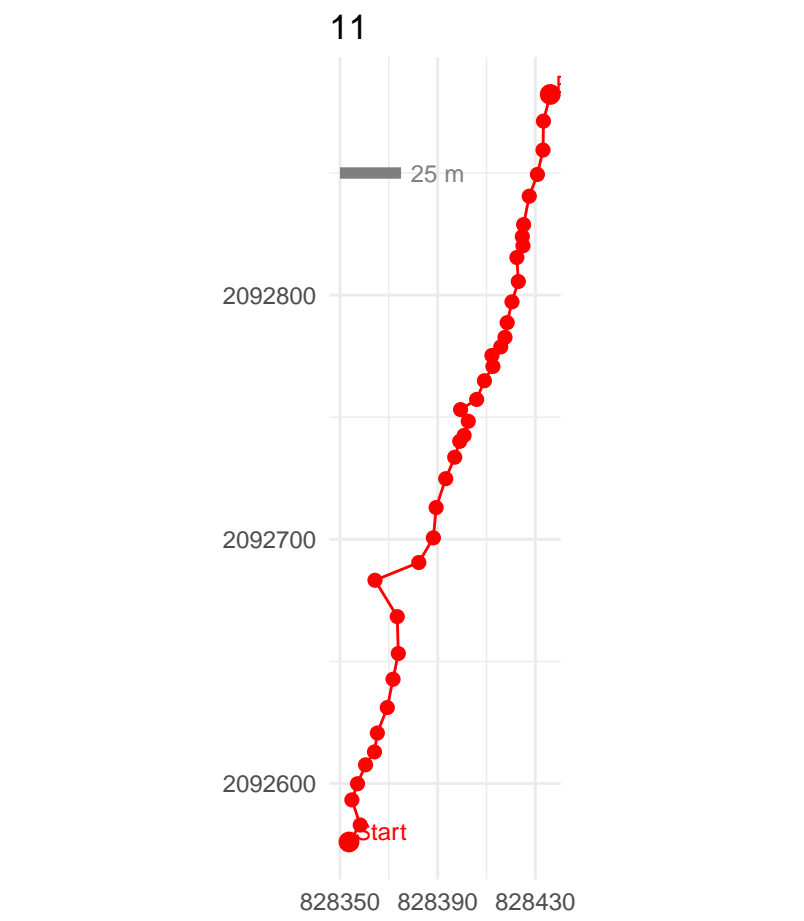
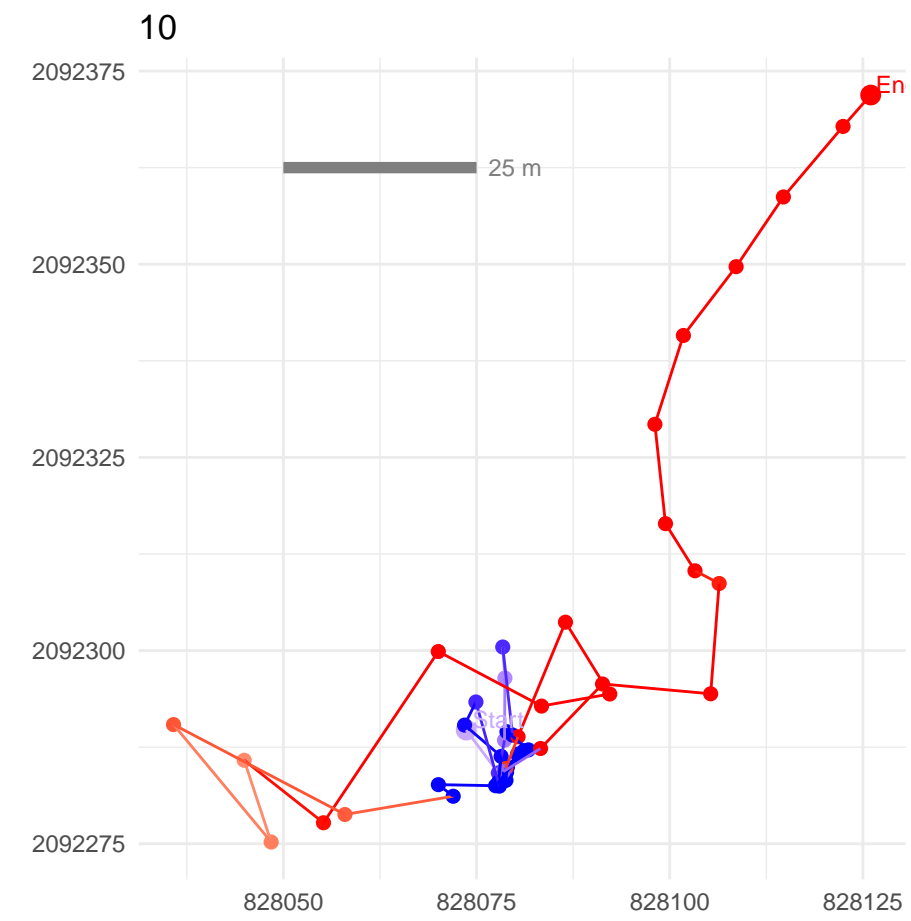
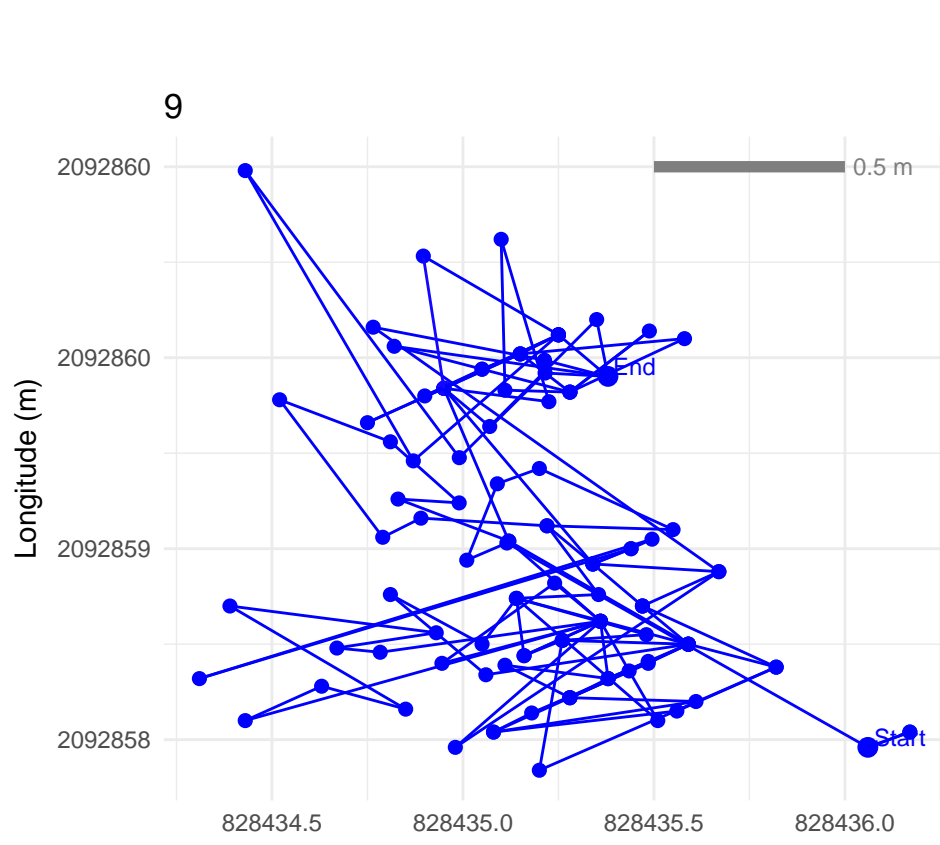
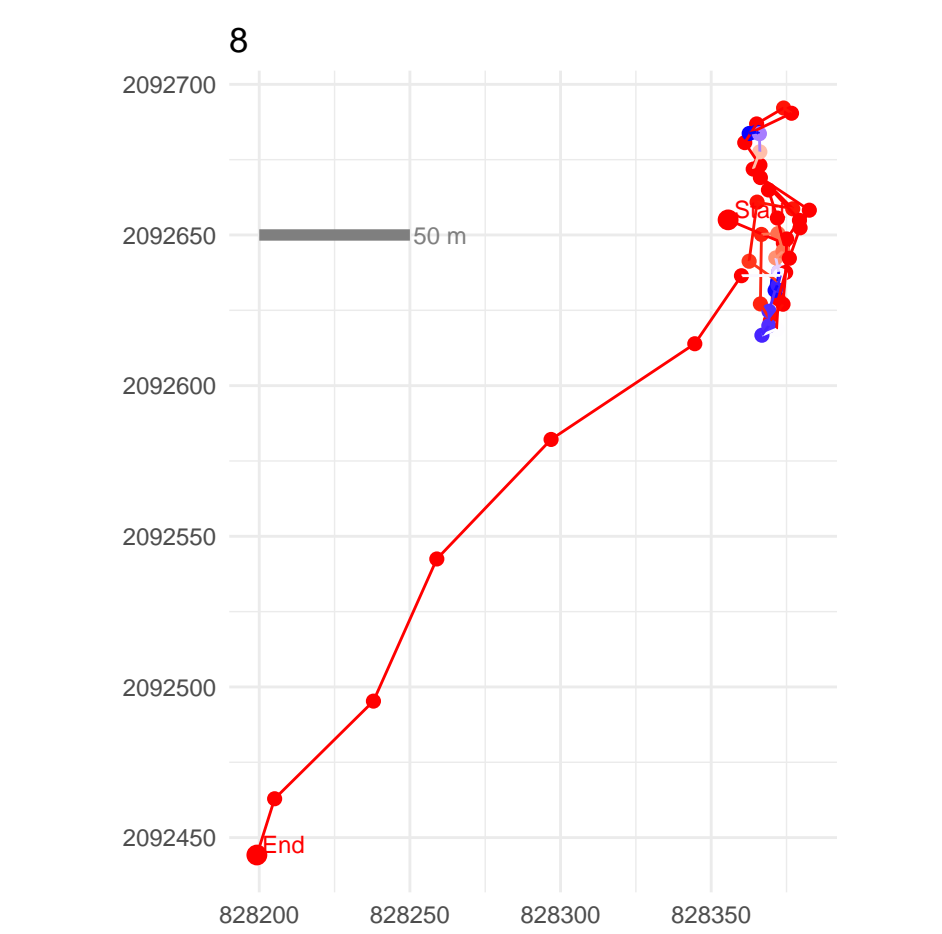
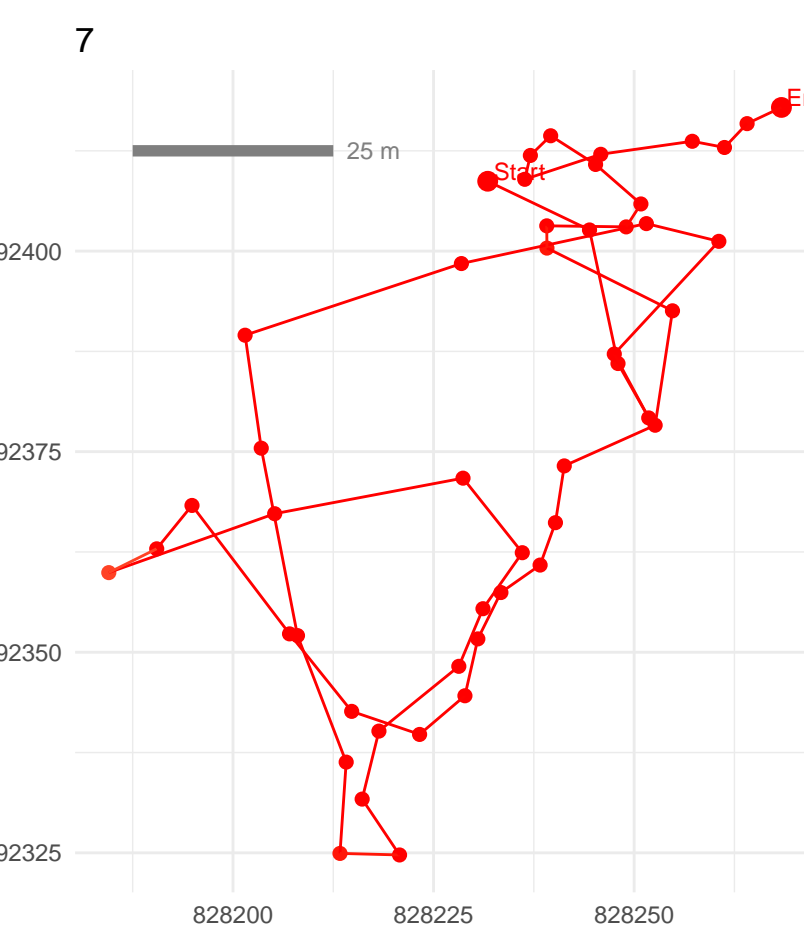
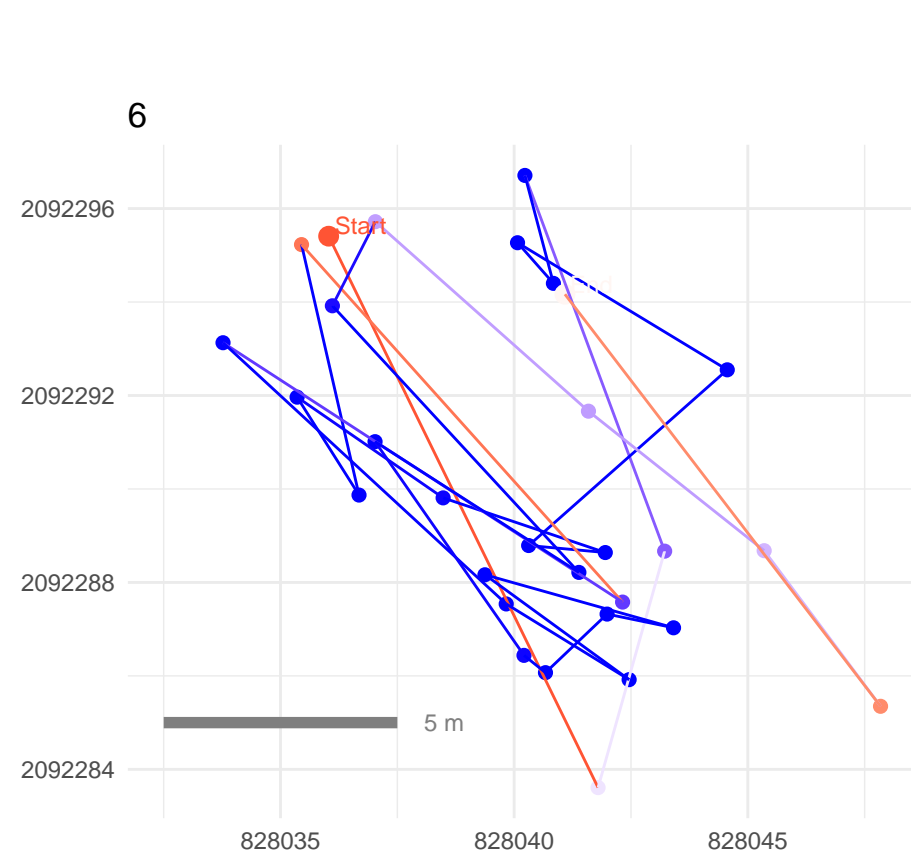
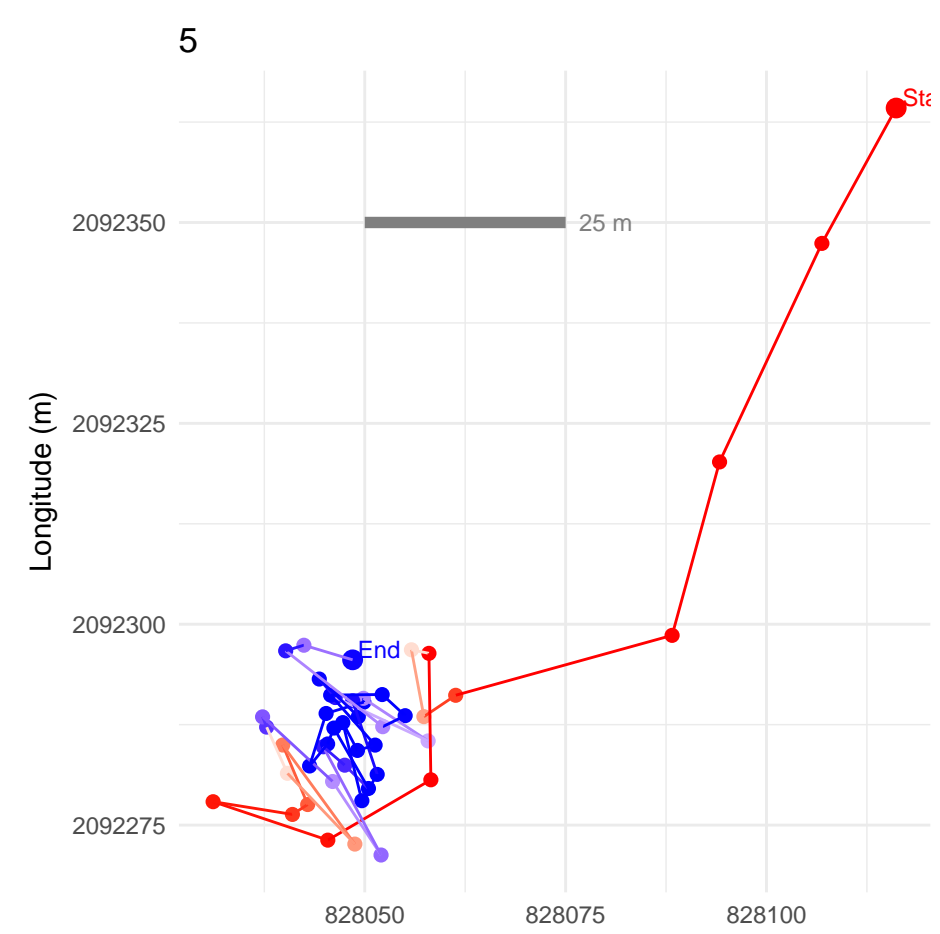
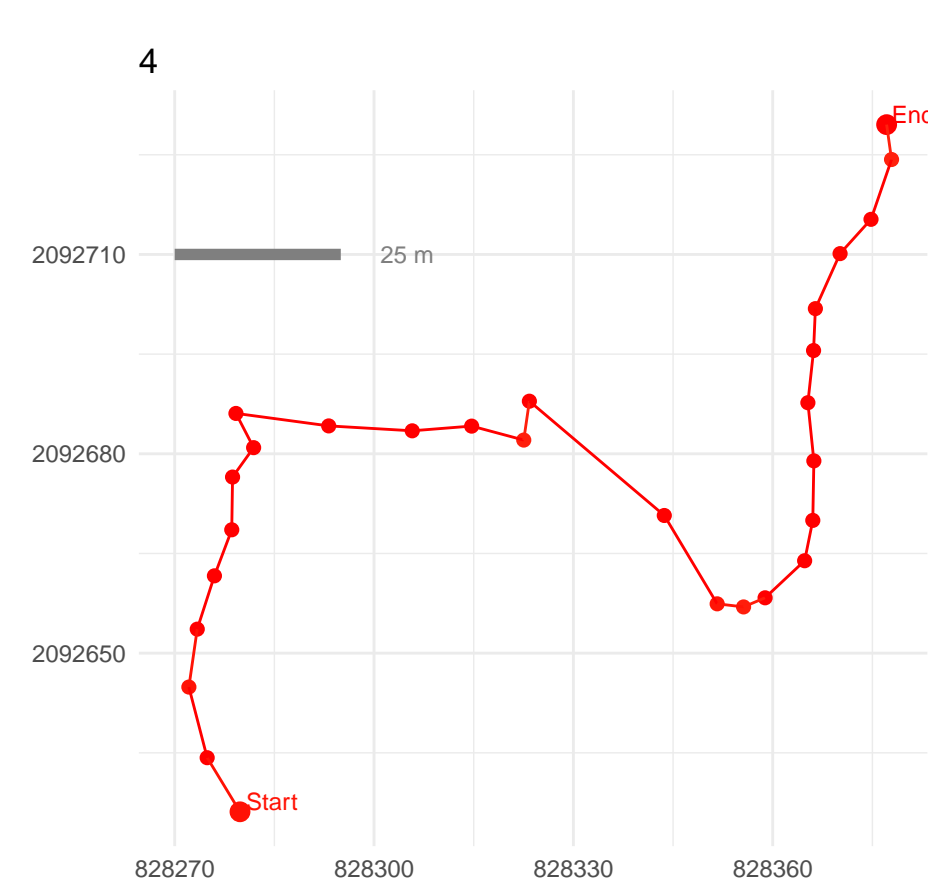
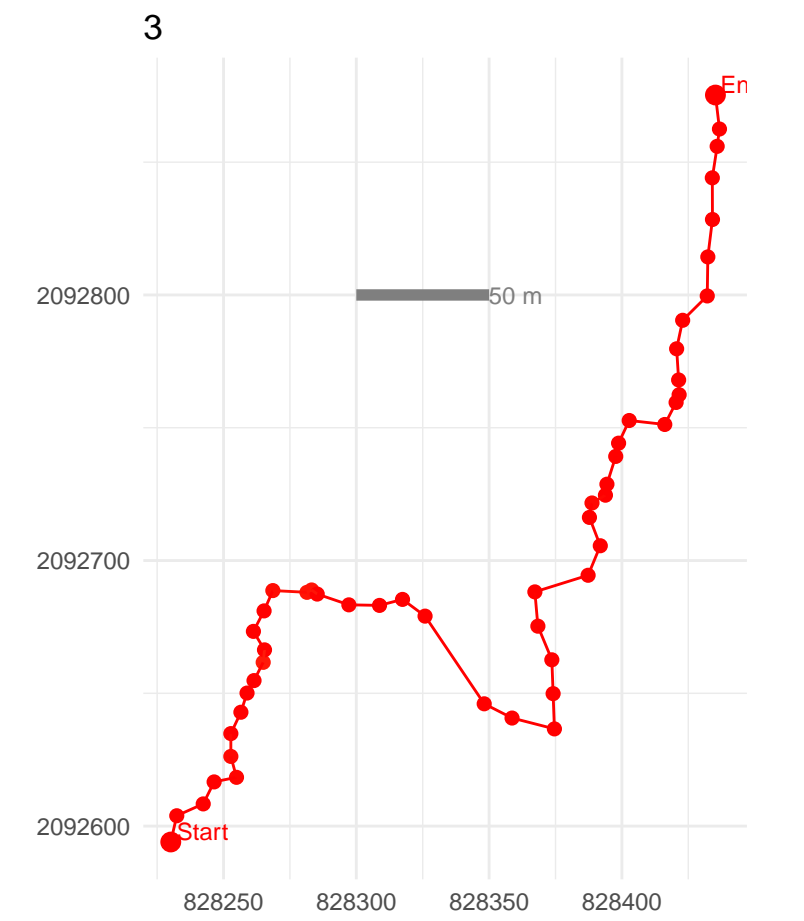
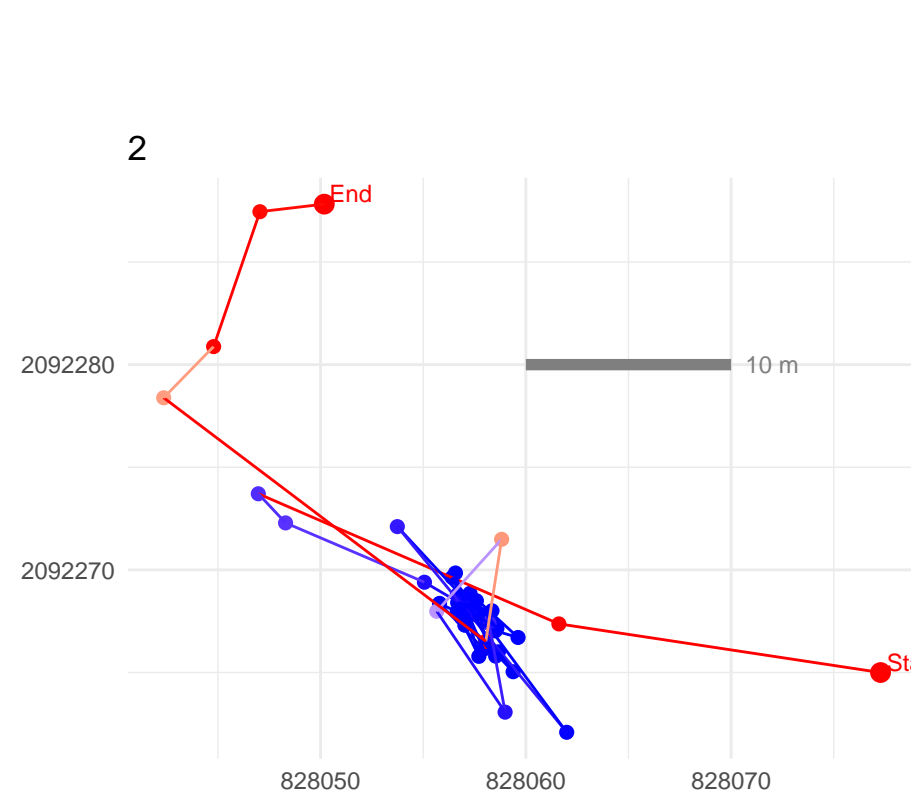
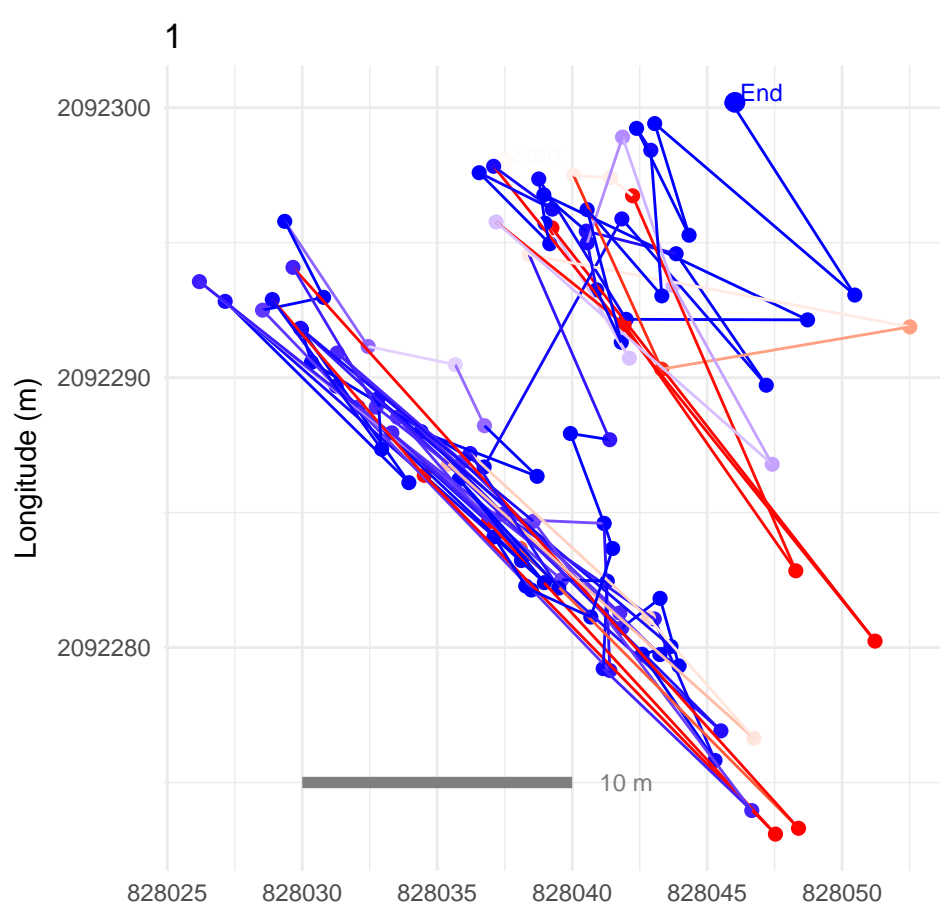


Travelling 2

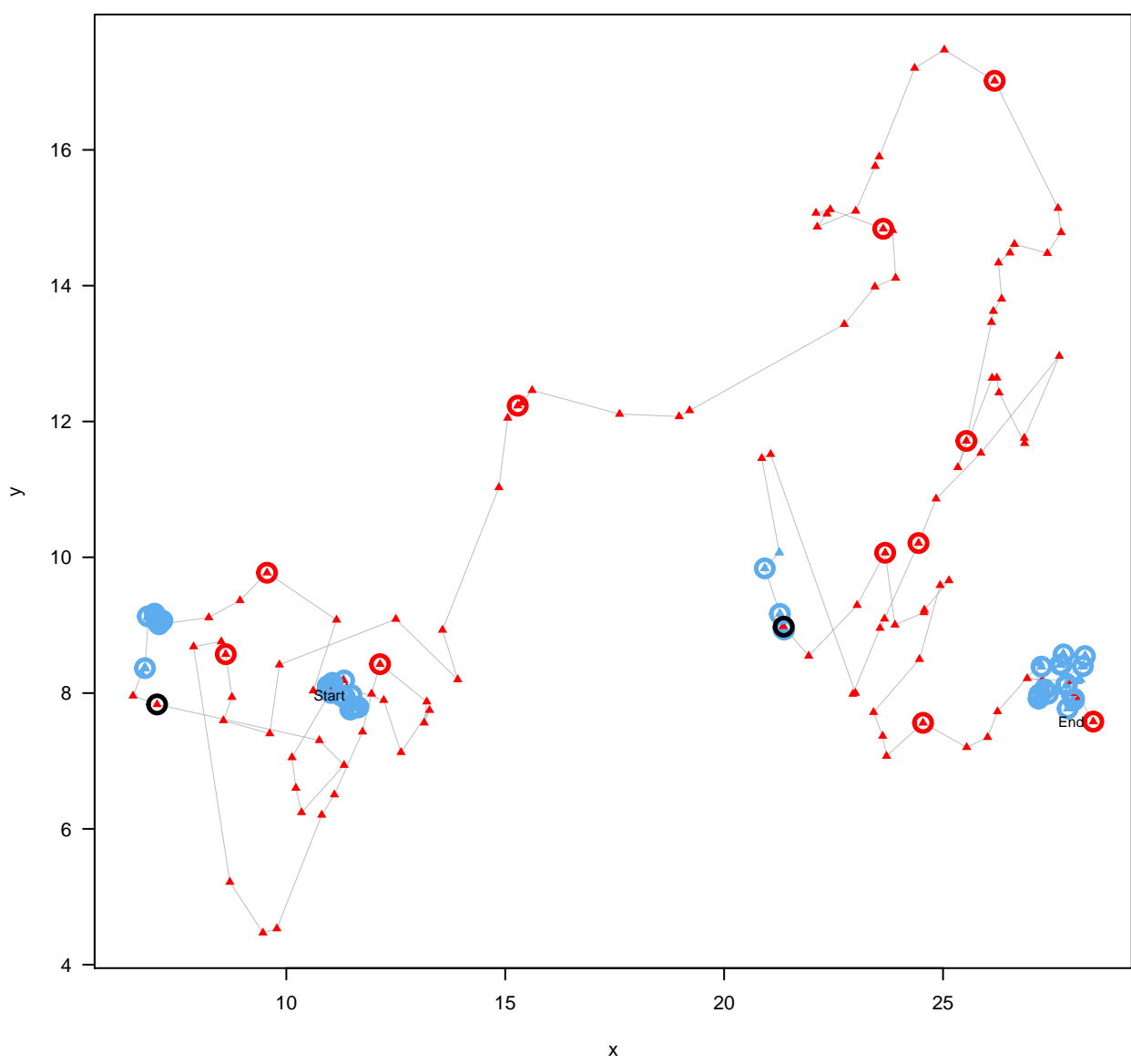




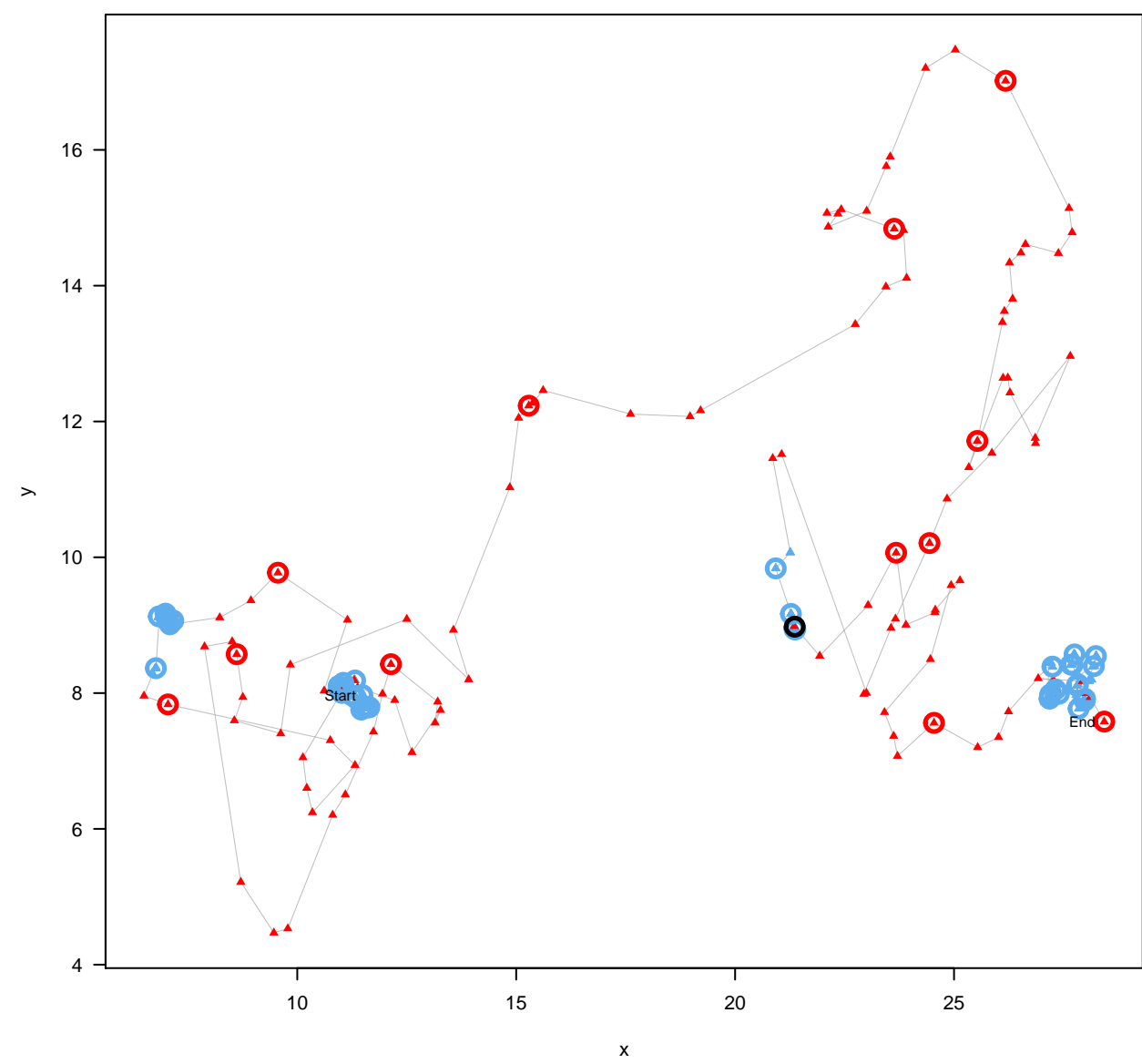




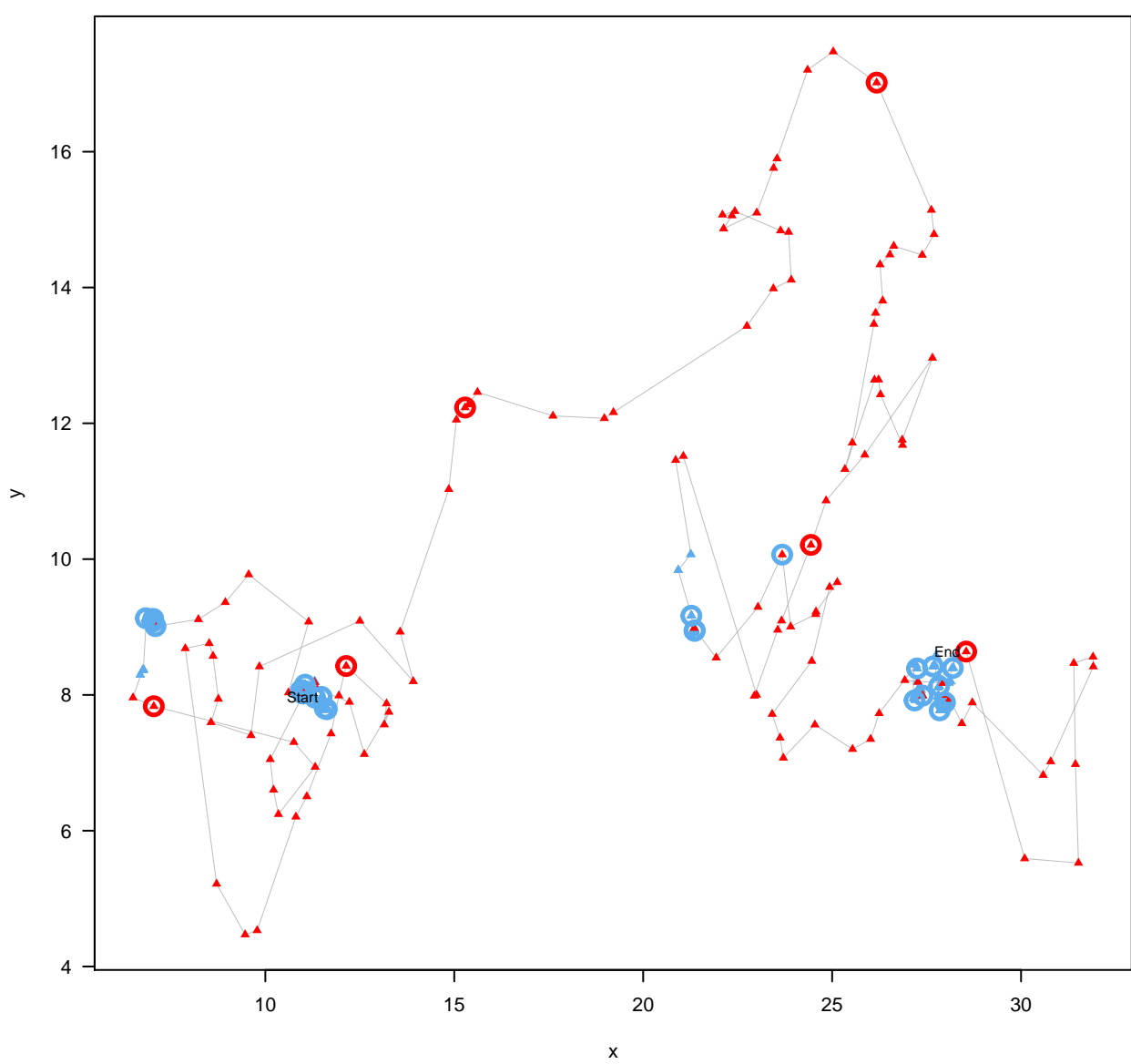
time-step=30s, without U1 and U2



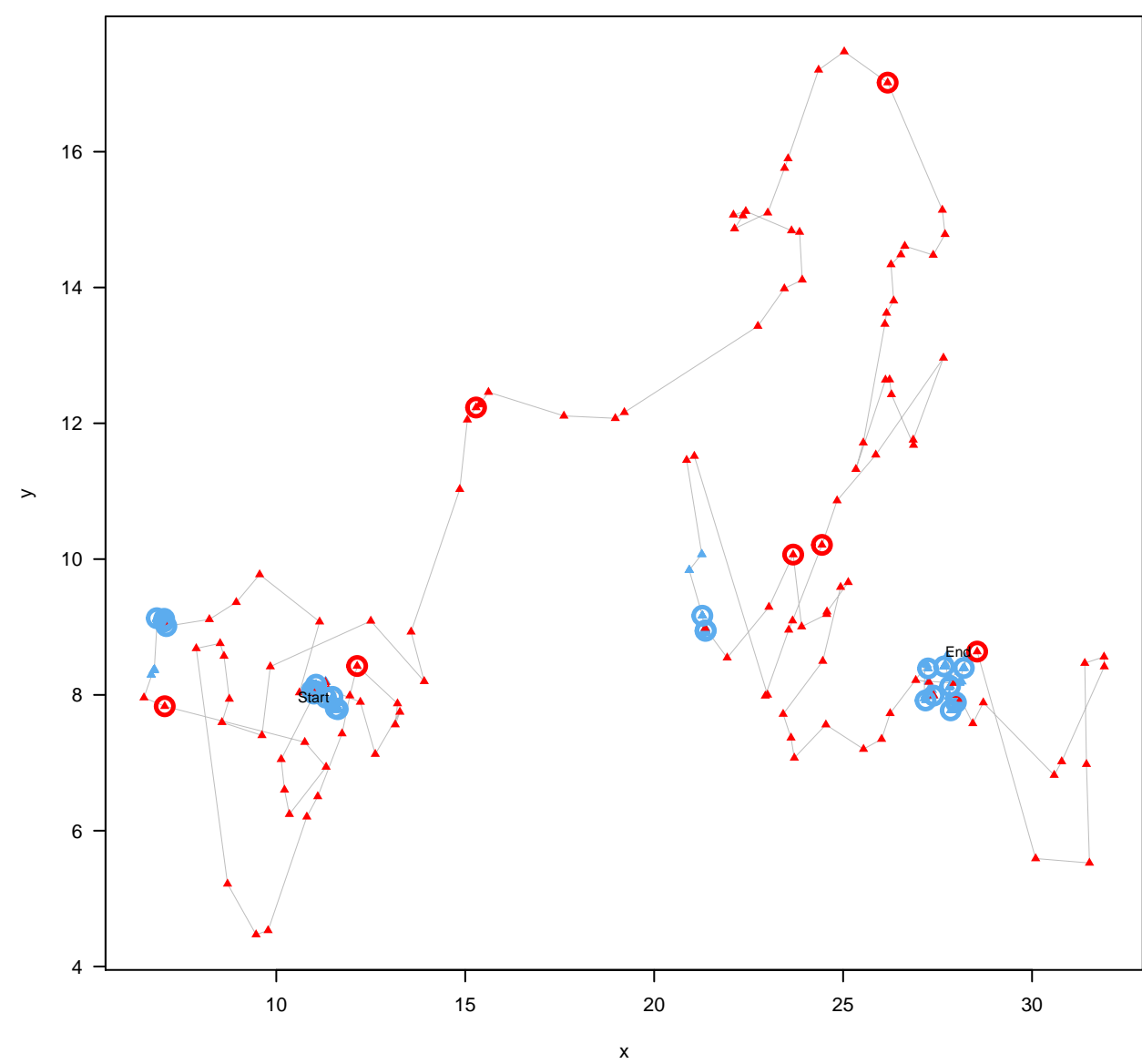
time-step=30s, with U1 and U2



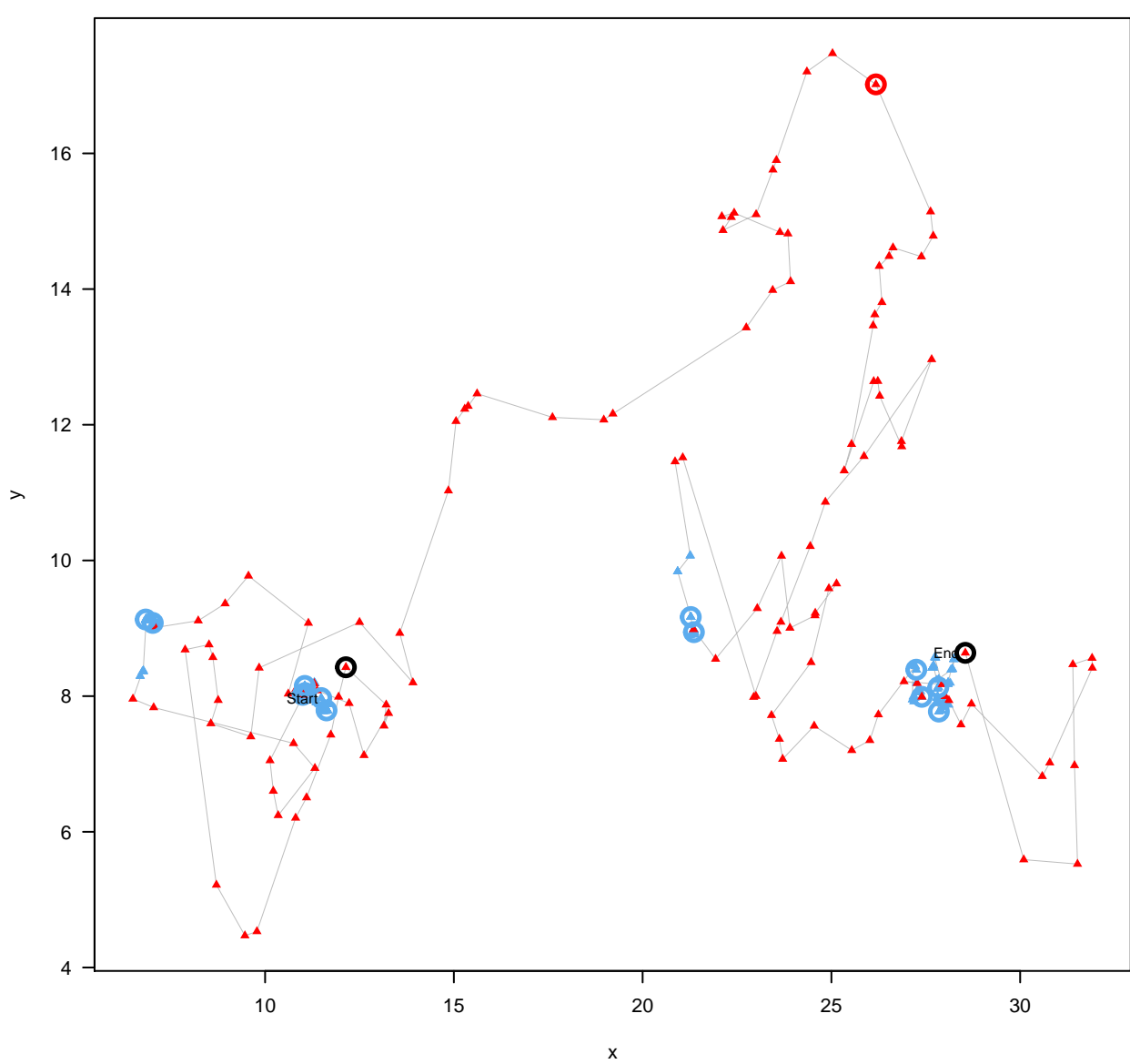
time-step=60s, without U1 and U2



time-step=60s, with U1 and U2



time-step=120s, without U1 and U2



time-step=120s, with U1 and U2

