# Supplementary Material for: Gaussian process enhanced semi-automatic approximate Bayesian computation: parameter inference in a stochastic differential equation system for chemotaxis

Agnieszka Borowska, Diana Giurghita and Dirk Husmeier
School of Mathematics and Statistics, University of Glasgow, U.K.

## A Simulated data examples

Each of Figures A.1–A.4 illustrates the dynamics of a given output – cell contour, GI, S, LA and LI, respectively – for 12 different parameter values. These parameters were generated from the Latin hypercube and their exact values are reported in Table A.1. For the 3 local signals we plot time series of means (where the mean at a given time point is taken over space).

| Index | $f_a$ | $r_c$ | $k_b$ | $d_b$ | $D_b$ | $k_M$ | $s_a$ | $b_a$ | $D_a$ | $d_a$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.0025 | 0.0552 | 0.0042 | 0.0076 | 0.0478 | 0.2237 | 1.19E-04 | 0.1417 | 0.0409 | 0.0399 |
| 2 | 0.0027 | 0.1193 | 0.0029 | 0.0208 | 0.0282 | 0.2162 | 8.70E-05 | 0.1826 | 0.0199 | 0.0146 |
| 3 | 0.0010 | 0.0738 | 0.0048 | 0.0071 | 0.0604 | 0.3006 | 1.31E-04 | 0.1161 | 0.0431 | 0.0171 |
| 4 | 0.0027 | 0.0971 | 0.0037 | 0.0079 | 0.0545 | 0.1390 | 1.07E-04 | 0.1249 | 0.0174 | 0.0207 |
| 5 | 0.0020 | 0.0561 | 0.0029 | 0.0173 | 0.0750 | 0.1442 | 9.38E-05 | 0.0596 | 0.0183 | 0.0118 |
| 6 | 0.0009 | 0.1035 | 0.0054 | 0.0083 | 0.0755 | 0.1033 | 1.15E-04 | 0.0886 | 0.0386 | 0.0302 |
| 7 | 0.0014 | 0.0696 | 0.0050 | 0.0228 | 0.0661 | 0.3094 | 1.27E-04 | 0.1508 | 0.0293 | 0.0215 |
| 8 | 0.0018 | 0.1092 | 0.0038 | 0.0223 | 0.0239 | 0.2357 | 1.38E-04 | 0.1097 | 0.0391 | 0.0395 |
| 9 | 0.0029 | 0.1109 | 0.0042 | 0.0213 | 0.0261 | 0.1916 | 3.51E-05 | 0.1888 | 0.0446 | 0.0212 |
| 10 | 0.0030 | 0.1140 | 0.0039 | 0.0095 | 0.0416 | 0.2319 | 1.25E-04 | 0.0824 | 0.0362 | 0.0277 |
| 11 | 0.0011 | 0.0894 | 0.0021 | 0.0201 | 0.0578 | 0.2210 | 9.59E-05 | 0.1958 | 0.0441 | 0.0133 |
| 12 | 0.0030 | 0.0442 | 0.0026 | 0.0171 | 0.0682 | 0.2365 | 1.39E-04 | 0.1375 | 0.0251 | 0.0201 |

**Table A.1:** Values of $\boldsymbol{\theta}$ generated from the Latin hypercube used to obtain simulations plotted in Figure 11 and Figures A.1–A.4.

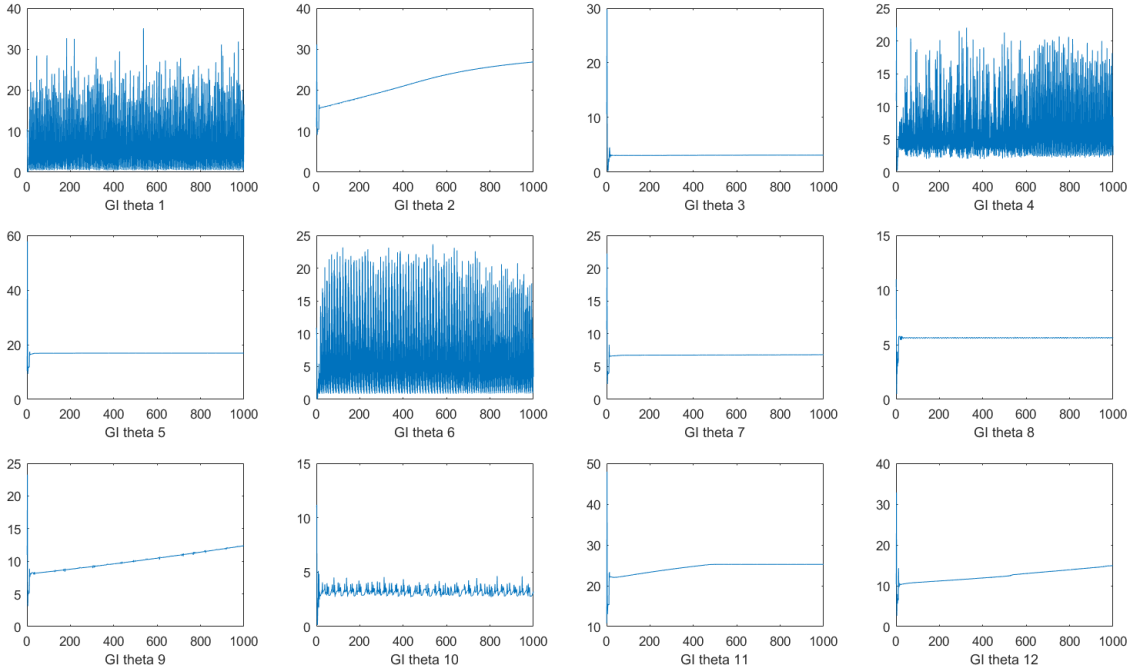



**Figure A.1:** Examples of time series of GI for different parameter vectors.

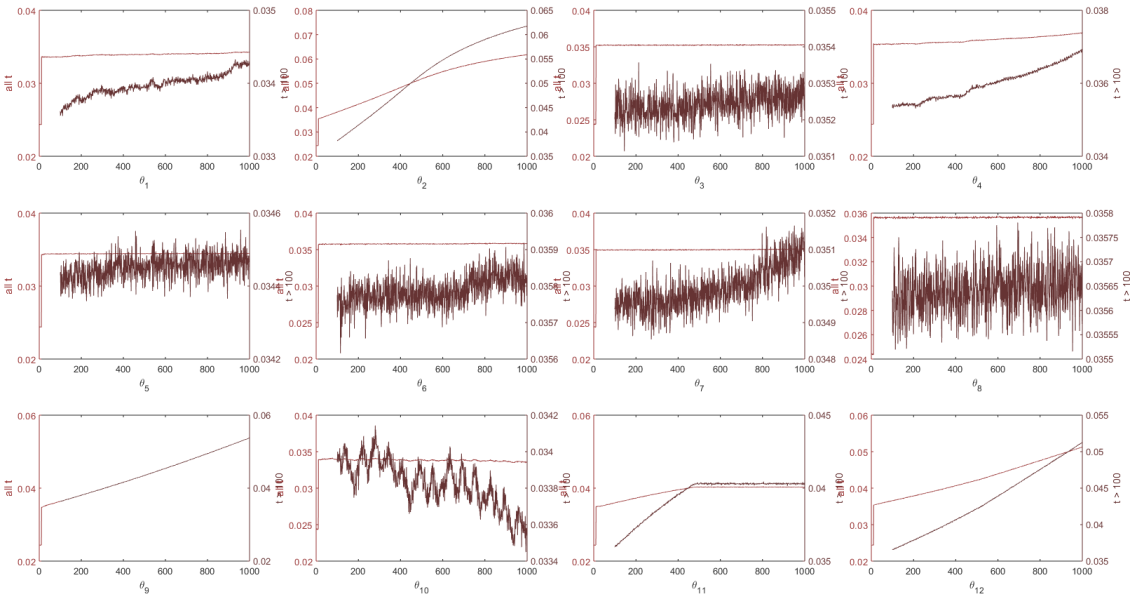



**Figure A.2:** Examples of time series of means over space of S for different parameter vectors for $t = 1, 2, \ldots, 1000$ (left axis) and for $t \geq 100$ (right axis).

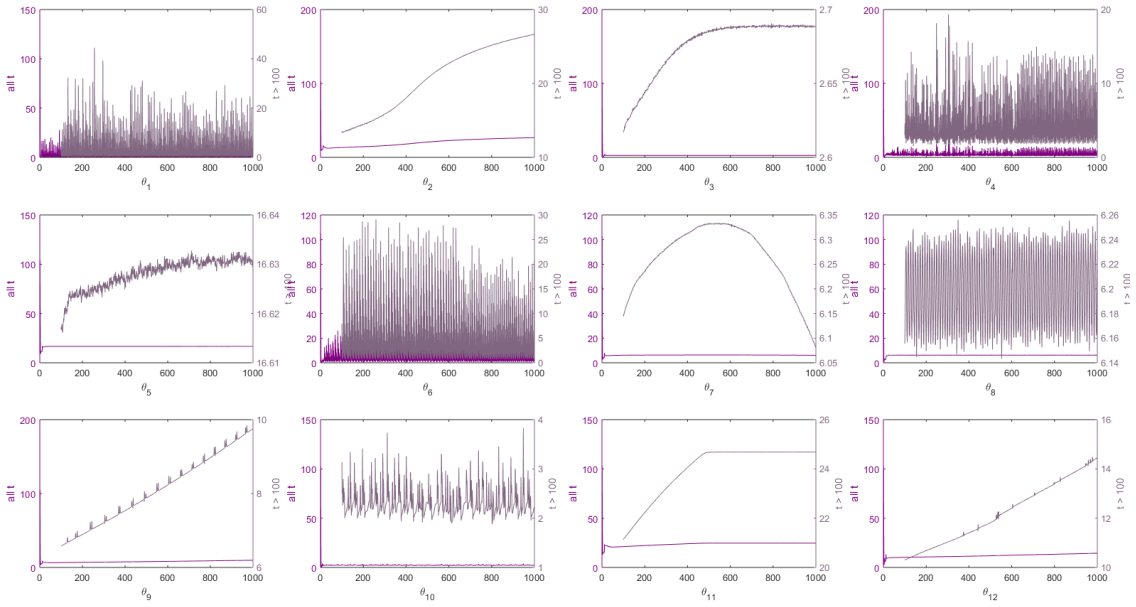



**Figure A.3:** Examples of time series of means over space of LA for different parameter vectors for $t = 1, 2, \ldots, 1000$ (left axis) and for $t \geq 100$ (right axis).

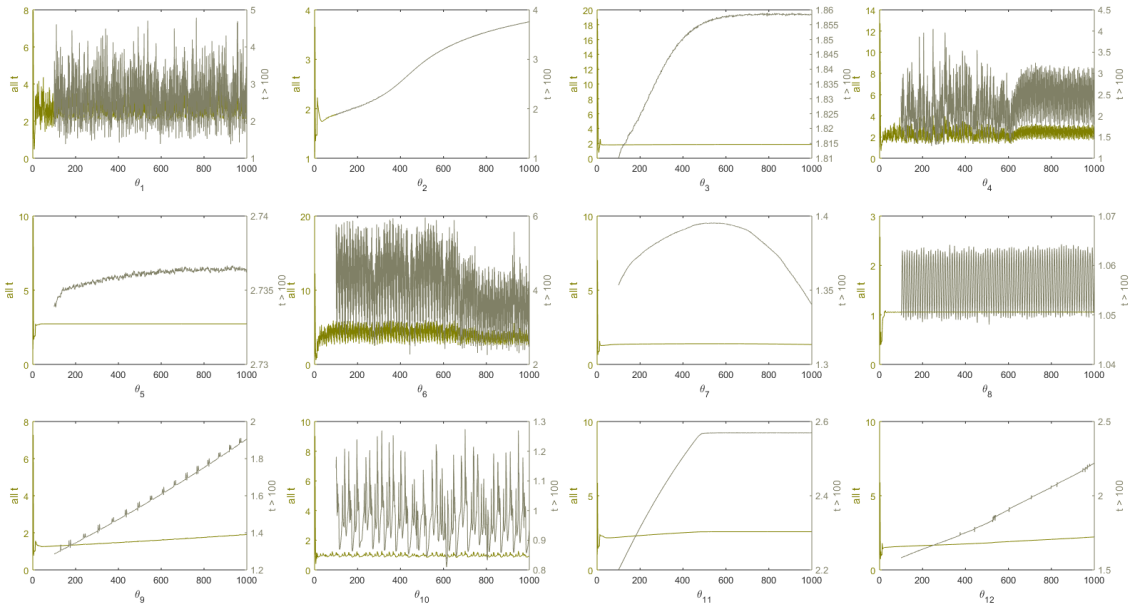



**Figure A.4:** Examples of time series of means over space of LI for different parameter vectors for $t = 1, 2, \ldots, 1000$ (left axis) and for $t \geq 100$ (right axis).

# B Generalised Procrustes analysis

Figure B.1 illustrates the problem of applying the generalised Procrustes analysis (GPA) applied to cell shapes. GPA results in a collection of scaled, rotated and translated shapes. GPA aligns shapes based on matching anatomical or mathematical landmarks (Dryden and Mardia, 2016). However, as noted by Tweedy et al. (2013), cell contours are highly variable and due to pseudopod formation and removal, there are no recurring features on cell membranes which could serve as landmarks for cell shape comparison. In our application we tried using finite element nodes as approximate landmarks, yet this approximation turned out to be too crude: the resulting cell shapes are not properly aligned. For instance, notice the presence of elongated cells which instead of overlapping are perpendicular to each other.

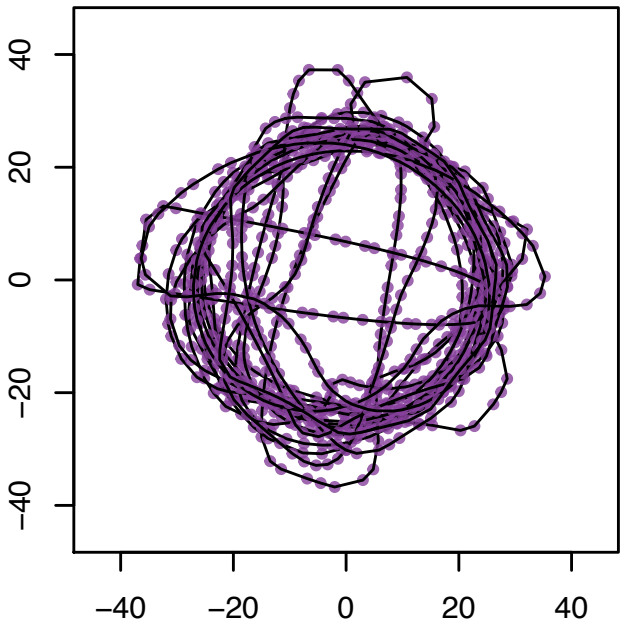



**Figure B.1:** Generalised Procrustes analysis applied to cell shapes.

# C Additional results

## C.1 Cell movements underlying Figure 12 (zero-crossings in PCA space)

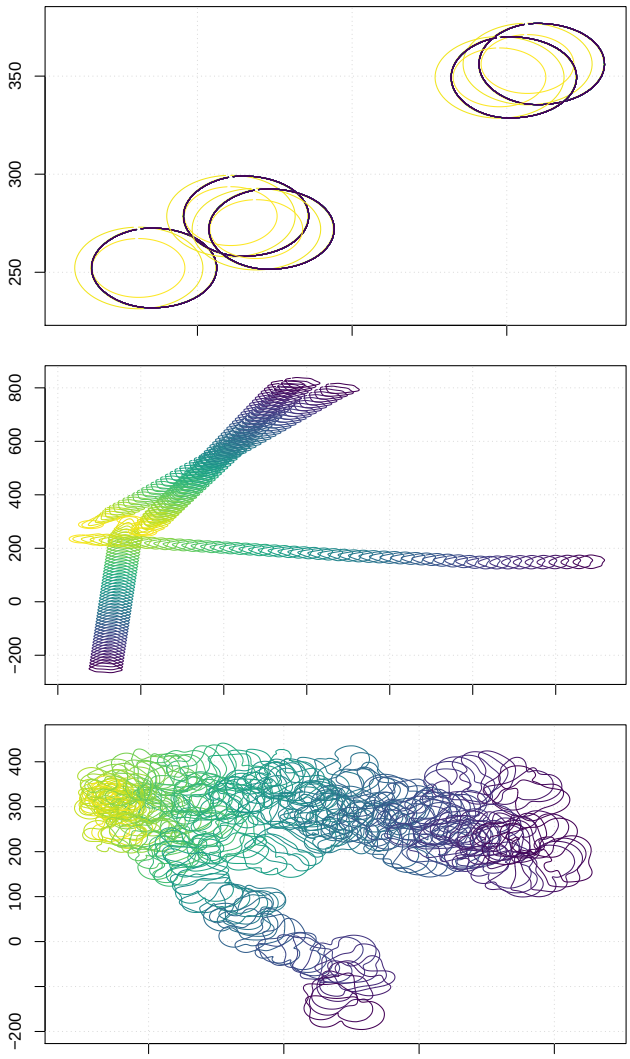
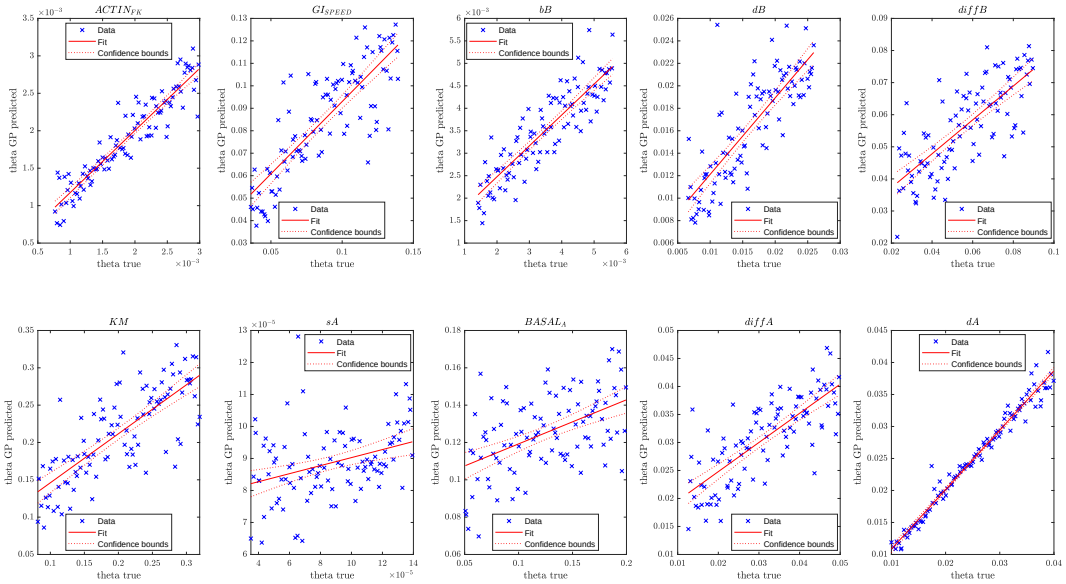
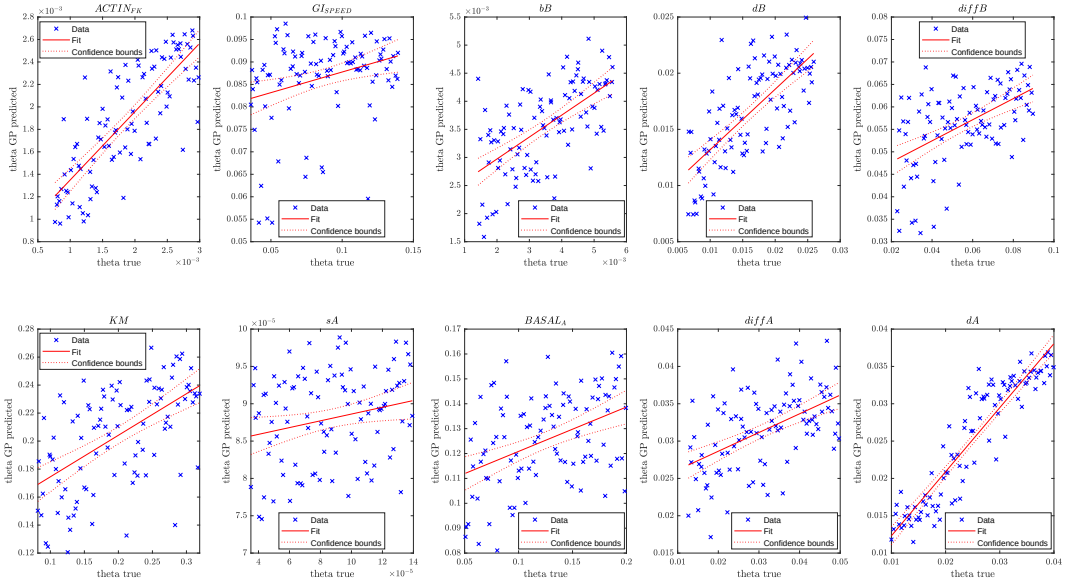



**Figure C.1:** Cell movements underlying Figure 12 (zero-crossings in PCA space), with the top plot corresponding to the left one in Figure 12, the middle one – to the middle one in Figure 12, the bottom one – to the right one in Figure 12. Each plot shows five independent realisation of cell movement generated with the same parameter $\boldsymbol{\theta}$.

## C.2 GP regression out-of-sample performance




**(a)** *Fully Observed Data*




**(b)** *Partially Observed Data.*

**Figure C.2:** Out-of-sample performance of the final kernels: predicted values against true values with the fitted regression line.

### C.3 Real data vs predicted data

***Fully Observed Data*** Explanation for Figures C.3–C.4 (from left to right).
Top row: contour evolution in time for $t = 1, 2, \ldots, 1000$ (the shade indicates the time point, from early – light, to late – dark); contour evolution in time for $t = 1, 101, \ldots, 901$; time series of GI for $t = 1, 2, \ldots, 1000$ (left axis) and for $t \geq 100$ (right axis); time series of area (left axis) and perimeter (right axis); time series of area-perimeter ratios.
Second row: cell contour at $t = 100$; value of S over the cell membrane at $t = 100$; value of S over the cell membrane at $t = 500$; value of S over the cell membrane at $t = 1000$; time series of means over space of S for $t = 1, 2, \ldots, 1000$ (left axis) and for $t \geq 100$ (right axis).
Third row: cell contour at $t = 500$; value of LA over the cell membrane at $t = 100$; value of LA over the cell membrane at $t = 500$; value of LA over the cell membrane at $t = 1000$; time series of means over space of LA for $t = 1, 2, \ldots, 1000$ (left axis) and for $t \geq 100$ (right axis).
Bottom row: cell contour at $t = 1000$; value of LI over the cell membrane at $t = 100$; value of LI over the cell membrane at $t = 500$; value of LI over the cell membrane at $t = 1000$; time series of means over space of LI for $t = 1, 2, \ldots, 1000$ (left axis) and for $t \geq 100$ (right axis).

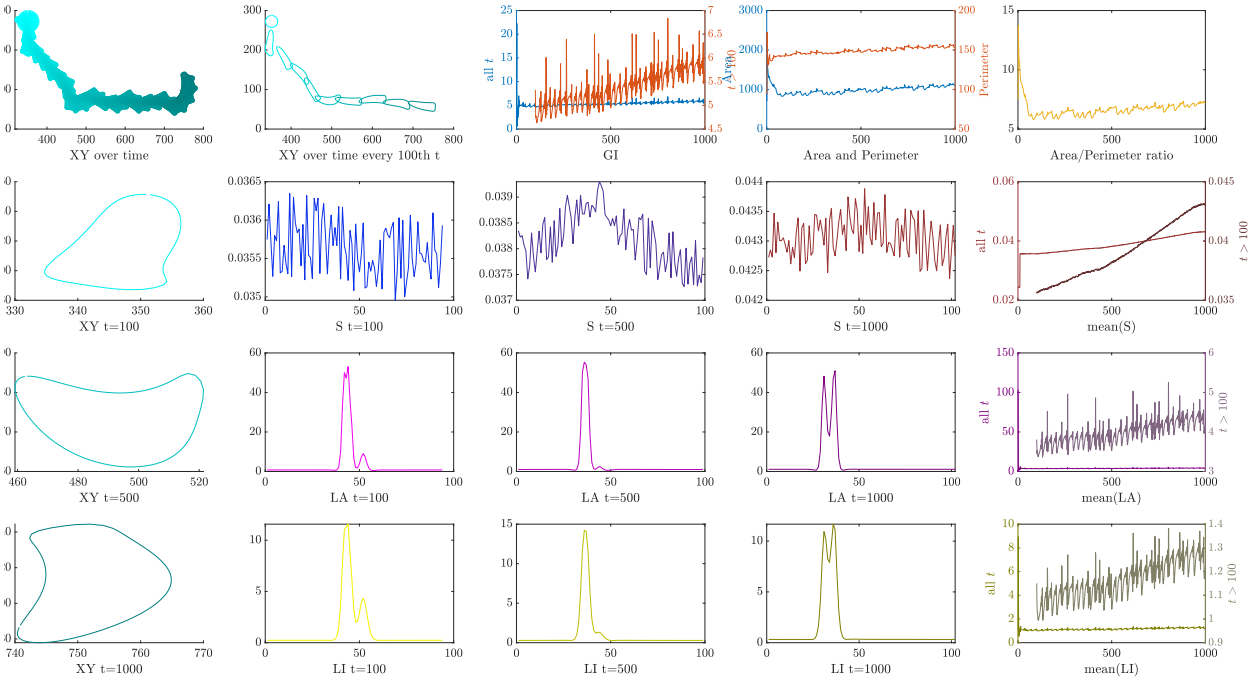
***Partially Observed Data*** Explanation for Figures C.5–C.6 (from left to right).
Top row: contour evolution in time for $t = 1, 2, \ldots, 1000$ (the shade indicates the time point, from early – light, to late – dark); contour evolution in time for $t = 1, 101, \ldots, 901$.
Bottom row: time series of area (left axis) and perimeter (right axis); time series of area-perimeter ratios.

## Bibliography

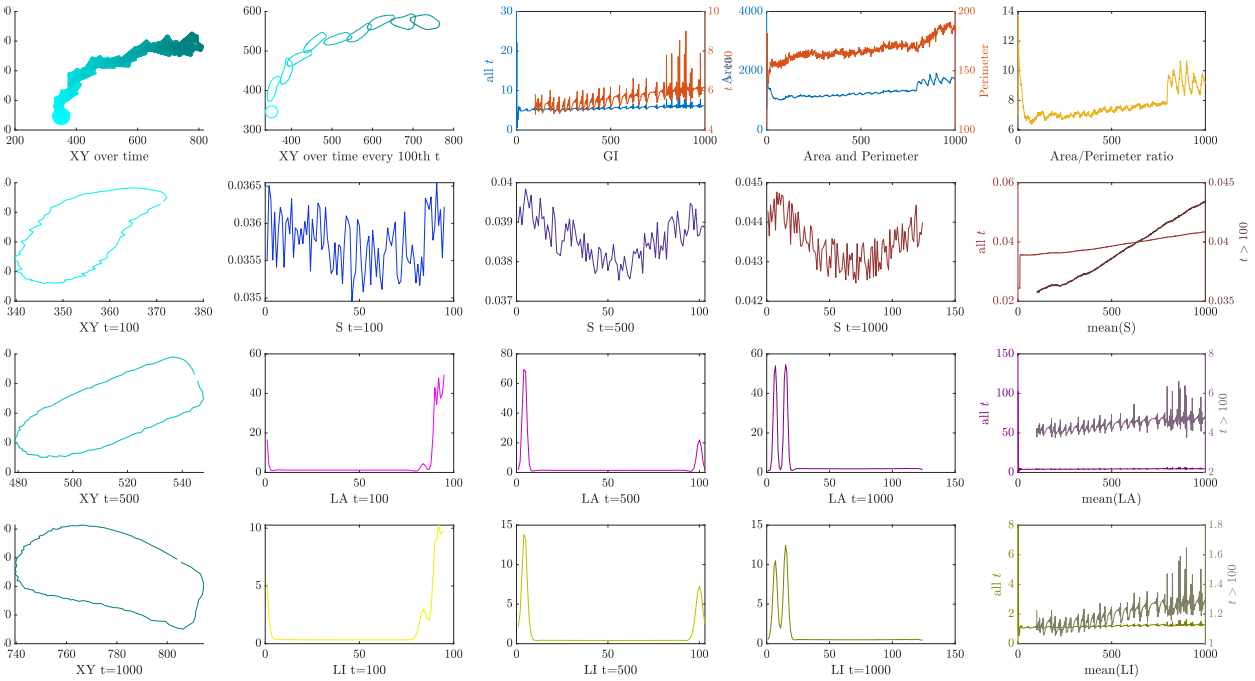
I. L. Dryden, K. V. Mardia, Statistical Shape Analysis, with Applications in R. Second Edition, John Wiley and Sons, Chichester, 2016.

L. Tweedy, B. Meier, J. Stephan, D. Heinrich, R. G. Endres, Distinct cell shapes determine accurate chemotaxis, Scientific Reports 3 (2013) 2606.

## GP predicted



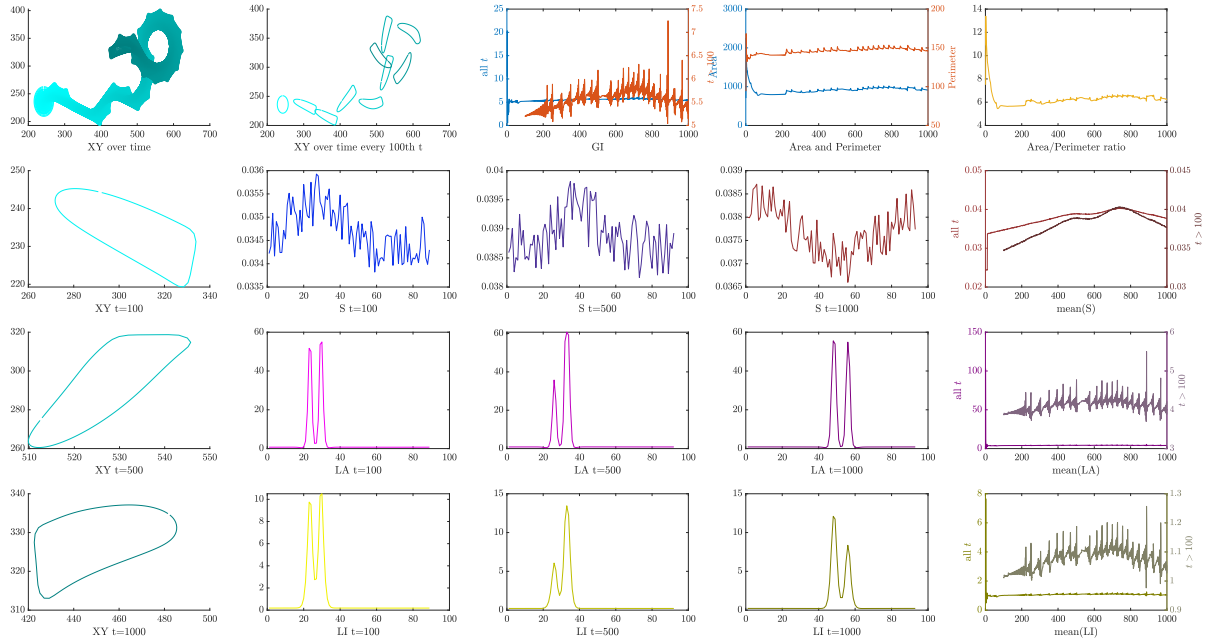

**(a)** GP predicted data.



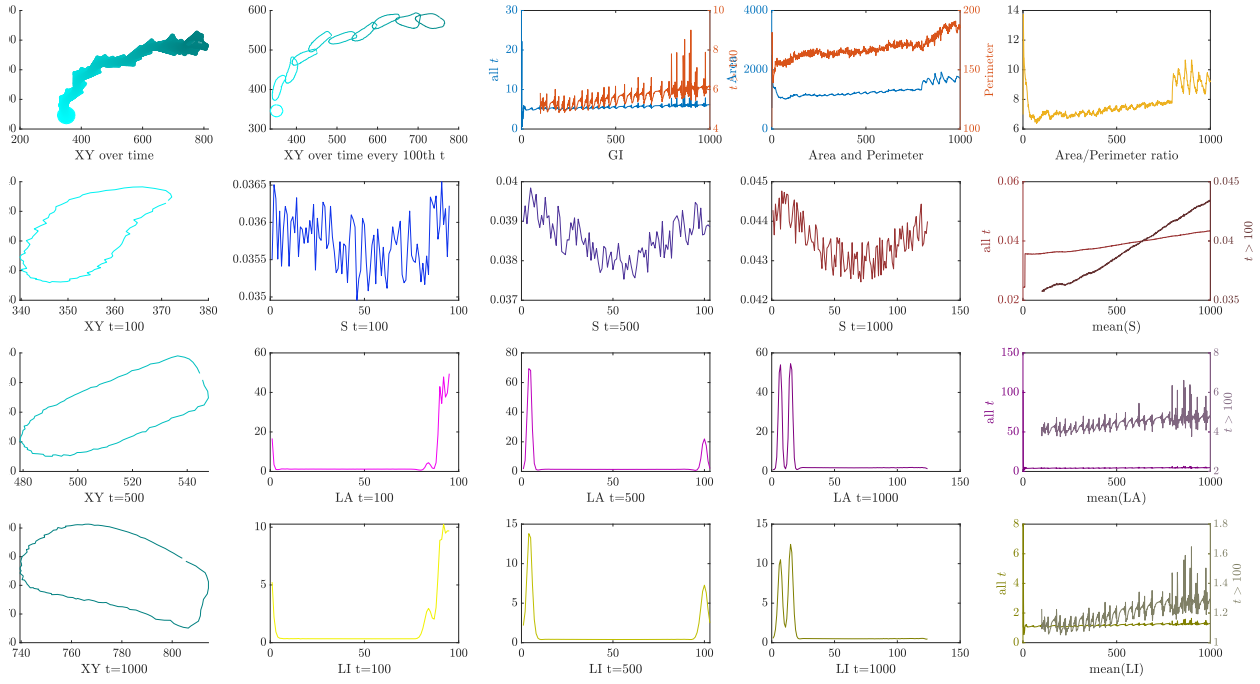

**(b)** Real data.

**Figure C.3:** *Fully Observed Data.*

## ABC predicted



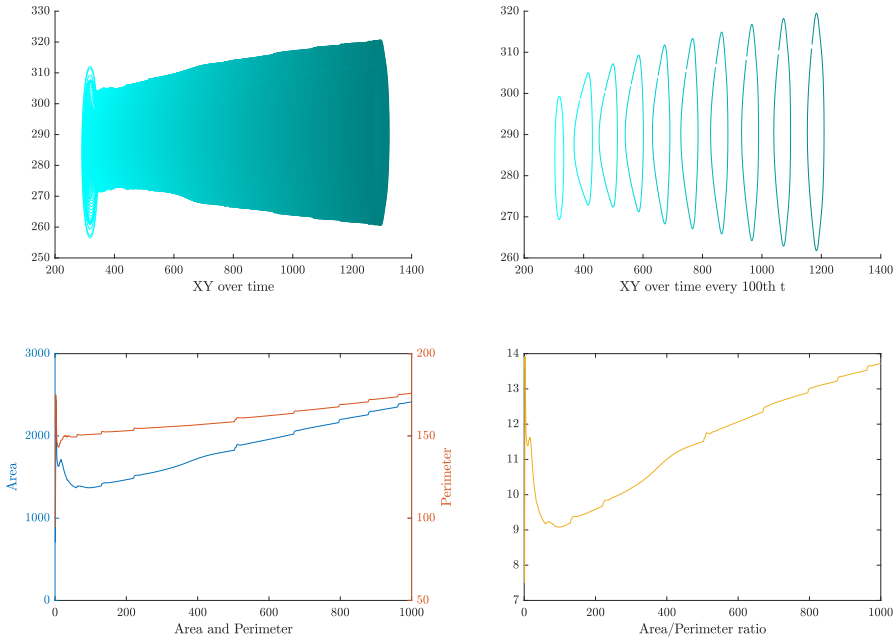

**(a)** ABC predicted data.



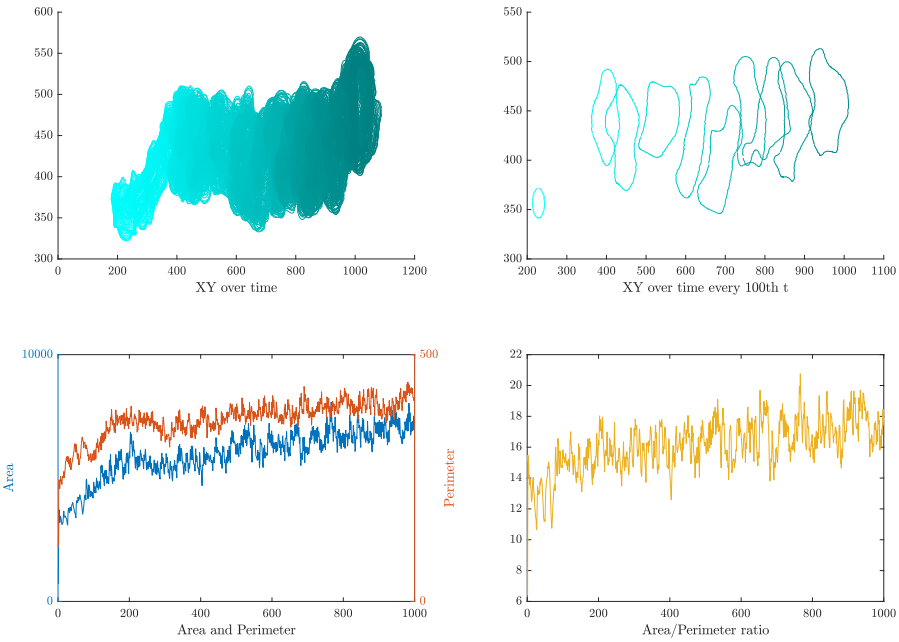

**(b)** Real data.

**Figure C.4:** *Fully Observed Data.*

## GP predicted



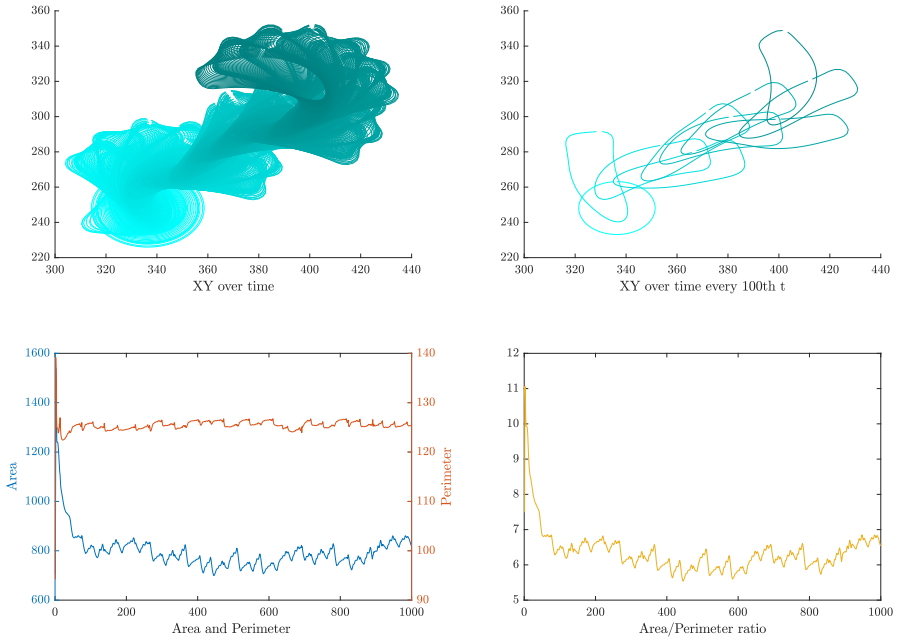

**(a)** GP predicted data.



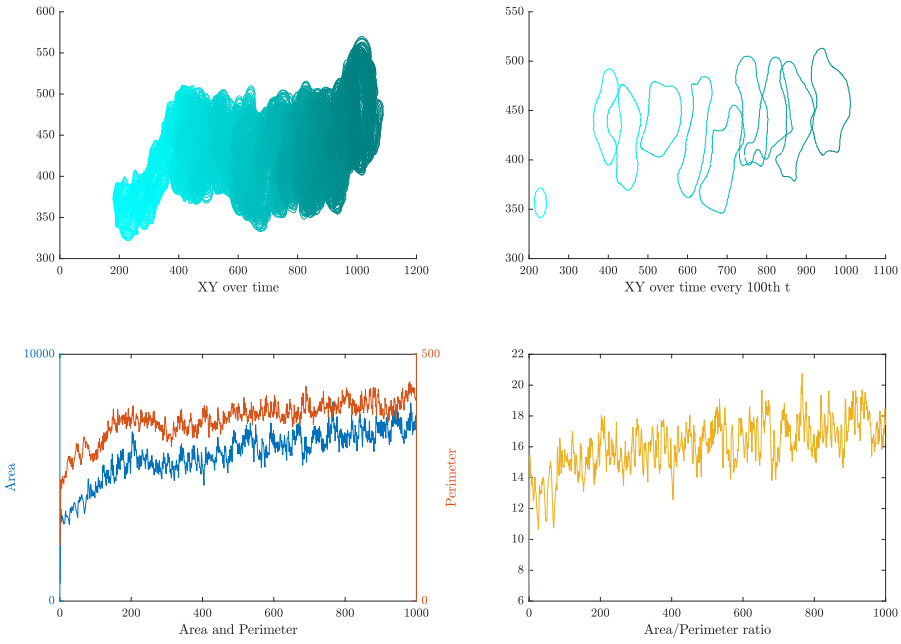

**(b)** Real data.

**Figure C.5:** *Partially Observed Data.*

## ABC predicted




**(a)** ABC predicted data.




**(b)** Real data.

**Figure C.6:** *Partially Observed Data.*