Tackling the challenges of fitting movement models to marine data

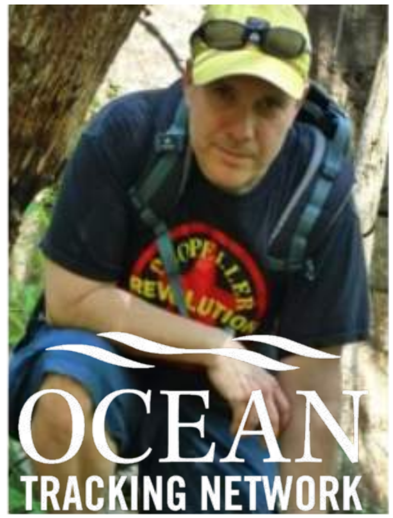
Marie Auger-Méthé
Dalhousie University

OCEAN TRACKING NETWORK

Co-authors

Joanna Mills Flemming

Ian Jonsen



Christoffer Albertsen

Andrew Derocher

Glenn Crossin

Katharine Studholme

Predicting the impacts of environmental changes

Photos: Dennis Bromage, David McNew

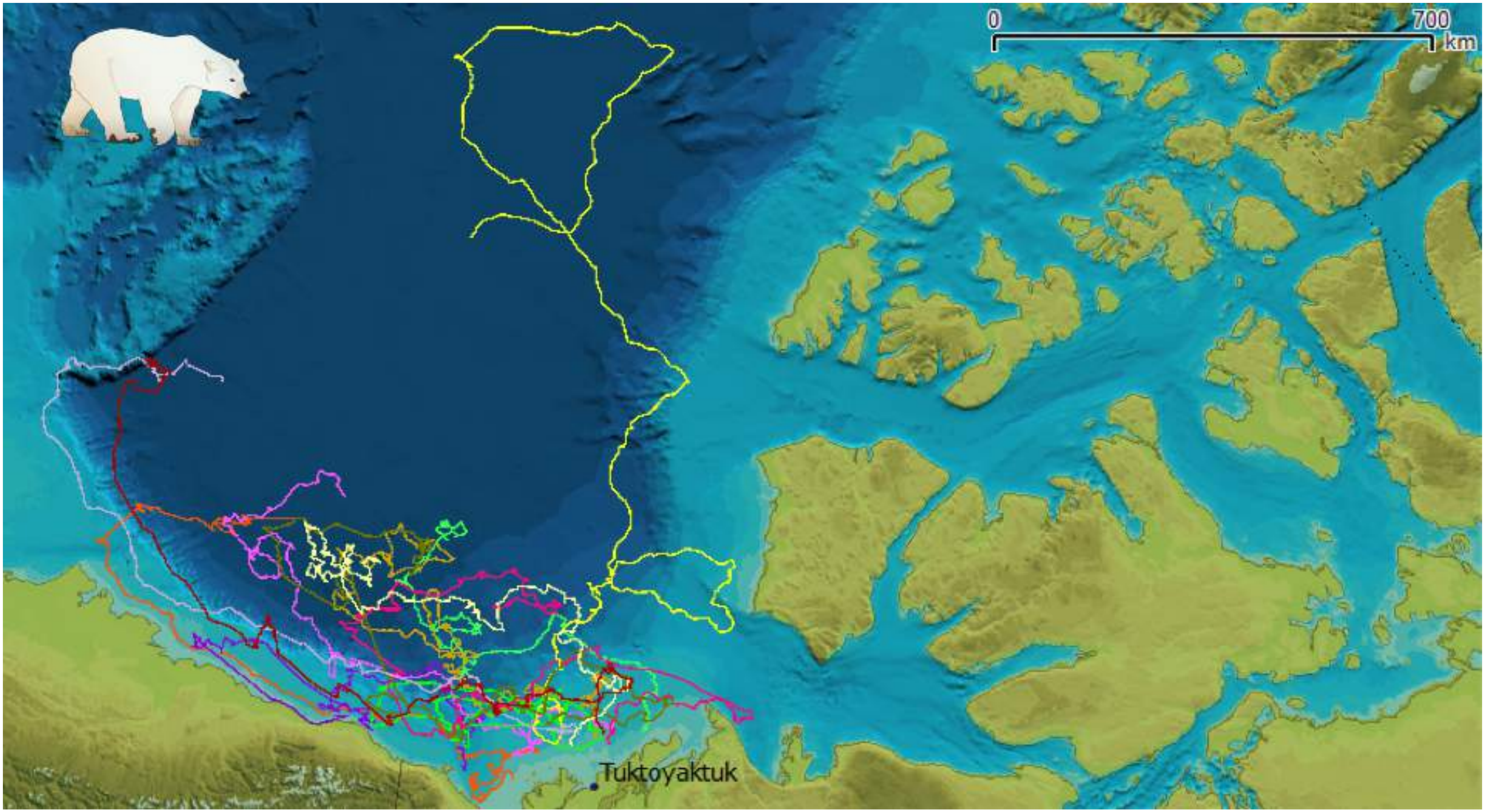
Monitor behaviours

Photos: Wayne Lynch, Ascension Island Turtle Group

Telemetry

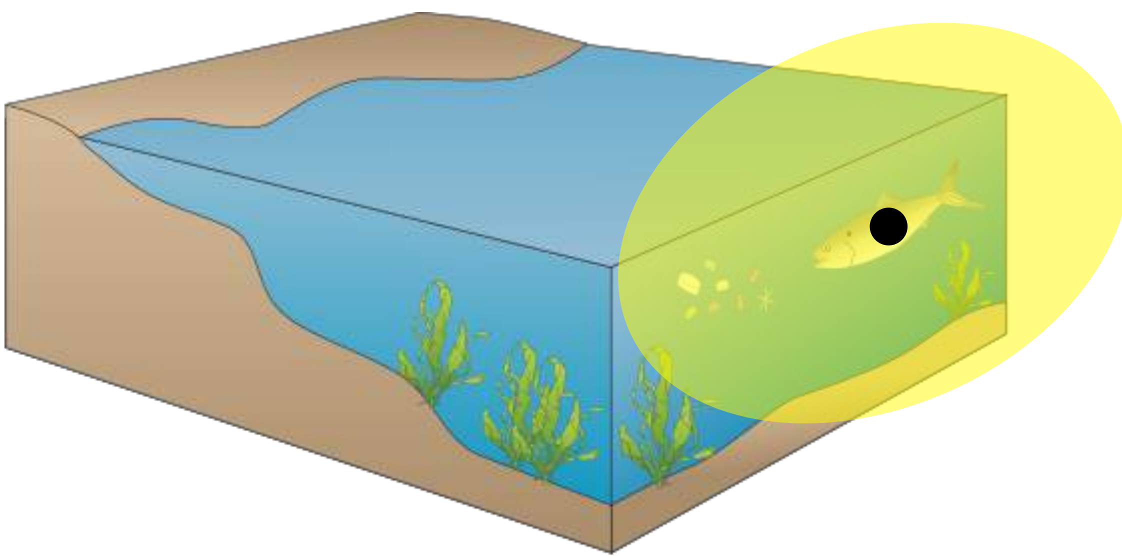
Photos: A. Park, C. Franklin, P. Lopez, N. Papathanasopoulou, R. Cameron, S. Anderson, R. Schuckard, L. Thorngren, L. Boehme,C. Jay

Apply models to movement data



Challenging for marine species

- Marine data have large measurement errors

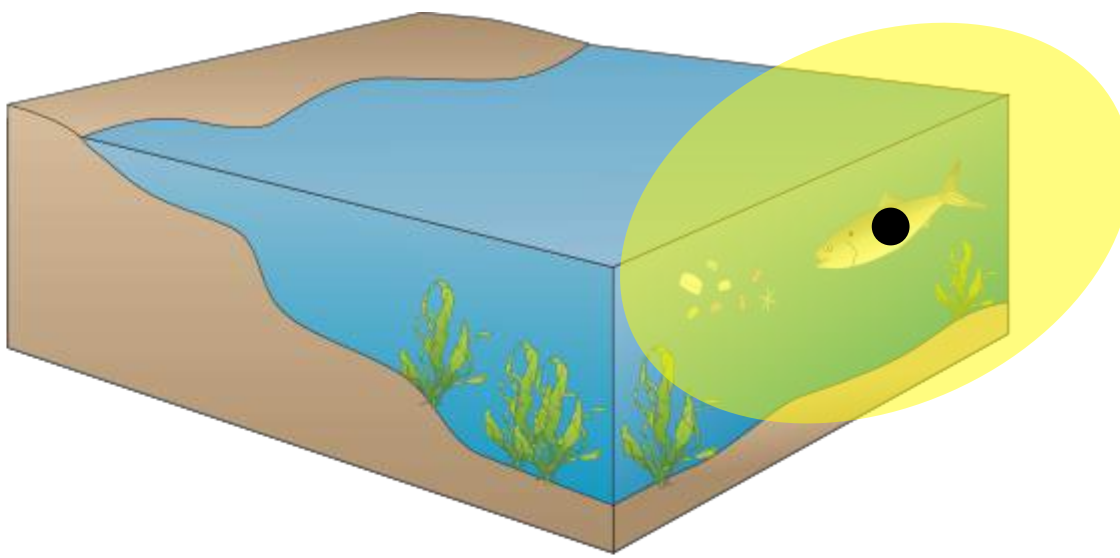


Challenging for marine species

- Marine data have large measurement errors

$$\text{Measurement eqn: } \mathbf{y}_t = f(\mathbf{x}_t) + \eta_t,$$

$$\text{Process eqn: } \mathbf{x}_t = g(\mathbf{x}_{t-1}) + \epsilon_t,$$

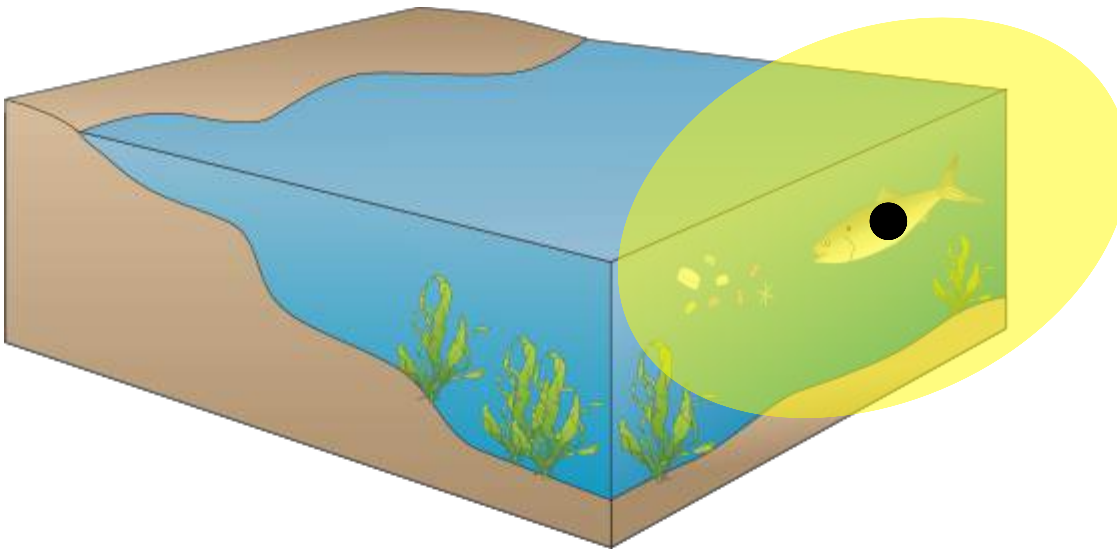


Challenging for marine species

- Marine data have large measurement errors

Measurement eqn: $\mathbf{y}_t = f(\mathbf{x}_t) + \eta_t,$

Process eqn: $\mathbf{x}_t = g(\mathbf{x}_{t-1}) + \epsilon_t,$

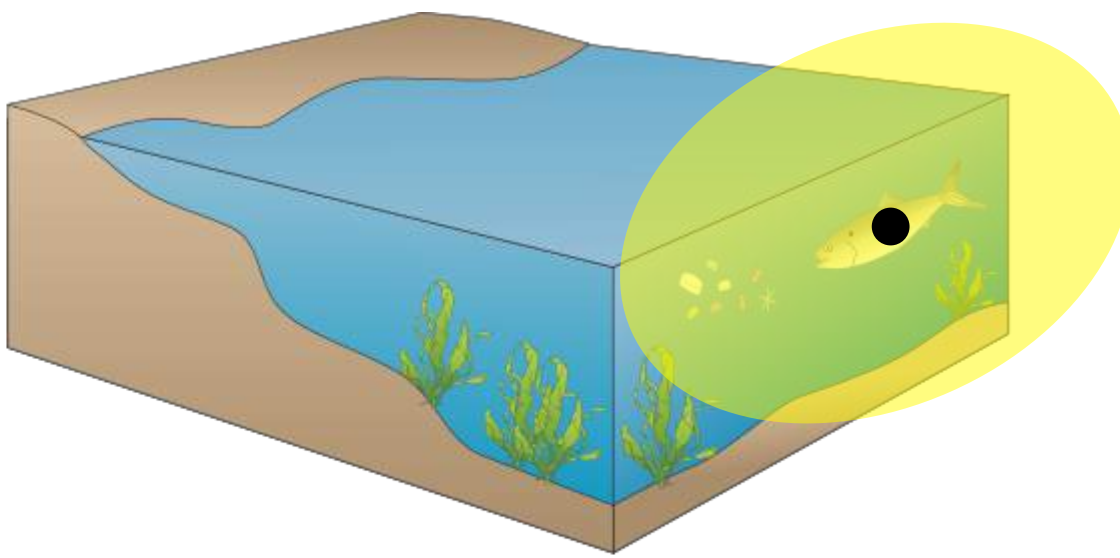


Challenging for marine species

- Marine data have large measurement errors

$$\text{Measurement eqn: } \mathbf{y}_t = f(\mathbf{x}_t) + \eta_t,$$

$$\text{Process eqn: } \mathbf{x}_t = g(\mathbf{x}_{t-1}) + \epsilon_t,$$



Challenging for marine species

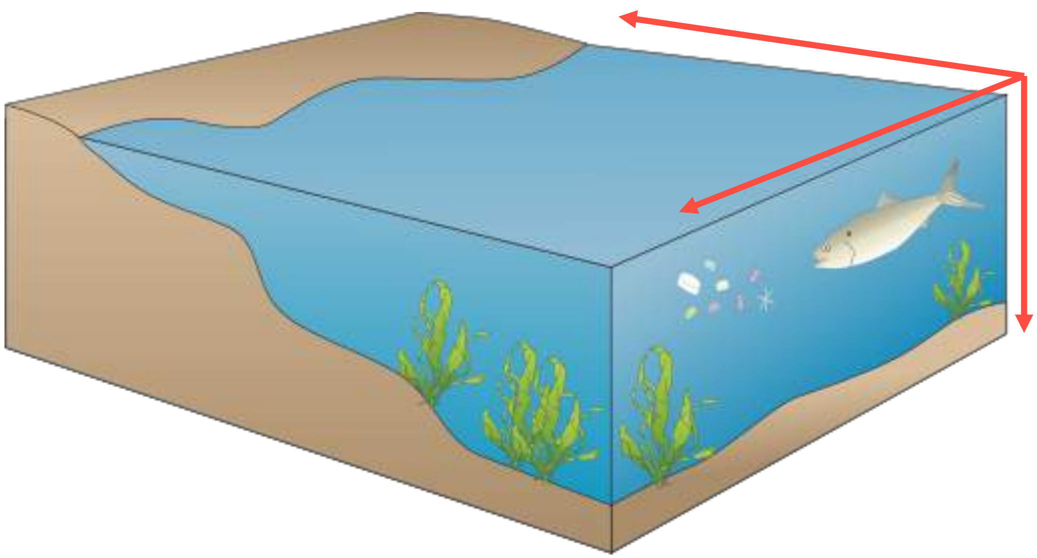
- Marine habitat has 3 dimensions

$$\text{Measurement eqn: } \mathbf{y}_t = f(\mathbf{x}_t) + \eta_t,$$

$$\text{Process eqn: } \mathbf{x}_t = g(\mathbf{x}_{t-1}) + \epsilon_t,$$

$$\mathbf{y}_t = \begin{pmatrix} y_{t,\text{lat}} \\ y_{t,\text{lon}} \\ y_{t,\text{depth}} \end{pmatrix}$$

$$\mathbf{x}_t = \begin{pmatrix} x_{t,\text{lat}} \\ x_{t,\text{lon}} \\ x_{t,\text{depth}} \end{pmatrix}$$



Challenging for marine species

- Dynamic habitat with currents

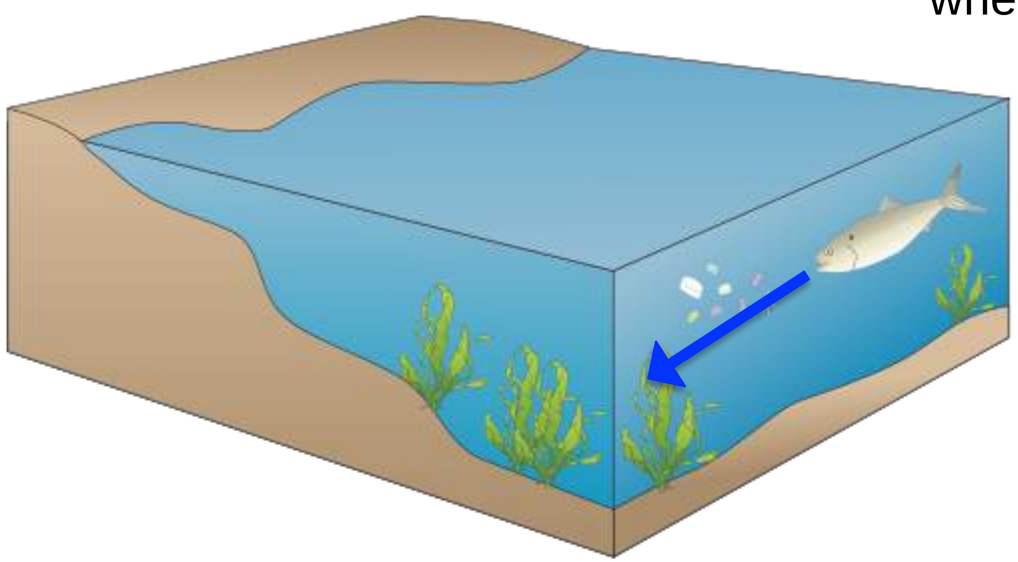
$$\text{Measurement eqn: } y_t = f(x_t, z_t) + \eta_t$$

Affect observed movement

$$\text{Process eqn: } x_t = g(x_{t-1}, z_t) + \epsilon_t$$

Affect voluntary movement

z_t current data at time t,
where the animal is observed



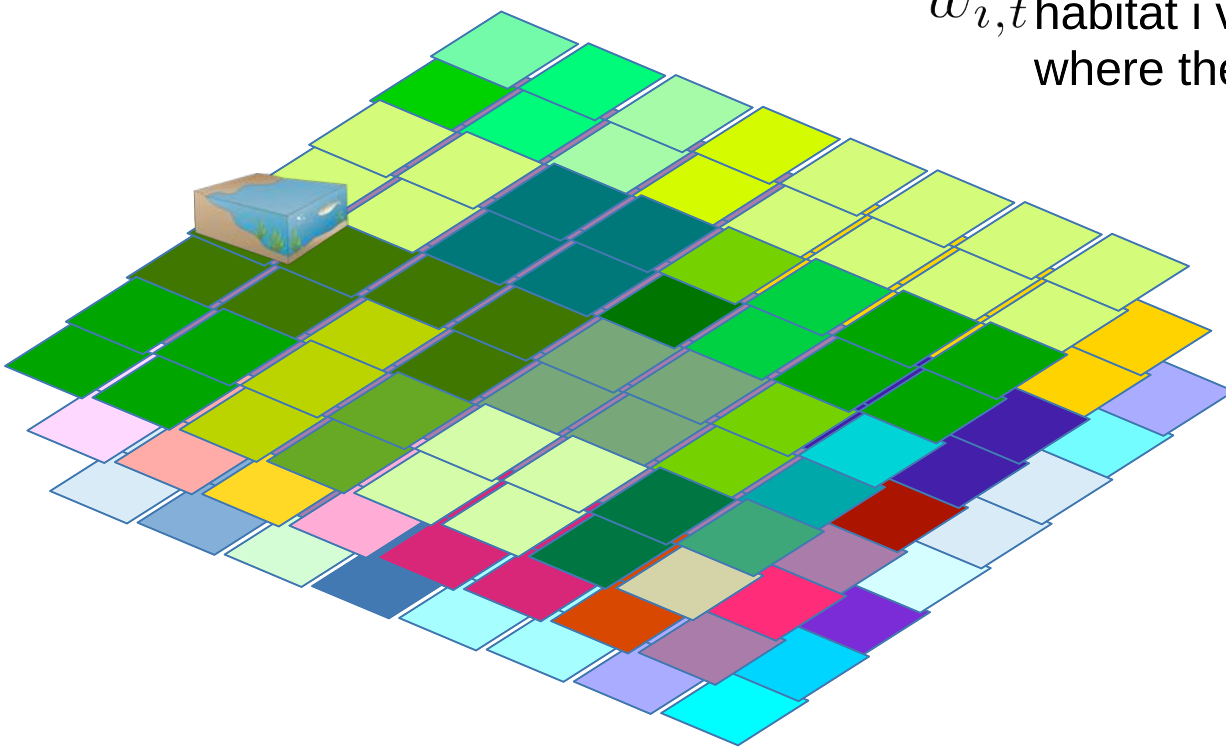
Challenging for marine species

- Interaction with habitat

$$\text{Measurement eqn: } y_t = f(x_t) + \eta_t$$

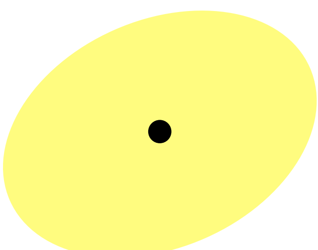
$$\text{Process eqn: } x_t = g(x_{t-1}, w_{1,t}, w_{2,t}, w_{3,t}) + \epsilon_t$$

$w_{i,t}$ habitat i value at time t where the animal is observed

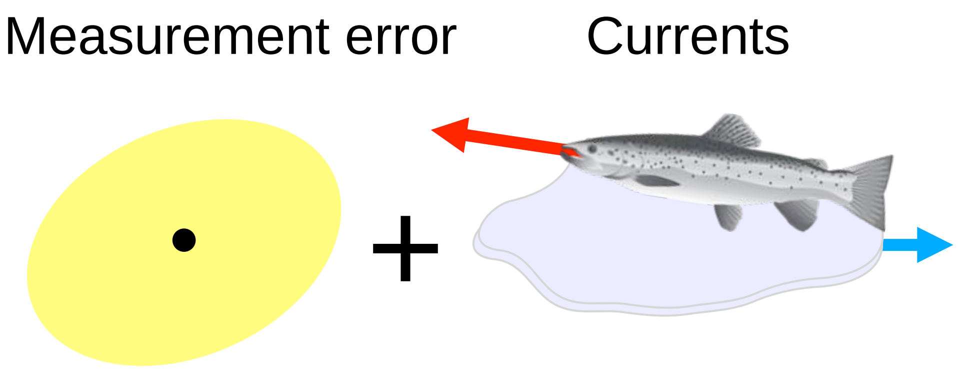


Challenging for marine species

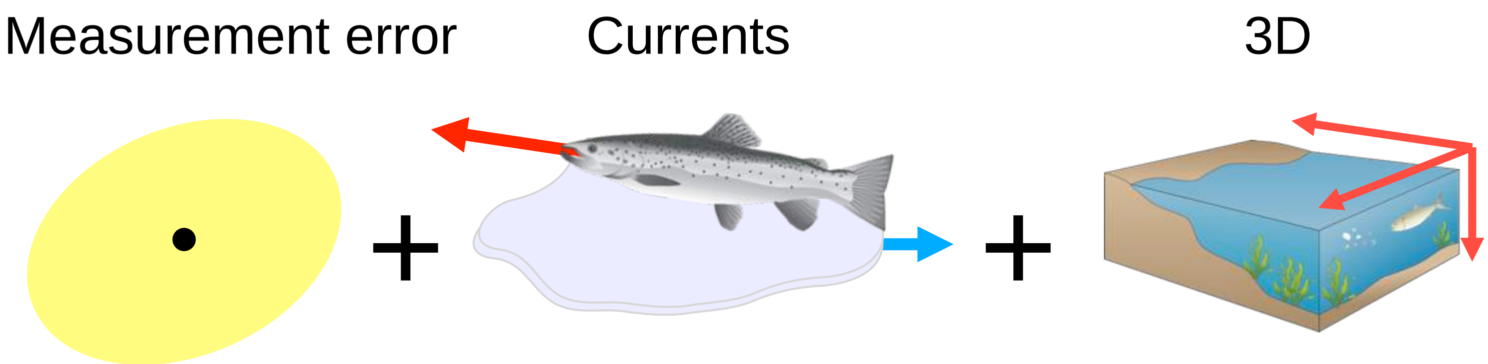
Measurement error



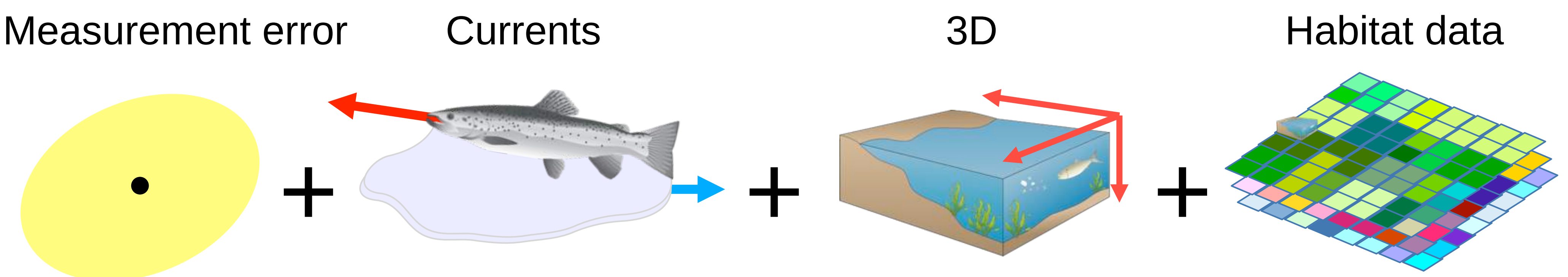
Challenging for marine species



Challenging for marine species

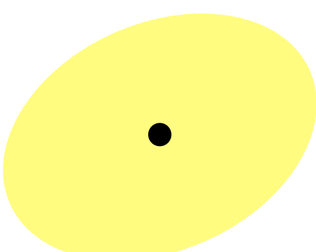


Challenging for marine species



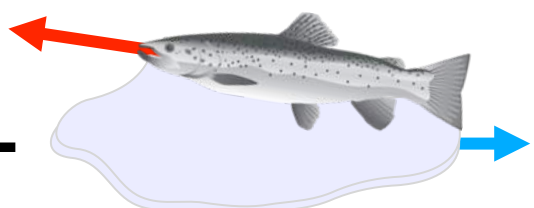
Challenging for marine species

Measurement error



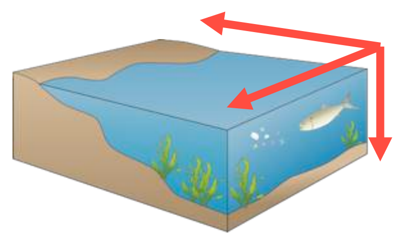
+

Currents



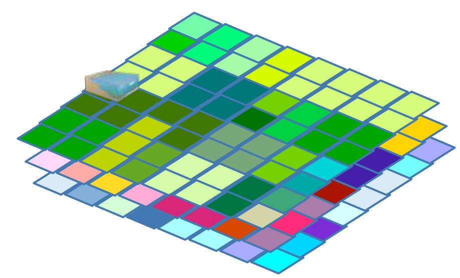
+

3D



+

Habitat data



=

complex model

Maximising the likelihood

$$L(\Theta|\mathbf{y}) = \int_{R^n} p(\mathbf{y}|\mathbf{x},\Theta)p(\mathbf{x}|\Theta)d\mathbf{x}$$

parameters to estimate: $\Theta \in R^m$

unobserved states: $\mathbf{x} \in R^n$

Maximising the likelihood

$$L(\Theta|\mathbf{y}) = \int_{R^n} p(\mathbf{y}|\mathbf{x},\Theta)p(\mathbf{x}|\Theta)d\mathbf{x}$$

parameters to estimate: $\Theta \in R^m$

unobserved states: $\mathbf{x} \in R^n$

- Estimating parameters

  -Maximum likelihood estimation (MLE)

$$\widehat{\Theta} = \arg\max_{\Theta} L(\Theta|\mathbf{y})$$

Maximising the likelihood

$$L(\Theta|\mathbf{y}) = \int_{R^n} p(\mathbf{y}|\mathbf{x},\Theta)p(\mathbf{x}|\Theta)\,d\mathbf{x}$$

parameters to estimate: $\Theta \in R^m$

unobserved states: $\mathbf{x} \in R^n$

- Estimating parameters
  -Maximum likelihood estimation (MLE)

$$\widehat{\Theta} = \arg\max_{\Theta} L(\Theta|\mathbf{y})$$

Maximising the likelihood

$$L(\Theta|\mathbf{y}) = \int_{R^n} p(\mathbf{y}|\mathbf{x},\Theta)p(\mathbf{x}|\Theta)d\mathbf{x}$$

parameters to estimate: $\Theta \in R^m$

unobserved states: $\mathbf{x} \in R^n$

- Estimating parameters
  -Maximum likelihood estimation (MLE)

$$\widehat{\Theta} = \arg\max_{\Theta} L(\Theta|\mathbf{y})$$

- Varied methods:
  -Analytical solution
  -Simulation based (Particle Filter, Bayesian MCMC)

Maximising the likelihood

$$L(\Theta|\mathbf{y}) = \int_{R^n} p(\mathbf{y}|\mathbf{x}, \Theta) p(\mathbf{x}|\Theta) d\mathbf{x}$$

parameters to estimate: $\Theta \in R^m$

unobserved states: $\mathbf{x} \in R^n$

- Estimating parameters
  - -Maximum likelihood estimation (MLE)

$$\widehat{\Theta} = \arg\max_{\Theta} L(\Theta|\mathbf{y})$$

- Varied methods:
  - -Analytical solution
  - -Simulation based (Particle Filter, Bayesian MCMC)

Slow!

Template model builder (TMB)

- Developed by Kasper Kristensen
- C++ template with R interface

Assessing TMB as a tool to fit movement models to marine data

- Comparing TMB to Bayesian MCMC (JAGS)
- Identifying robust models in TMB

First difference correlated random walk

Measurement eqn: $\mathbf{y}_t = \mathbf{x}_t + \eta_t$

$$\eta_{t,\mathrm{E}} \sim t(\psi\sigma_{\eta,\mathrm{E},j}, df_{\mathrm{E},j})$$
$$\eta_{t,\mathrm{N}} \sim t(\psi\sigma_{\eta,\mathrm{N},j}, df_{\mathrm{N},j})$$

Process eqn: $\mathrm{x}_t = \mathbf{x}_{t-1} + \gamma\mathbf{T}(\theta)(\mathbf{x}_{t-1} - \mathbf{x}_{t-2}) + \epsilon_t, \quad \epsilon_{\mathbf{t}} \sim N(0, \Sigma_\epsilon)$

First difference correlated random walk

Measurement eqn: $\mathbf{y}_t = \mathbf{x}_t + \eta_t$

$$\eta_{t,\mathrm{E}} \sim t(\psi \sigma_{\eta,\mathrm{E},j}, df_{\mathrm{E},j})$$
$$\eta_{t,\mathrm{N}} \sim t(\psi \sigma_{\eta,\mathrm{N},j}, df_{\mathrm{N},j})$$

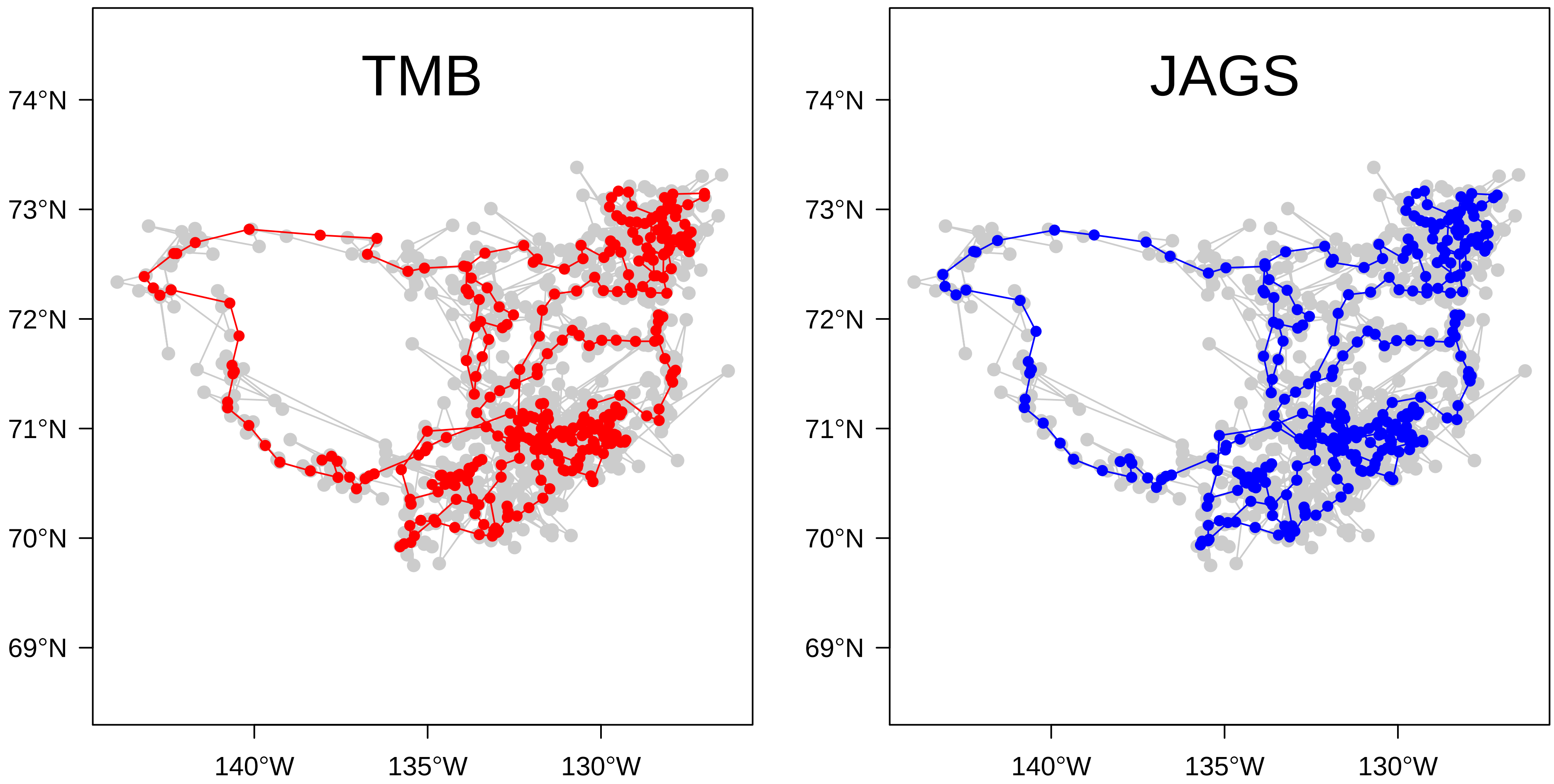
Process eqn: $\mathrm{x}_t = \mathbf{x}_{t-1} + \gamma \mathbf{T}(\theta)(\mathbf{x}_{t-1} - \mathbf{x}_{t-2}) + \epsilon_t, \quad \epsilon_{\mathbf{t}} \sim N(0, \Sigma_\epsilon)$

Polar bear GPS & Argos

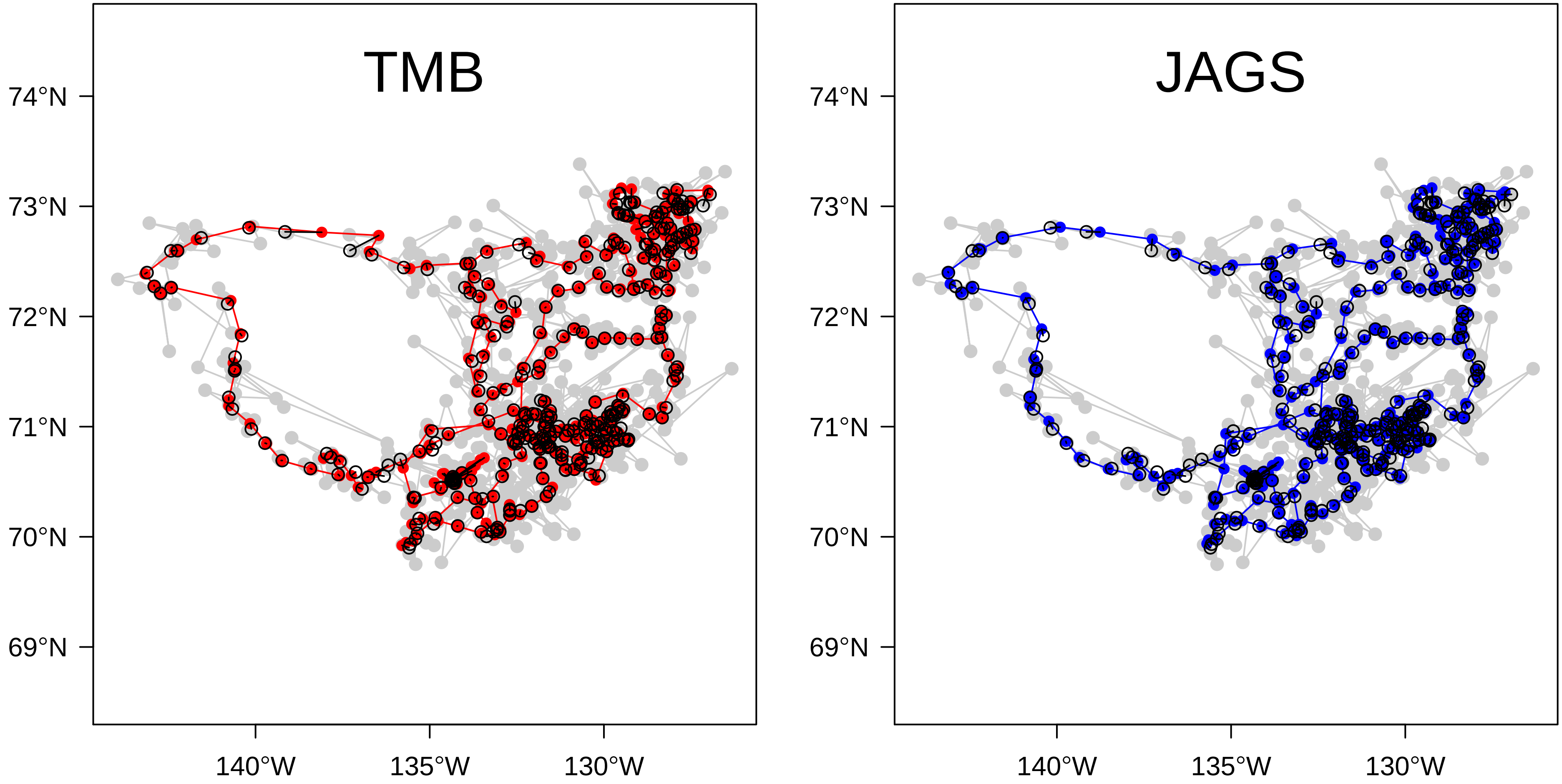
- Argos data:
  - large measurement error!
  - use for many marine animals

- GPS data:
  - “true” states

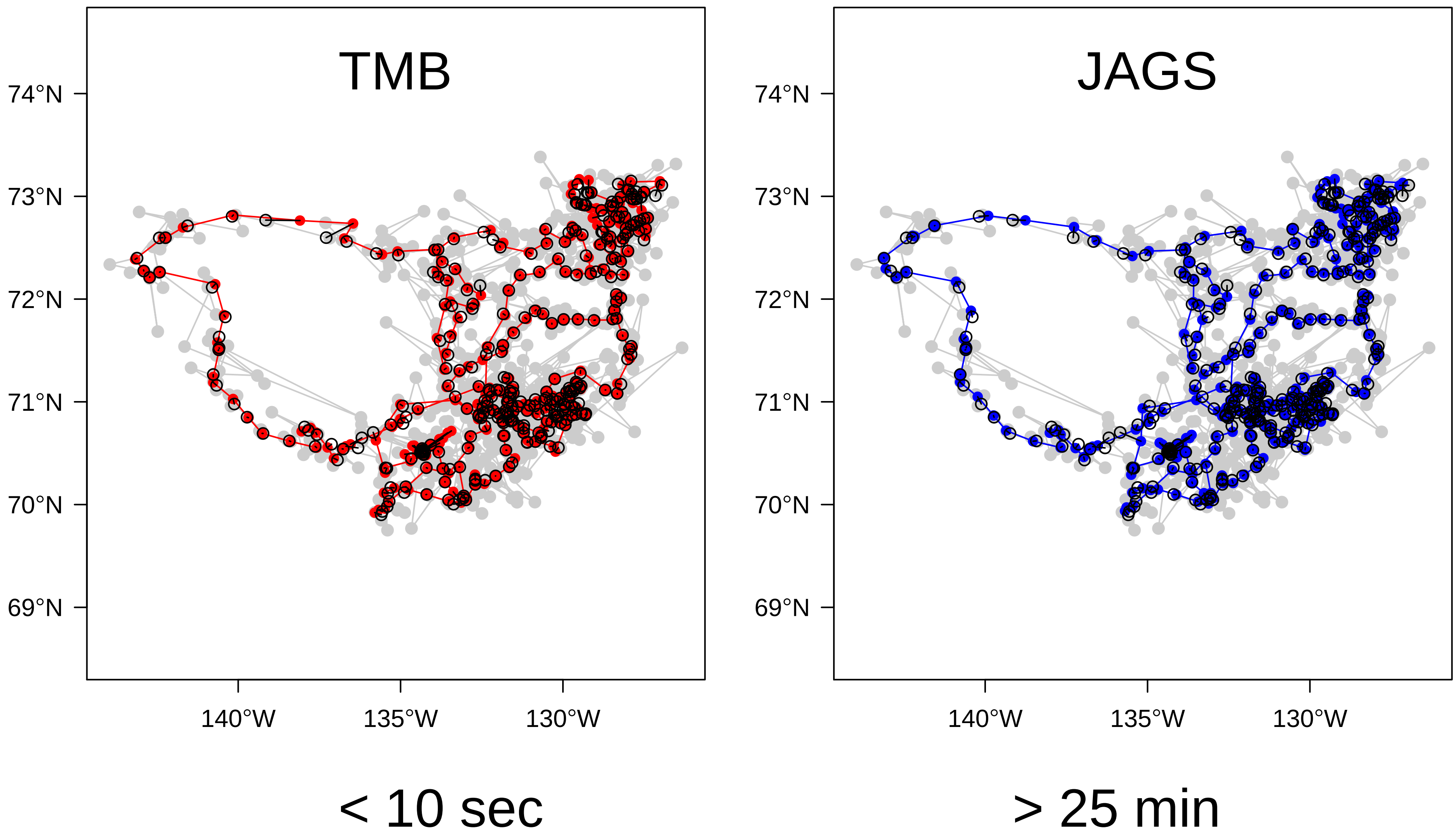
Similar parameter and state estimates



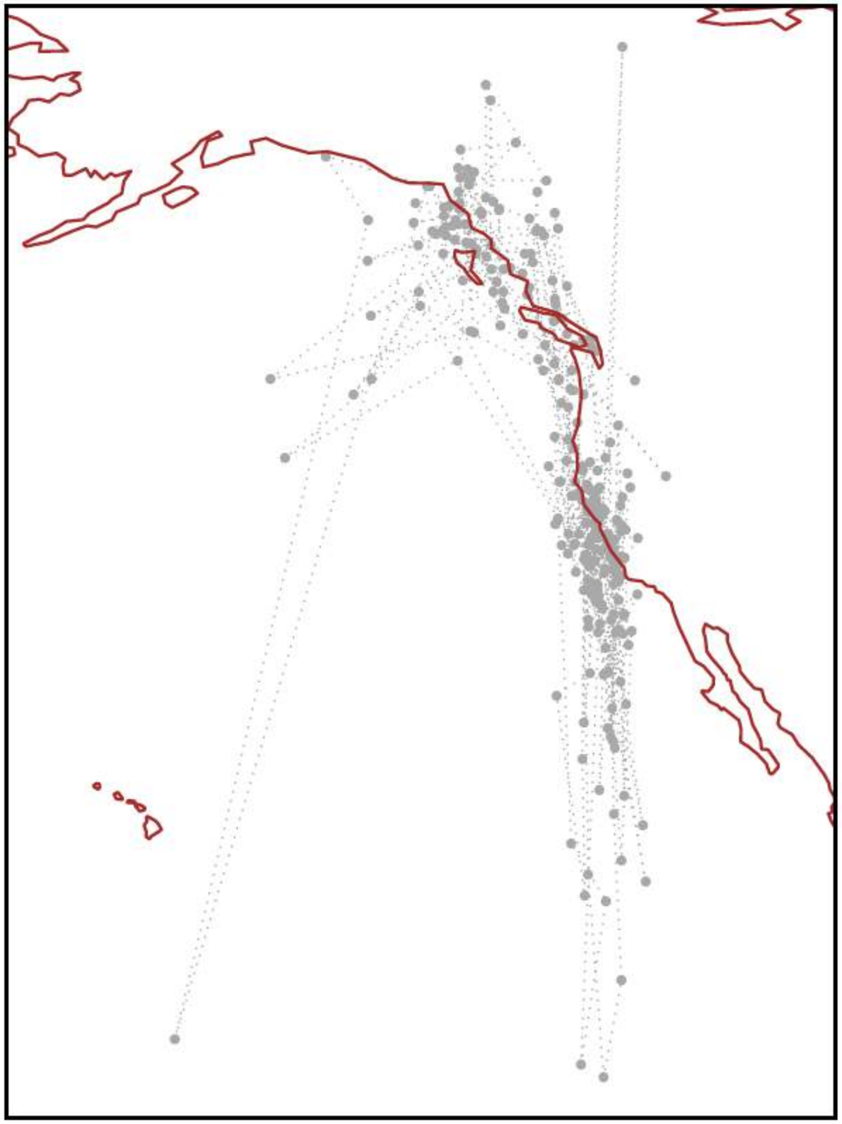
Similar parameter and state estimates



Similar parameter and state estimates but much faster



Rhinoceros auklet light geolocation



First difference correlated random walk

Measurement eqn: $\mathbf{y}_t = \mathbf{x}_t + \eta_t$

Process eqn: $\mathrm{x}_t = \mathbf{x}_{t-1} + \gamma \mathbf{T}(\theta)(\mathbf{x}_{t-1} - \mathbf{x}_{t-2}) + \epsilon_t, \quad \epsilon_{\mathbf{t}} \sim N(0, \Sigma_\epsilon)$

First difference correlated random walk

Measurement eqn: $\mathbf{y}_t = \mathbf{x}_t + \eta_t$

Process eqn: $\mathrm{x}_t = \mathbf{x}_{t-1} + \gamma \mathbf{T}(\theta)(\mathbf{x}_{t-1} - \mathbf{x}_{t-2}) + \epsilon_t, \quad \epsilon_{\mathbf{t}} \sim N(0, \Sigma_\epsilon)$

Model 1:

$$\eta_{\mathbf{t}} \sim N(0, \Sigma_\eta)$$

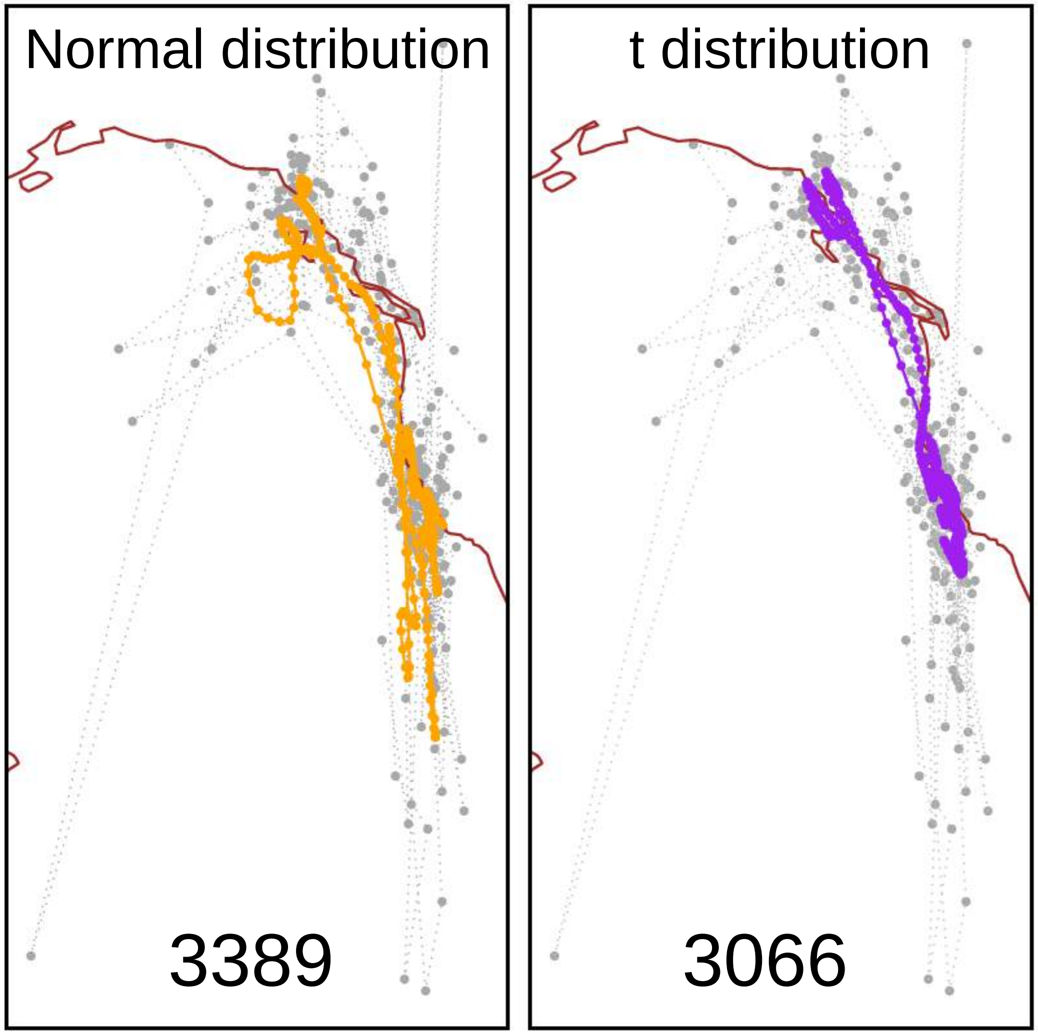
Model 2:

$$\eta_{t,\mathrm{E}} \sim t(\sigma_{\eta,\mathrm{E}}, df_{\mathrm{E}})$$
$$\eta_{t,\mathrm{N}} \sim t(\sigma_{\eta,\mathrm{N}}, df_{\mathrm{N}})$$

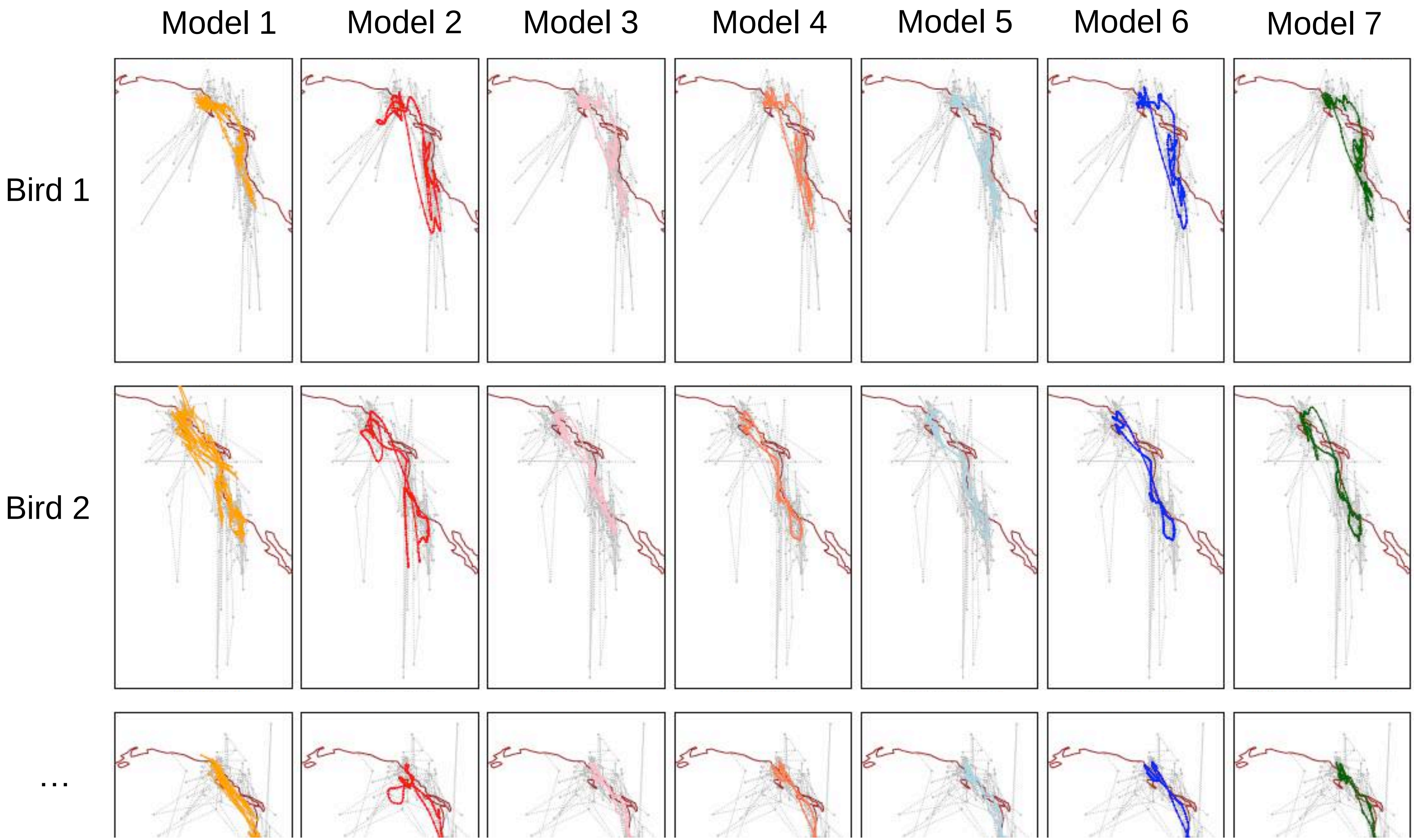
Model selection

$$\mathrm{AIC} = -2\ln L(\widehat{\Theta}|\mathbf{y}) + 2k$$

AIC:



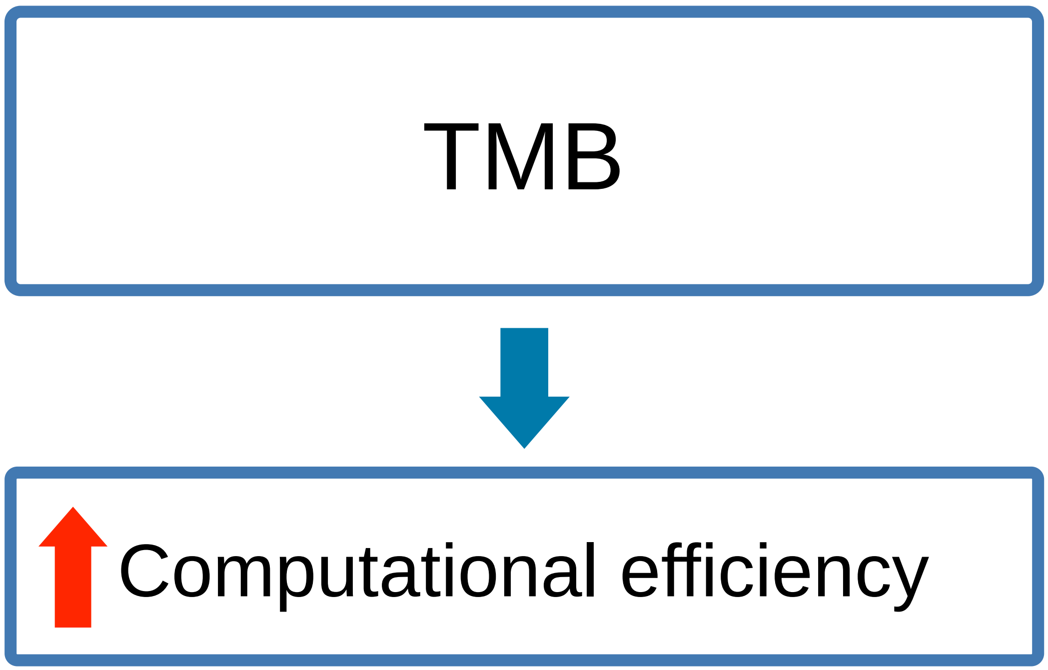
Can compare many models



Conclusion

TMB

Conclusion



Conclusion

TMB

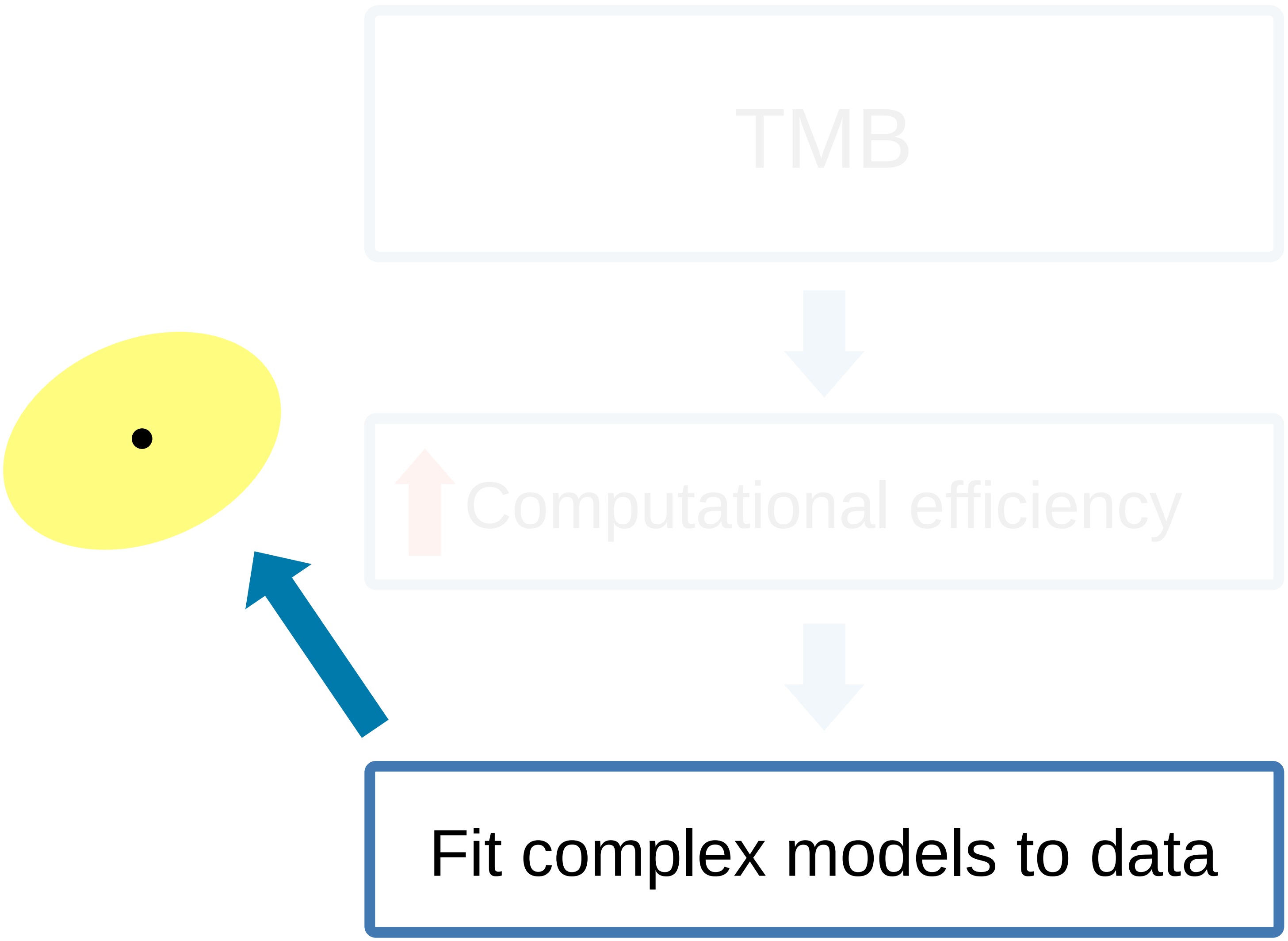
↓

⬆ Computational efficiency

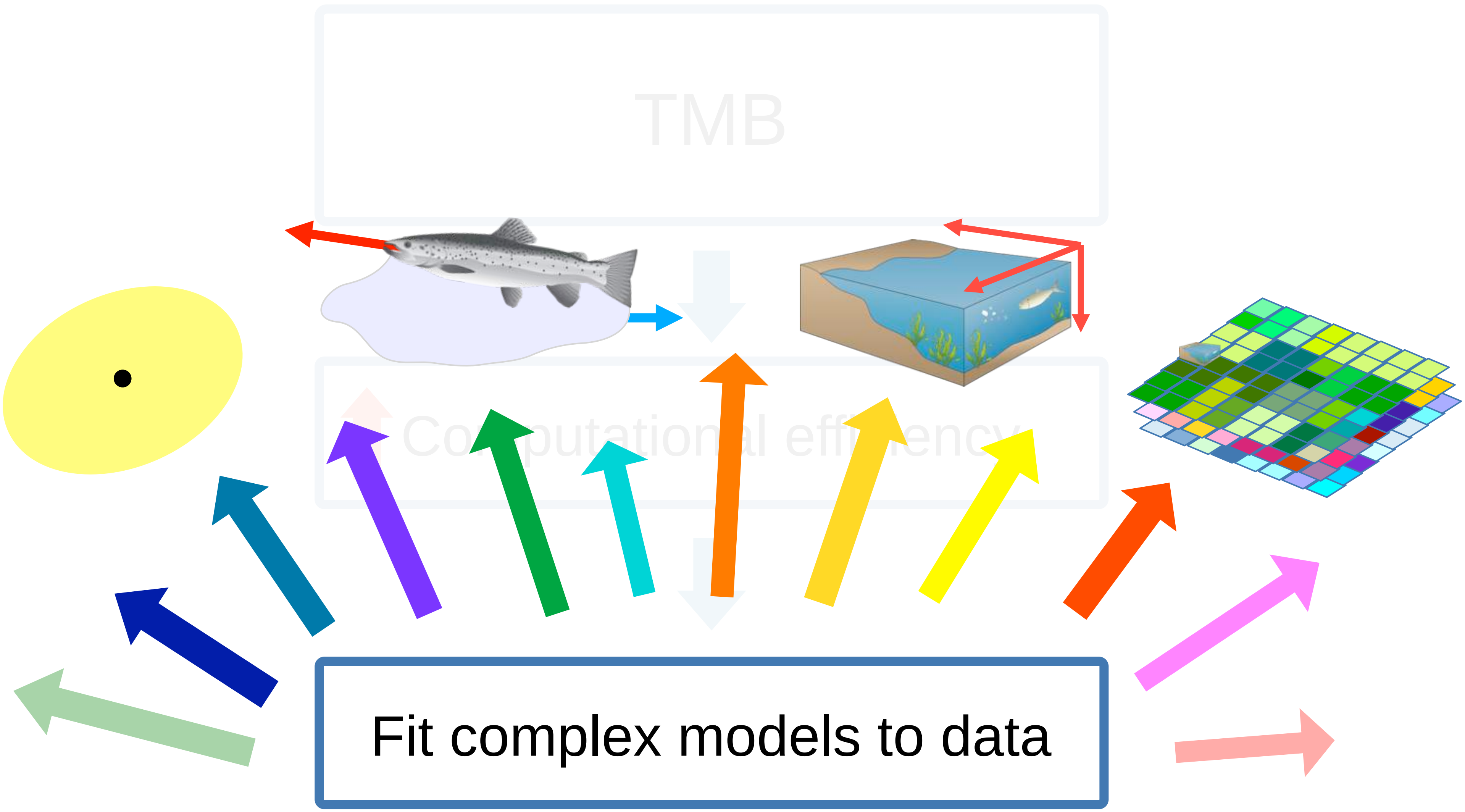
↓

Fit complex models to data

Conclusion



Conclusion



CRSNG
NSERC

INNOVATION.CA
CANADA FOUNDATION FOR INNOVATION | FONDATION CANADIENNE POUR L'INNOVATION

Thanks!
Questions?