Pelagic Fisheries Research Program

http://www.soest.hawaii.edu/PFRP/

A mosaic of models for light-based geolocation: How to choose, what to be careful about, and future directions

Anders Nielsen & John Sibert

anders.nielsen@hawaii.edu

Light based geolocation

- We got:
  - Light, depth, and temperature
  - Measured for instance every minute
  - From archival tags the entire record can be retrieved
  - From satellite transmitting tags only a summary
- We want:
  - A track of geographic positions (geolocations)
  - Some idea about the uncertainties
  - Perhaps some quantitative movement parameters
- Problems:
  - Indirect measurements: Light $\rightarrow$ solar altitude $\rightarrow$ geolocation
  - High and correlated uncertainties from changing weather and incomplete depth corrections

This talk

- Will talk about:
  - Raw geolocations
  - Kftrack
  - Kfsst/ukfsst
  - Trackit (with and without SST)
- Will not talk about:
  - Satellite methods
  - Tidal location models
  - Sunrise/sunset times models
  - SST matching algorithms
  - EASy-FishTracker
  - ...

Similarities of kftrack, kfsst, ukfsst, and trackit

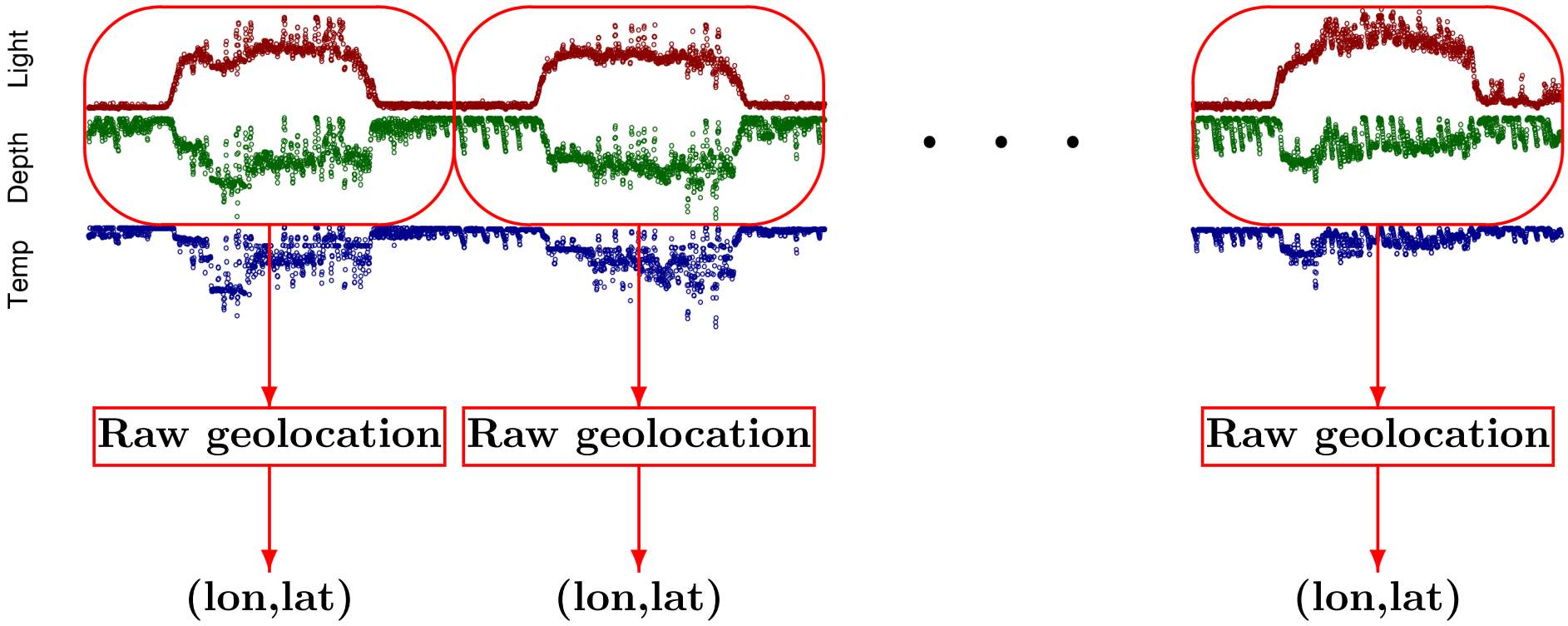
- Underlying movement model
  - Assume the same random walk model
  - Includes drift and diffusion
- At any given position the observation model
  - Predicts the observation
  - Describes the observation error
- All model parameters are maximum likelihood estimated
- Any point on the most probable track is a weighted average of:
  - What is learned from the current observation
  - What is learned from the entire track

The differences are

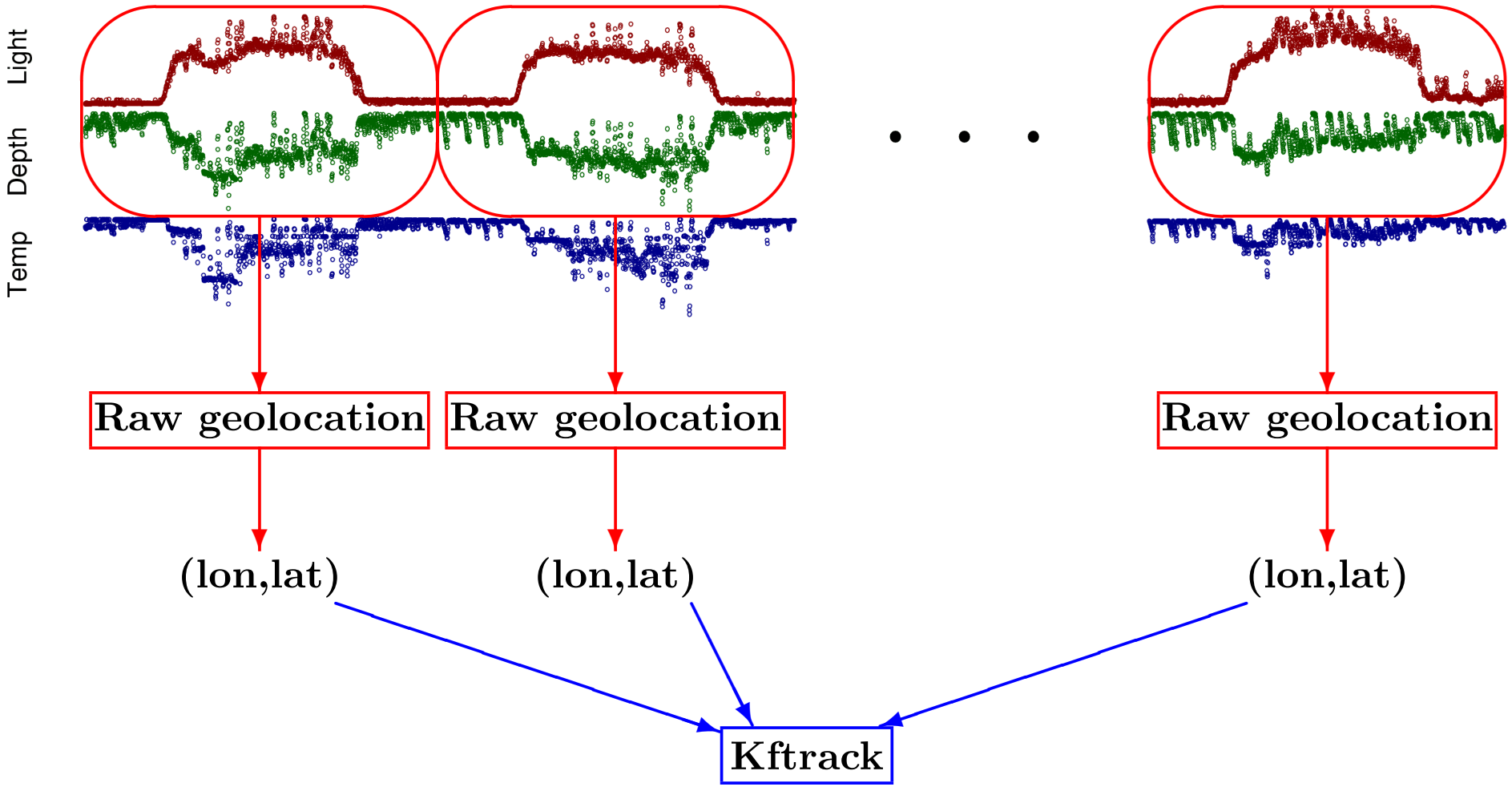
- What they take as data
  - Raw geolocations (lon,lat) used by kftrack
  - Raw geolocations and SST (lon,lat,SST) used by kfsst and ukfsst
  - Light readings and SST used by trackit
- Technical details in handling non-linearities
  - Extended Kalman filter used by kftrack and kfsst
  - Unscented Kalman filter used by ukfsst and trackit

	(lon,lat)	**SST**	**Light**	**EKF**	**UKF**
kftrack	X			X	
kfsst	X	X		X	
ukfsst	X	X			X
trackit		X	X		X

If possible, running more than one can often be very instructive ...

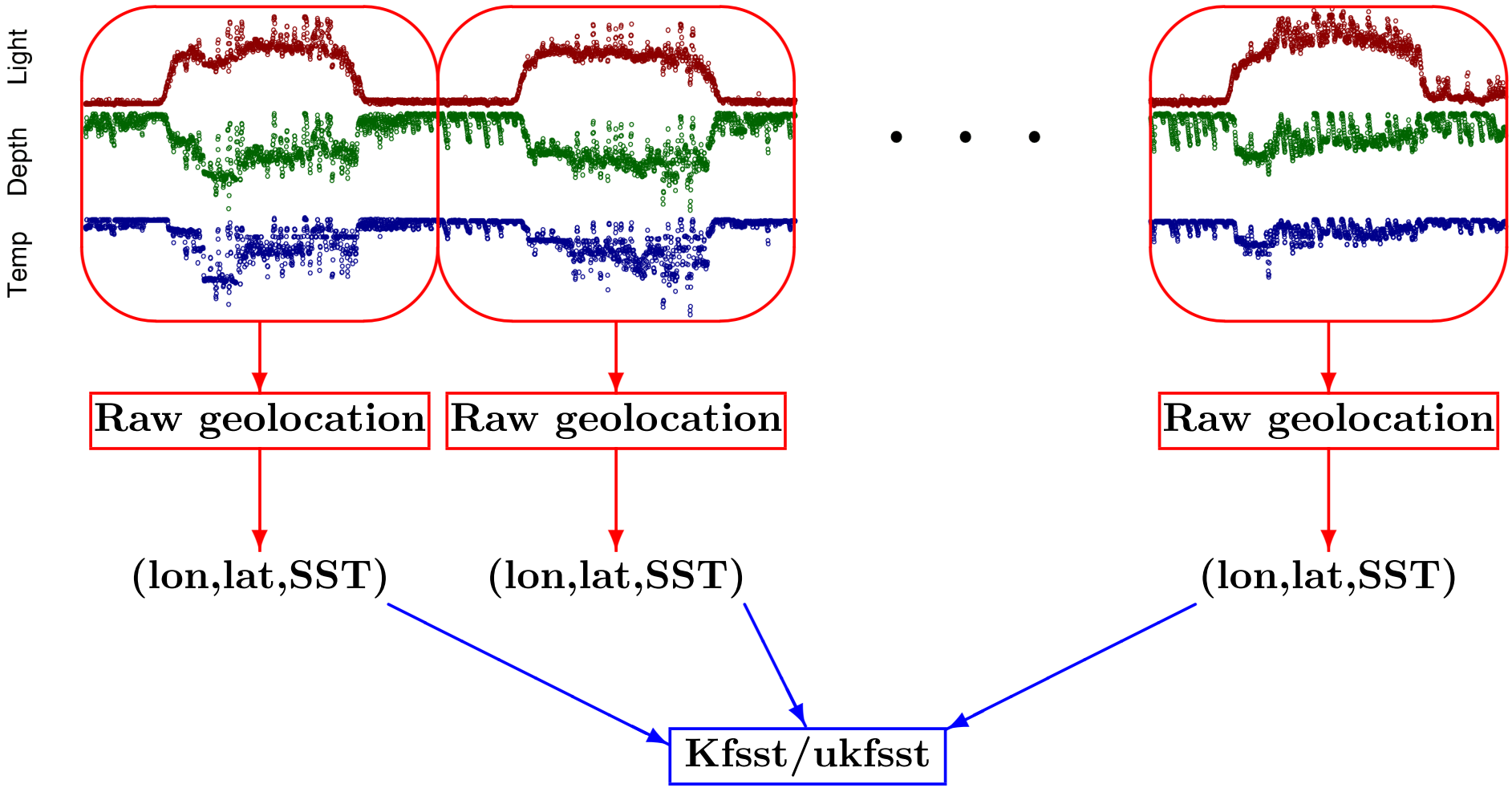


- Algorithms (partly) proprietary, but in essence:
  - Associate a certain light level crossing with a solar angle (problematic)
  - Calculate position from two of those
- These raw geolocations are often imprecise and biased
- Especially latitude around equinox



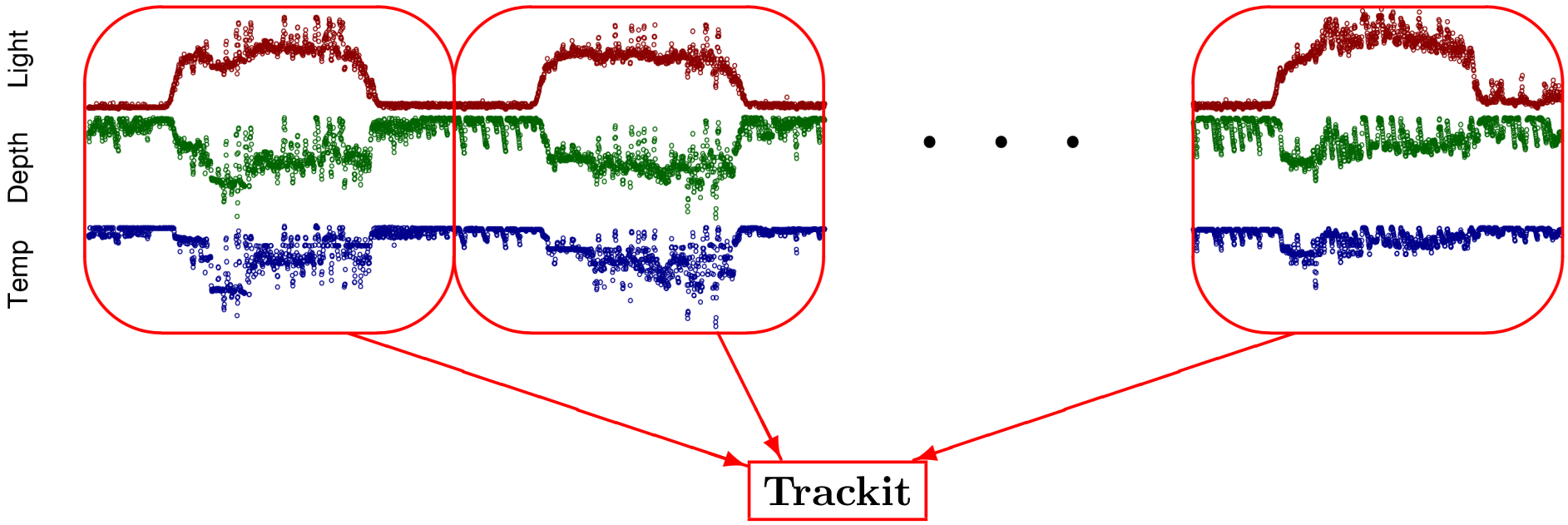

```
day month year lon lat
21 1 1999 201.750 18.65
22 1 1999 204.520 20.00
23 1 1999 206.086 22.00
24 1 1999 204.399 23.50
.
.
.
```

```
library(kftrack)
track<-read.table('datafile.dat', header=TRUE)
fit<-kftrack(track)
plot(fit, map=TRUE, ci=TRUE)
```




```
day month year lon lat sst
11 4 2001 201.722 18.875 24.73
16 4 2001 201.190 24.150 24.37
18 4 2001 202.950 12.890 24.73
22 4 2001 199.110 28.790 24.37
.
.
.
```

```
library(kfsst) or library(ukfsst)
track<-read.table('datafile.dat', header=TRUE)
sst.path<-get.sst.from.server(track)
sst.file<-write.sst.field(sst.path) delete line in ukfsst
fit<-kfsst(track)
plot(fit, map=TRUE, ci=TRUE)
```

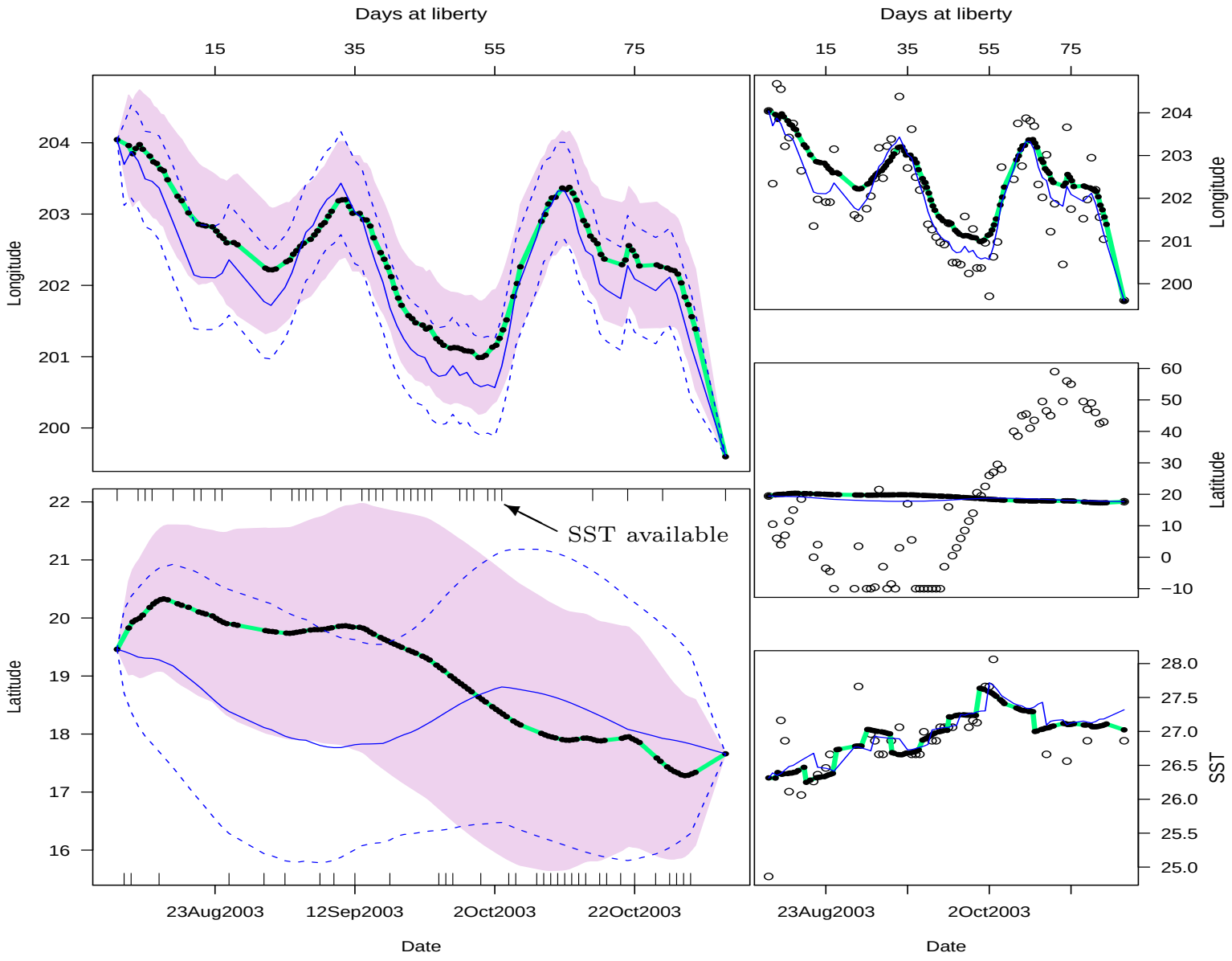




```
year month day hour min sec depth light temp
2002 9 11 15 35 53 15.5 69 21.1
2002 9 11 15 36 53 15.0 75 21.1
2002 9 11 15 37 53 16.0 85 21.3
.
.
.
```

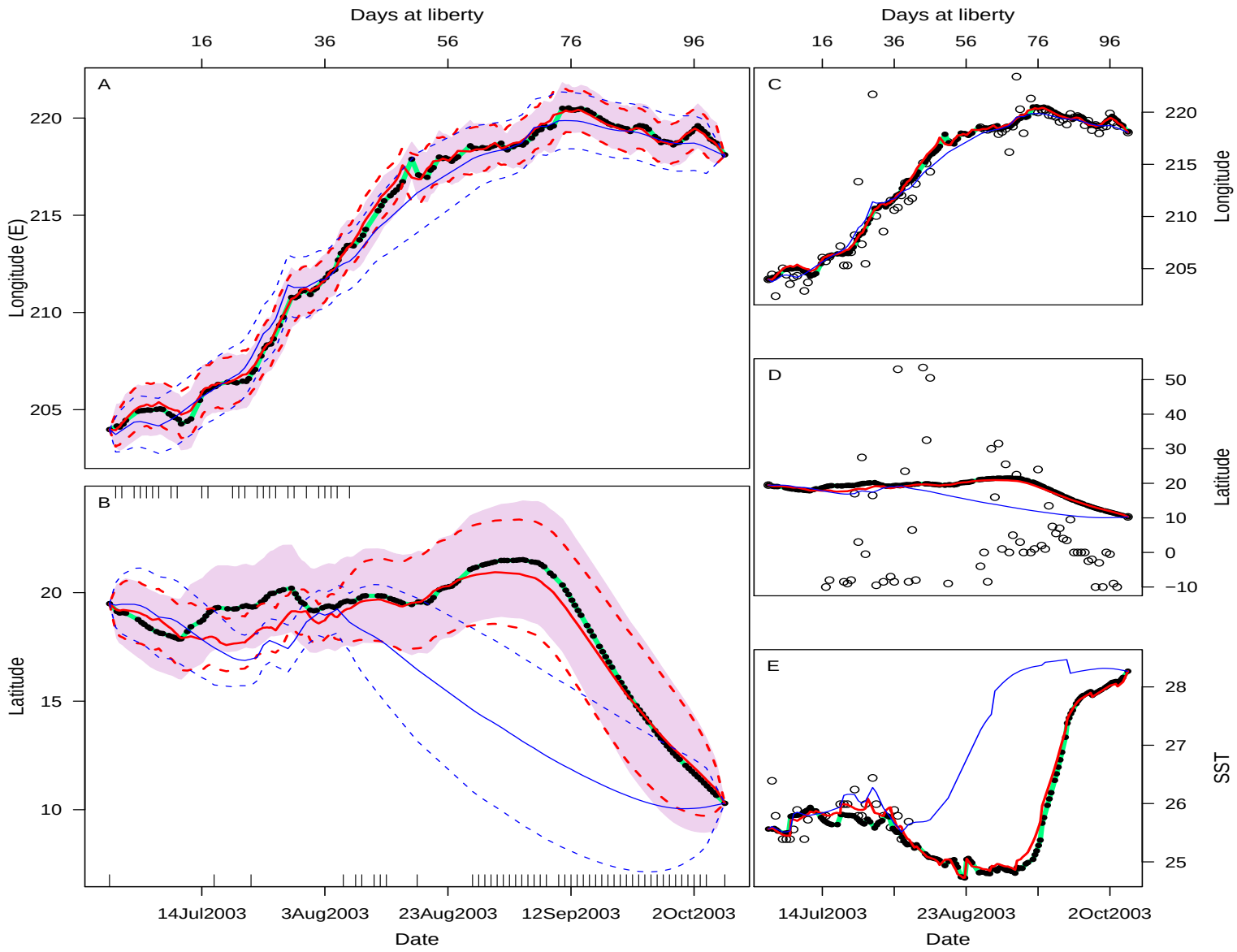
```
year month day hour min sec sst
2002 9 15 3 0 0 22.4
2002 9 17 2 0 0 23.1
2002 9 21 5 0 0 22.9
.
.
.
```

```
library(trackit)
track<-read.table('datafile.dat', header=TRUE)
sst<-read.table('sstdatafile.dat', header=TRUE)
sst.path<-get.sst.from.server(track,150,250,0,40)
prep.track<-prepit(track, sst=sst,
    fix.first=c(198.55,22.85,2002,9,10,0,0,0),
    fix.last=c(200.13,21.95,2003,5,21,0,0,0))
fit<-trackit(prep.track)
plot(fit)
fitmap(fit)
```


Be careful about trusting raw geolocations



Be careful about trusting raw geolocations - 2



Be careful about convergence

```
#R-KFtrack fit
#Mon Nov 12 16:22:14 2007
#Number of observations: 76
#Negative log likelihood: 322.056
#The convergence criteria was met
```

- Convergence should be obtained
- Ways to help the optimizer if a track is problematic
  - Simplify model (especially for short tracks)
    (e.g. `fit<-kftrack(track, bx.a=FALSE, by.a=FALSE)`)
  - Supply better initial values
    (e.g. `fit<-kftrack(track, D.i=500)`)
  - Remove extreme outliers
    (e.g. `track<-track[abs(track$lat)<90,]`)
  - A combination
  - Also check data

Be careful about selecting satellite SST data

- In open ocean coarse resolution is fine
- Near the coast a fine resolution is needed
- In areas with frequent cloud cover consider the blended source
- See the options in the documentation

```
?get.sst.from.server
```

```
?get.blended.sst
```

Remember to report back

- Like to hear when it is working
- Need to hear when it is not

Future directions

- Grid based methods interesting
  - allows other distributions than Gaussian
  - allows strong non-linearities (land areas)
  - very computational demanding
- Numerous tracks in one model
  - The right thing to do if some (or all) parameters are common
  - More confidence in estimated parameters
  - Possible to allow more flexible movement patterns
- Conventional tracks and archival tags in one model
  - First step in using all tagging data in fish stock assessment models
- All packages can be downloaded from:

http://www.soest.hawaii.edu/tag-data/software/

Thank you for listening!

Combining individual and population based models

An appealing idea

- The parameters are the same (drift and diffusion)
- All tagged fish from the same population should be equal representatives, no matter what type of tag
- Might get better individual tracks when parameter estimates get better
- How much more is learned from an (expensive) archival tag

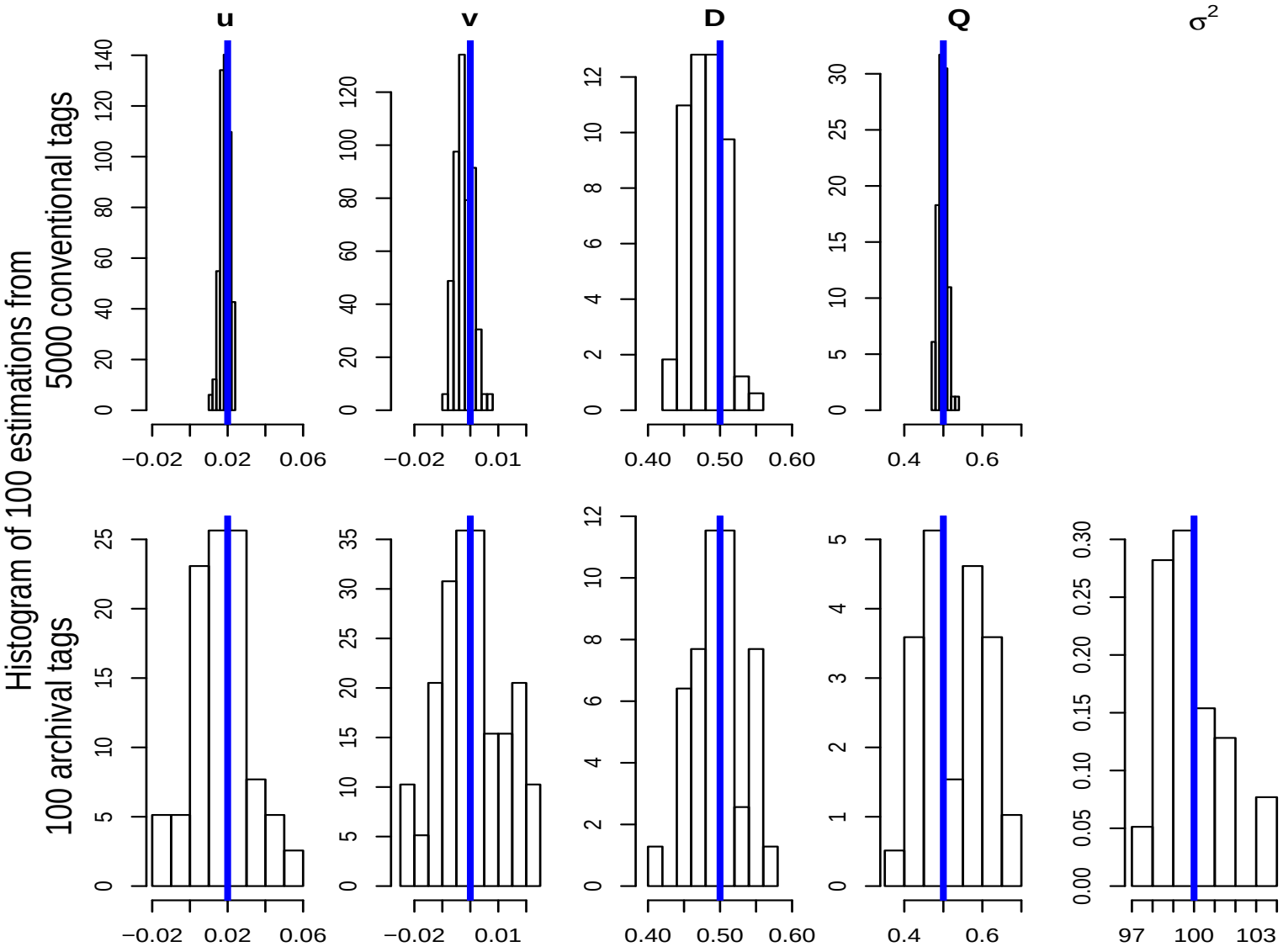
Simulation study

- 100 data sets are simulated each with 5100 simulated individuals
- 5000 with conventional tags 100 with archival tags (randomly assigned)
- Realistic effort pattern, fishing mortality, and natural mortality are applied

Parameter estimation

- A–D–R model is used for the conventional tags
- The Kalman filter likelihood was extended to include survival and recapture probabilities, and the individuals that were not recaptured

Results


Histogram of 100 estimations from
5000 conventional tags
100 archival tags
u
v
D
Q
σ^2

Parameter	True	Conventional tags		Archival tags		All tags	
name	value	bias	std. dev.	bias	std. dev.	bias	std. dev.
u	0.02	-0.00143	0.00259	-0.00364	0.01467	-0.00188	0.00182
v	0.00	-0.00232	0.00302	0.00059	0.01150	-0.00134	0.00299
D	0.50	-0.02078	0.02403	-0.00244	0.03733	-0.01695	0.01640
Q	0.50	-0.00228	0.01221	0.02835	0.08274	-0.00011	0.01164
σ^2	100.00			-0.13608	1.52063	0.09419	1.34204

- Best results from combined model
- In this setting we get almost the same amount of information about drift $(u, v)'$ from one conventional tag as from one archival tag. This will likely change in a more complex setting.
- Archival tags provide more information per tag about diffusion D than conventional tags