

Analysis of wallaby and kangaroo line transect data

David L. Borchers and Martin J. Cox

December 22, 2014

1 Summary

1. Data were collected by an observer walking line transects with a 400 m line spacing.
2. Line transects were orientated east-west and north-south through the survey area.
3. Two species were present in the survey area: wallaby and kangaroo.
4. During the north-south transects the observer left the transect to confirm group size.
5. Data were split by species and transect direction.
6. So far data have been analysed for wallaby and kangaroo for the east-west transects using the 'h1' hazard function form and uniform and half-normal perpendicular density gradients (π_x , Tables 1 & 2). The normal form for π_x is a poor fit, resulting in either problems inverting the hessian (wallaby data set) or massive variance for the parameter estimates (kangaroo data set).
7. I have had no luck fitting the 'h2' hazard function, initial values seem to be the problem

2 Overview

The wallaby and kangaroo sighting line transect data were provided by Colin Southwell in a Microsoft access database. Here, we use data from Wallaby creek although other data sets are available. The Access database files were converted to xlsx files. The two species analysed are the little wallaby (denoted as **RNW** in the files) and 'big' kangaroos (denoted as **EGK** in the files). The wallabys are small and like dense habitat, whereas the kangaroos are comfortable in the open.

2.1 User defined variables

```
> w=200 #m (half inter-transect spacing)
> ystart=400 #m
```

We start by loading the R workspace with the model fit objects because I can't yet work out how to make Sweave cache:

```
> load('~Dropbox/packages/2D distance sampling with time/data/landSurveys/workspace.RData')
```

Next we read in sightings and transect data. To do so, we will need the `xlsx` package.

```
> library(xlsx)
> transects=read.xlsx('~Dropbox/packages/2D distance sampling with time/data/landSurveys/W
> sightings=read.xlsx('~Dropbox/packages/2D distance sampling with time/data/landSurveys/W
```

We make use of the `psych:scatterplot` function to display the sightings data.

```
> require(psych)
```

Source the package R code:

```
> source("~/dropbox/packages/2D distance sampling with time/R/2DLTfunctions.r")
```

The line transects were 400 m. Data were collected by observers on foot in the pre-GPS era, so no grid coordinates are available for the transects. **I have a note from my discussions with Colin that the observer left the N-S and S-N direction transects, so we'll concentrate our analyses on the E-W and W-E transects.** Remove the N-S and S-N transects.

```
> nrow(transects)
```

```
[1] 78
```

```
> transects=subset(transects,TDRN %in% c('EW','WE'))
> nrow(transects)
```

```
[1] 47
```

Examine transect data:

```
> summary(transects)
```

	SIDT	TIDT	TNNU	TBRG	TDRN
Min.	:4	Min. :12.00	Min. : 7.00	Min. : 90.0	EW:20
1st Qu.:	:4	1st Qu.:17.00	1st Qu.:26.50	1st Qu.: 90.0	NS: 0
Median	:4	Median :21.00	Median :43.00	Median : 90.0	SN: 0
Mean	:4	Mean :20.47	Mean :42.21	Mean :166.6	WE:27
3rd Qu.:	:4	3rd Qu.:24.50	3rd Qu.:58.50	3rd Qu.:270.0	
Max.	:4	Max. :28.00	Max. :77.00	Max. :270.0	

TLGT	DATE	TDUR	OBID	REPL
Min. :0.5000	Min. :1986-08-23	Min. :17.00	CS:47	Min. :3.000
1st Qu.:0.8250	1st Qu.:1986-08-28	1st Qu.:26.50		1st Qu.:3.000
Median :1.0000	Median :1986-08-30	Median :34.00		Median :4.000
Mean :0.9649	Mean :1986-08-28	Mean :33.15		Mean :3.574
3rd Qu.:1.1000	3rd Qu.:1986-08-31	3rd Qu.:39.00		3rd Qu.:4.000
Max. :1.2000	Max. :1986-09-02	Max. :60.00		Max. :4.000

TDAY
Min. :1.000
1st Qu.:2.000
Median :2.000
Mean :2.043
3rd Qu.:3.000
Max. :3.000

SIDT site ID.

TNNU transect number

TBRG transect bearing, deg (0 = north; 90 = east)

TDRN transect direction, grid (e.g. east to west = EW)

TLGT transect length, km

TDUR

We need the transects to be observed at a constant speed so checking transect speed:

and the direction of transects:

```
> table(transects$TBRG)
```

```
90 270
27 20
```

We will now subset the observation data to include only EW and WE transects:

```
> nrow(sightings)
```

```
[1] 1124
```

```
> sub.sight=subset(sightings,sightings$TNNU %in% unique(transects$TNNU))
```

```
> nrow(sub.sight)
```

```
[1] 629
```

merge the sightings and transect data:

```
> sub.sight=merge(sub.sight,transects,'TNNU')
```

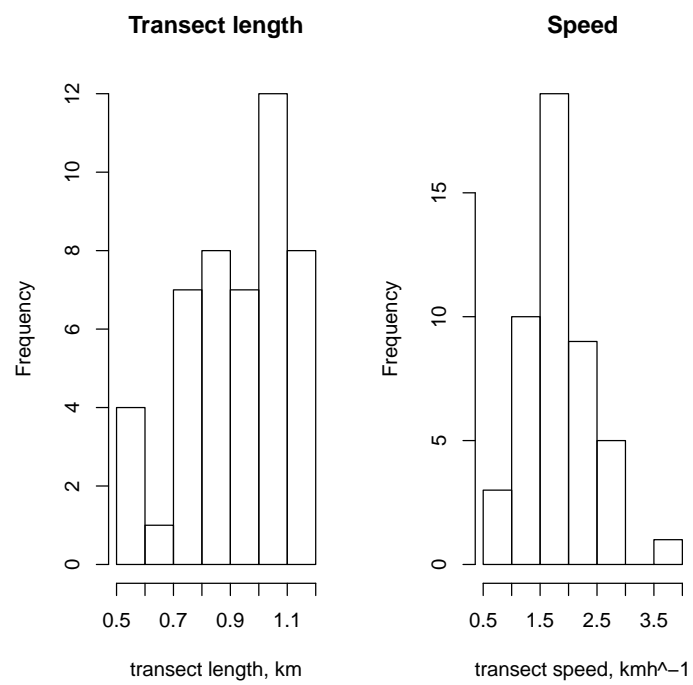


Figure 1: Transect lengths and survey speeds.

2.2 Sightings data

> summary(sub.sight)

TNNU	SIDT.x	STNU	HOUR	SPEC
Min. : 7.00	Min. :4	Min. : 1.000	Min. : 6.00	EGK:402
1st Qu.:29.00	1st Qu.:4	1st Qu.: 2.000	1st Qu.: 8.00	RNW:227
Median :44.00	Median :4	Median : 4.000	Median :12.00	
Mean :42.63	Mean :4	Mean : 4.095	Mean :11.94	
3rd Qu.:57.00	3rd Qu.:4	3rd Qu.: 6.000	3rd Qu.:16.00	
Max. :77.00	Max. :4	Max. :13.000	Max. :17.00	
ANGL	RADL	PERP	GPSZ	
Min. : 0.0	Min. : 8.0	Min. : 0.00	Min. : 1.000	
1st Qu.: 21.0	1st Qu.: 55.0	1st Qu.: 20.00	1st Qu.: 2.000	
Median : 44.0	Median : 88.0	Median : 53.00	Median : 3.000	
Mean : 47.1	Mean :101.7	Mean : 64.69	Mean : 5.124	
3rd Qu.: 68.0	3rd Qu.:140.0	3rd Qu.: 89.00	3rd Qu.: 6.000	
Max. :168.0	Max. :330.0	Max. :249.00	Max. :23.000	
MVST	SBMV	FSAG	TWAW	
Min. :0.00000	Min. :0.0000	Min. :0.000	Min. :0.0000	
1st Qu.:0.00000	1st Qu.:0.0000	1st Qu.:0.000	1st Qu.:0.0000	
Median :0.00000	Median :0.0000	Median :0.000	Median :0.0000	
Mean :0.08903	Mean :0.4817	Mean :0.787	Mean :0.6216	
3rd Qu.:0.00000	3rd Qu.:1.0000	3rd Qu.:1.000	3rd Qu.:0.0000	
Max. :9.00000	Max. :9.0000	Max. :9.000	Max. :9.0000	
HABT	GPRP	SIDT.y	TIDT	TBRG
Mode:logical	Min. : 15.00	Min. :4	Min. :12.00	Min. : 90.0
NA's:629	1st Qu.: 25.00	1st Qu.:4	1st Qu.:17.00	1st Qu.: 90.0
	Median : 55.00	Median :4	Median :19.00	Median : 90.0
	Mean : 66.16	Mean :4	Mean :19.67	Mean :176.7
	3rd Qu.: 95.00	3rd Qu.:4	3rd Qu.:22.00	3rd Qu.:270.0
	Max. :255.00	Max. :4	Max. :28.00	Max. :270.0
TDRN	TLGT	DATE	TDUR	OBID
EW:303	Min. :0.500	Min. :1986-08-23	Min. :17.00	CS:629
NS: 0	1st Qu.:0.900	1st Qu.:1986-08-28	1st Qu.:31.00	
SN: 0	Median :1.100	Median :1986-08-30	Median :35.00	
WE:326	Mean :1.012	Mean :1986-08-28	Mean :37.46	
	3rd Qu.:1.100	3rd Qu.:1986-08-31	3rd Qu.:44.00	
	Max. :1.200	Max. :1986-09-02	Max. :60.00	
REPL	TDAY			
Min. :3.000	Min. :1.000			
1st Qu.:3.000	1st Qu.:1.000			
Median :4.000	Median :2.000			
Mean :3.518	Mean :2.052			
3rd Qu.:4.000	3rd Qu.:3.000			
Max. :4.000	Max. :3.000			

ANGL Sighting angle 0 = along transect; 90 - abeam
RADL Range finder distance
PEPR Perpendicular distance calculated from ANGL and RADL
GPSZ Group size.
MVST Movement: 0 = still; 1 = movement after detected.
SBMV Subsequent movement, after detection
TWAW 0 = away, 1 = towards **this is wrong**

Add a y-coordinate to each sighting:

```
> sub.sight$Y=sqrt(sub.sight$RADL**2-sub.sight$PERP**2)
```

Subset the sightings data based on truncation distance, w , and maximum y-coordinate range `ystart`:

```
> nrow(sub.sight)
```

```
[1] 629
```

```
> sub.sight=subset(sub.sight, PERP < w & Y < ystart)
> nrow(sub.sight)
```

```
[1] 612
```

Now we split the data by species:

```
> RNWdat=subset(sub.sight,SPEC=='RNW')
> nrow(RNWdat)
```

```
[1] 223
```

```
> EGKdat=subset(sub.sight,SPEC=='EGK')
> nrow(EGKdat)
```

```
[1] 389
```

Display the sightings

3 Analysis

In the analysis section we will attempt to fit three models to each species, each model will have a different perpendicular density distribution: (i) uniform; (ii) half-normal and (iii) normal. We will use the 'h2' hazard rate form **I have had no luck fitting the 'h1' hazard rate function, initial values seem to be the problem.**

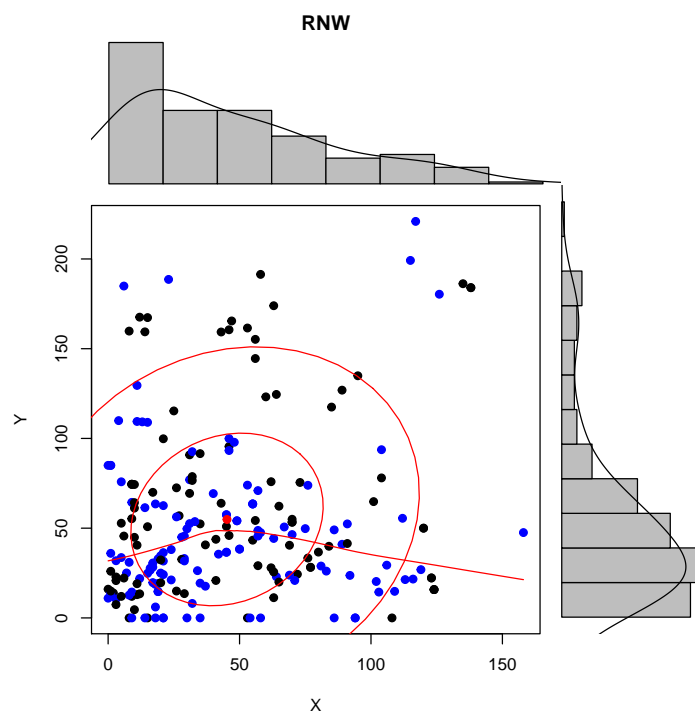


Figure 2: Observations for wallabys (RNW).

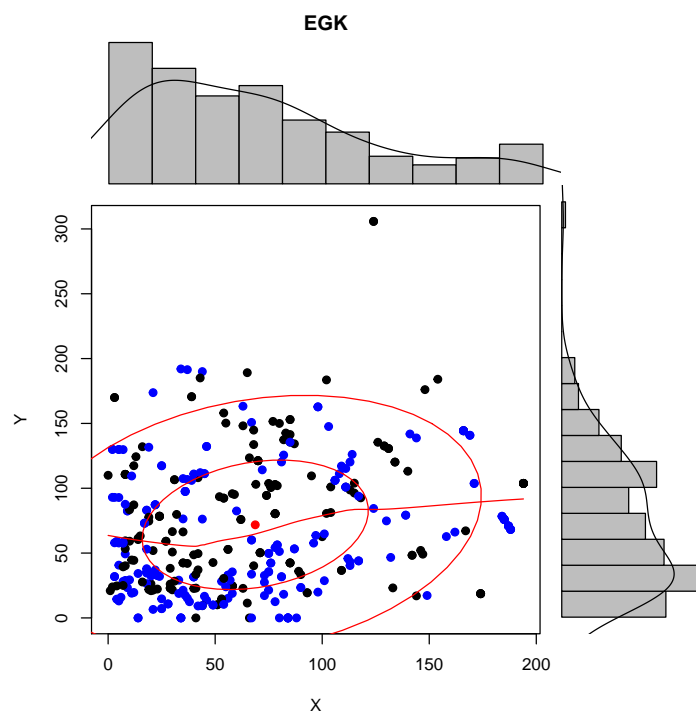


Figure 3: Observations for kangaroos (EGK).

3.1 Analysis of the EW and WE wallaby (RNW) sightings

```

> #fit uniform perpendicular density function:
> modRNWunif=fityx(y=RNWdat$Y,x=RNWdat$PERP,b=log(c(1,1)),hr=h1,ystart=ystart,
+                 pi.x=pi.const,logphi=NULL,w=w,hessian=TRUE)
> #fit half-normal perpendicular density function:
> modRNWhn=fityx(y=RNWdat$Y,x=RNWdat$PERP,b=log(c(1,1)),hr=h1,ystart=ystart,
+                 pi.x=pi.hnorm,logphi=150,w=w,hessian=TRUE)
> #normal perpendicular density function:
> modRNWn=fityx(y=RNWdat$Y,x=RNWdat$PERP,b=log(c(1,1)),hr=h1,ystart=ystart,
+                 pi.x=pi.norm,logphi=c(0.1,10),w=w,hessian=TRUE)

```

Table 1: Wallaby (RNW) data set parameter estimates for the **h1** form of the hazard function and uniform and half-normal perpendicular density distribution.

Density distribution	Hazard rate h1 parameters (CV) \hat{b}	π_x parameters (CV) $\hat{\phi}$	dAIC
Uniform	5.02 (0.05); 0.92 (0.02)	-	4.1
Half-normal	4.75 (0.06); 0.89 (0.03)	163.19 (0.23)	0

3.2 Analysis of the EW and WE wallaby (EGK) sightings

```

> #fit uniform perpendicular density function:
> modEGKunif=fityx(y=EGKdat$Y,x=EGKdat$PERP,b=log(c(1,1)),hr=h1,ystart=ystart,
+                 pi.x=pi.const,logphi=NULL,w=w,hessian=TRUE)
> #fit half-normal perpendicular density function:
> modEGKhnn=fityx(y=EGKdat$Y,x=EGKdat$PERP,b=log(c(1,1)),hr=h1,ystart=ystart,
+                 pi.x=pi.hnorm,logphi=150,w=w,hessian=TRUE)
> #normal perpendicular density function:
> modEGKn=fityx(y=EGKdat$Y,x=EGKdat$PERP,b=log(c(1,1)),hr=h1,ystart=ystart,
+               pi.x=pi.norm,logphi=c(0.1,10),w=w,hessian=TRUE)

```

Table 2: Kangaroo (EGK) data set parameter estimates for the **h1** form of the hazard function and uniform and half-normal perpendicular density distribution.

Density distribution	Hazard rate h1 parameters (CV) \hat{b}	π_x parameters (CV) $\hat{\phi}$	dAIC
Uniform	4.68 (0.05); 0.83 (0.03)	-	0
Half-normal	-51.08 (NA); 0.24 (NA)	427.34 (NA)	242
Normal	4.68 (NA); 0.83 (NA)	2.58 (NA); 12.51 (NA)	2

4 Next steps

1. Find out why π_x normal distribution is failing.
2. Find out why hazard function form `h2` isn't working.
3. Truncate the EGK data set at $y=200$ (see Fig. 3).
4. Look at speeds for North-South transects. Perhaps these are more consistent than East-West transects.