ANNEX 1: Field data forms and field form instructions

1. **MIKE SITE FORM: FOREST ELEPHANT POPULATION**

2. **MIKE FORM INSTRUCTIONS: ELEPHANT POPULATION FORM**

1. MIKE SITE FORM
FOREST ELEPHANT POPULATION

1. **Name of site:**
2. **Monitoring trip identification number:**
3. **Monitoring team leader:**
4. **Start date of monitoring trip:**
5. **End date of monitoring trip:**
6. **Point of departure of monitoring trip:**
7. **Point of arrival of monitoring trip:**
8. **Number of effective monitoring days:**
9. **Number of men in monitoring team:**

Type	**Number**
Monitoring team	
Porters	
Total	

10. **Sector monitored:**
11. **Draw route of the monitoring trip and mark recces and transects on attached map or satellite image (give reference of used map of image):**
12. **Total length of recces:**
13. **Total length of transects:________**
14. **Total distance (calculated/estimated): ____/______**
15. **Survey coverage: inventory codes and lengths**

Inventory Type	Code 1	Km 1	Code 2	Km 2	Code 3	Km 3	Code 4	Km 4	Code 5	Km 5	Code 6	Km 6
Recce deplacement												
Recce												
Transect												
Other												

16. Team activities and habitat:

DATES:						MIKE SITE AND SECTOR:									
Date	**Time**	**Recce or Transect ID**	**GPS Coordinates**	**Grid**	**Altitude**	General Habitat						**Specific Habitat Type Landscape feature**	**Topo-graphy**	**Team**	**Notes**
						TYP	**SUB**	**CAN**	**US1**	**US2**	**VIS**				

17. Recce observations:

RECCE ID:	RECCE TYPE:	LENGT H:	**DATES:**	**MIKE SITE AND SECTOR:**
INITIAL COMPASS BEARING:			**OBSERVERS:**	
GPS start: *Long.* *Lat.*			*GPS end:* *Long.* *Lat.*	

Date	Time	Distance from start	GPS coordinates (optional)	Animal Species	Type of Sign	Age of Sign	Number of animals or sign	Human Sign				Notes
								IND	TYP	AGE	QTY	

18. Transect observations:

TRANSECT ID:	TRANSECT TYPE:	LENGTH:	**DATES:**	**MIKE SITE AND SECTOR:**
COMPASS BEARING:			**OBSERVERS:**	
GPS start: *Long.* *Lat.*			*GPS end:* *Long.* *(Optional)* *Lat.*	

Date	**Time**	**Distance from start**	**Perp. Dist.**	**Animal Species**	**Type of Sign**	**Age of Sign**	**Number of animals or sign**	**Human Sign**				**Notes**
								IND	**TYP**	**AGE**	**QTY**	

2. MIKE FORM INSTRUCTIONS
ELEPHANT POPULATION FORM

FORM INSTRUCTIONS

1. **Name of site:** fill in the full official MIKE site name.

2. **Monitoring trip identification number:** it is good management practice to give each monitoring trip a number or code before it is sent out. This may include the year, a sector code and a number. For example, the fifth monitoring trip in Lope during 1995 might be called LOPE/1995/05 (see below).

3. **Monitoring team leader:** give name and rank of person in charge of the team during its time in the field.

4. **Start date of monitoring trip:** record the first day of the monitoring trip giving the day in numbers, the month in letters and the year in numbers, for example 29 February 1998.

5. **End date of monitoring trip:** record the last day of the monitoring trip in format day, month, year, for example 1 April 1998.

6. **Point of departure of monitoring trip:** enter the name of the point of departure of the monitoring trip and GPS coordinates.

7. **Point of arrival of monitoring trip:** enter the name of the point of arrival of the monitoring trip and GPS coordinates.

8. **Number of effective monitoring days:** the number of days men actually spent monitoring. Placement days, that is days spent in transit to and from their start and end points, should not be included as effective monitoring days. For example a 12 day trip in which men were carried to the starting point of their trip for one day in a 4WD vehicle would be counted as a tip of 10 effective monitoring days.

9. **Number of men in the monitoring team:** give the number of active monitoring team members and number of porters

10. **Sector monitored:** give a brief description of the route of the monitoring trip, including the areas or sectors of the MIKE site visited.

11. **Draw route of the monitoring trip and mark recces and transects on attached map:** each trip should be provided with a photocopied map of the MIKE site on which to record their route. This map should be appended to the Report. Maps should be simplified, for example without contours although relief bands might be shown by different shading. Drainage systems are very useful features to include, as many untrained scouts can navigate using them. All recces and transects should

be shown on the map with their respective ID number.
If a satellite image is used, note the image type, path/row numbers and the year of the image on the form and draw the recces and transects on the image (normally this should have been done already).

12. **Total length of recces:** from the map of the monitoring route estimate how many kilometres were covered on the ground by recces

13. **Total length of transects:** from the map of the monitoring route estimate how many kilometers were covered by transects

14. **Total distance covered (calculated/estimated):** this is the total distance of the whole monitoring trip, including recces, transects and travel recces between recce/transects.
Give both calculated and estimated distance. For example: 15 km (calculated) + 26 km (estimated) = 41 km (estimated).

15. **Survey coverage: inventory codes and lengths:** provide an overview of all sampling units (recces and transects) in this table. Give code for each recce, recce-transect or transect with its respective length in km.

16. **Team activities and habitat:**

This form keeps track of team activities (transects and recces, rest, camp, etc.) and habitat along recces and transects. Enter dates and mike site and sector(s) at the top of the form.
For each entry, note date, time, recce or transect identification number (see below), GPS coordinates, grid cell if GPS coordinates are not recorded and altitude (with altimeter only in mountainous terrain).

Identification number: for each new monitoring trip, recce or transect give a unique identification code. The code contains the following elements:

<u>Monitoring trip:</u>
[Name of site/year/month (start month of monitoring trip)]
For example: *Lope/2000/6/M2 = second monitoring trip in June 2000 in Lope*

<u>Travel Recce or simple recce</u>
[Name of site/year/month (start month of monitoring trip)/RD number]
For example: *Lope/2000/6/M2/RD1 = first travel recce on the second monitoring trip in June 2000 in Lope*

<u>Recce – transect recce</u>
[Name of site/year/month (start month of monitoring trip)/RT number/Recce Number]
For example: *Lope/2000/6/M2/RT3/R2 = second recce segment of the third recce transect unit on the second monitoring trip in June 2000 in Lope*

Recce – transect transect
[Name of site/year/month (start month of monitoring trip)/RT number/Transect Number]
For example: *Lope/2000/6/M2/RT3/T4 = fourth transect segment of the third recce transect unit on the second monitoring trip in June 2000 in Lope*

The following team activities are recorded in the team column: start and end of a recce-transect unit (RT), start and end of a travel recce (RD), overnight camp (CA), rest during the day (RE).

Vegetation is recorded every 200 meters on a recce-transect and every 500 meters on a travel recce . Vegetation types are recorded in a hierarchical way to allow comparison across sites.

TYP: major vegetation type: forest (FOR), savannah (SAV) and agriculture (AGR).
SUB: substrate: dry (TFE), seasonally inundated (TIN), inundated (TIE).
CAN: canopy cover (including tree cover in open savannah): no or very few trees (0-1%), very open (1-15%), open (15-75%), closed (75-100%).
US1: the most abundant or dominant vegetation type in the understorey: herbs or grasses (HER), small trees and bushes (ARB), Lianas (LIA), Ferns (FOU), Palms (PAL), Bamboo (BAM), Other (AUT). Specify other vegetation types in the notes column.
US2: the second most abundant or dominant vegetation type in the understorey
VIS: visibility in the understorey at the height of an adult (1.6 – 1.8 meters): less than 5 meters (TDE), 5-10 meters (DEN), more than 10 meters (OUV).

Specific habitat type / Landscape feature: if you can give a more specific description of the habitat type, note this in the specific habitat type column. This will depend on the site you work in and different classifications exist in different places. Table 1 gives some examples of types that are being used. Try to stick to these, but you can use other types if yours doesn't figure in the list. In that case give a short description of that particular habitat on the back of the form.
Note other landscape features such as streams, caves, big rocks etc. in the notes column (table 1)

Topography: describe the general topography: plain (PLA), Valley bottom (FVA), steep slope (PRA), light slope (PLE), ridge (CRE), hill or mountain top (SOM).

Notes: give additional notes on team activities and vegetation.
• For vegetation characterized by (a) certain species or genus note the characteristic tree species (e.g. Gilbertiodendron, Aucoumea etc.) and understorey species (e.g. marantaceae).
• For savannah note age of burn: very recent (TRE, < 3 months), recent (REC, 3-12 months), old (ANC, > 12 months), not burned (NON), unknown (INC).
• For logged forest note whether logging is ongoing or past and give date or period of the last cut if known.

• For landscape features such as clearings in a forest or salt licks note if recently visited by animals.
• For cash crops specify species.

Table 1: *hierarchical classification of habitat and vegetation*

Principal vegetation types and habitat structure					**Specific vegetation type / Landscape element**	**Notes**
TYP	**SUB**	**CAN**	**US1 US2**	**VIS**		
Forest (FOR) Savannah (SAV) Agriculture (AGR)	Dry (TFE) Seasonally inundated (TIN) Inundated (TIE)	<1% <15% 15-75% >75%	herbs or grasses (HER) small trees and bushes (ARB) Lianas (LIA) Ferns (FOU) Palms (PAL) Bamboo (BAM) Other (AUT)	<5m (TDE) 5-10m (DEN) >10m (OUV)	**Natural habitat** Dense humid evergreen forest Dense humid deciduous forest Marantaceae forest Dense dry forest Open forest Mountain forest Secondary forest High altitude bamboo Low altitude bamboo Swamp or inondated forest Raphia forest Gallery forest Bosquet (isolated tree patch in savannah) Grass savannah Shrub savannah Tree savannah Tree or bush savannah Woodland Montane grassland Herbaceous swamp Papyrus swamp Mangrove Floodplain Other (specify in notes) <u>Agriculture</u> Field Fallow/forest mosaic Burned field Commercial or industrial plantation Tree plantation Other (specify in notes)	<u>Dominant tree species</u> <u>Dominant species in understorey</u> <u>Logging</u> active age of last logging <u>Savannah burns</u> very recent (TRE, < 3 months) recent (REC, 3-12 months) old (ANC, > 12 months) not burned (NON) unknown (INC)

					Landscape features Forest clearing with salt lick Forest clearing without salt lick Salt lick in savannah Salt lick in forest Waterfall Big rock Escarpment Stream River Lake Swamp	

17. Recce form:

Give the recce an unique identification number (see higher). Give recce type (travel recce, recce-transect recce,…) and length.
Give dates, MIKE site and sector and the names of the observers. Underline the name of the person who takes notes.
Record a GPS location at the beginning and end of each recce and note the initial compass bearing.
For each observation note the following:

- Date
- Time of observation
- Distance from the start of the recce as measured with hip-chain thread
- Animal species
- Type of sign:
  - Seen (OBS)
  - Heard (ENT)
  - Dung (CRO)
  - Nest (NID)
  - Spoor (EMP)
  - Wallow (BAU)
  - Carcass, animal killed by predator (CAR PRE)
  - Carcass, animal killed by poacher/hunter (CAR BRA)
  - Carcass, cause of death unknown (CAR INC)
  - Other (AUT): specify in notes
- Age of sign: fresh (FRA), recent (REC), old (VIE), very old (TVI)
  For elephant dung age classification see table 2
- Number of animals or sign

Table 2: *Dung age classes*

Fresh	**FRA**	sometimes still warm!, with shiny fatty acid sheen glistening on exterior and strong smell
Recent	**REC**	odour present (break the boli), there may be flies, but the fatty acid sheen has disappeared
Old	**VIE**	overall form still present although boli may be partly or completely broken down into an amorphous mass, no odour
Very Old	**TVI**	flattened, dispersed, tending to disappear

If you come across an elephant boulevard, note the orientation of the boulevard in the notes column (in degrees from the N). Only major boulevards should be recorded. Make a note if you follow a boulevard as part of the recce. You can also record other sign of elephant in the notes column. This includes: feeding sign, rubbing trees, scars on trees due to bark feeding, wallows etc. This information should only be recorded if there is very little or no other recent sign of elephants in the area.

Describe human presence and activity in the human sign column. Note the following:

- IND = indicator of human sign (refer to table 3).
- TYP = type of human sign (refer to table 3).
- AGE = age of human sign (refer to table 3)
- QTY = quantity of human sign (refer to table 3)

Try to give at least the indicator of human sign and if you have more specific information you can give type also (e.g. if you heard gun shots but you don't know what calibre they are, only fill in the indicator column (gunshots). If you do know the calibre you can fill in type as well). For some types it will always be possible to specify (e.g. car, human tracks etc.)).

Add further comments in the notes column. Here you can also give the activity the sign is associated with (hunting) if known.

Table 3: Human sign

INDICATOR	TYPE	AGE	QUANTITY
People (HOM)	- Armed people (HOM ARM) - Non-armed people (HOM NAR) - People heard (HOM ENT)		- Number of people

Human passage (PAS)	- Human tracks (TRA HUM) - Cut or broken branch (BRA CAS) - Major foot path (SEN GRA) - Small foot path - (SEN PET) - Vehicle track (TRA VEH) - Motorcycle, bicycle track (TRA MOT) - Dirt road car (PIS VEH) - Dirt road motorcycle (PIS MOT) - Tarmac road (ROU GOU) - Foot bridge (PNT PIE) - Car bridge (PNT VEH) - River crossing foot (GUE PIE) - River crossing car (GUE VEH) - River access (ACC RIV) - Car or other vehicle (VEH) - Boat (BAT) - Motorised canoe (PIR MOT) - Non-motorised canoe (PIR RAM) - Other (AUT): specify in notes	- Recent (R) - Old (V)	- Number of vehicles, boats...
Permanent settlements (CAM PER)	- Big game hunting camp (CAM GCH) - Small game hunting camp (CAM PCH) - Hunting camp undetermined (CAM ACH) - Camp for capturing animals (CAM CAP) - Fishing camp (CAM PEC) - Mining camp (CAM MIN) - Agricultural dwelling (CAM AGR) - Camp undetermined (IND)	- Occupied (A) - Recently abandoned (R) - Old, abandoned (V) - Unknown (I)	- Number of settlements
Temporary settlements (CAM TEM)	- Big game hunting camp (CAM GCH) - Small game hunting camp (CAM PCH) - Hunting camp undetermined (CAM ACH) - Camp for capturing animals (CAM CAP) - Fishing camp (CAM PEC) - Mining camp (CAM MIN) - Agricultural dwelling (CAM AGR) - Camp undetermined (IND)	- Occupied (A) - Recently abandoned (R) - Old, abandoned (V) - Unknown (I)	- Number of settlements
Gunshots (FUS)	- Automatic (FUS AUT) - Heavy calibre (FUS LOU) - Light calibre (FUS LEG) - Unknown (FUS IND)		- One - Some (2-5) - Many (>5)
Ammunition (MUN)	- Automatic (MUN AUT) - Heavy calibre (MUN LOU) - Light calibre (MUN LEG) - Unknown (MUN IND)		- One - Some (2-5) - Many (>5)
Traps, snares, nets etc. (PIE	- Wire (CAB) - Nylon (NYL) - Pitfall (FOS) - Hunting line (FIL CHA) - Fish trap (PIE POI) - Fish net (FIL POI) - Fish dam (BAR POI) - Other trap (AUT): specify notes	- Active (A) - Old (V)	- Number
Agriculture (AGR)	- Field being prepared (CHA PRE) - Field (CHA)	- Active (A) - Recent (R)	

	- Fallow (JAC) - Tree plantation (ARB) - Commercial or industrial plantation (CUL COM)	- Old (V)	
Wood extraction (BOI)	- Commercial logging (FOR IND) - Prospecting for commercial logging (FOR PRO) - Wood cutting (COU BOI)	- Active (A) - Recent (R) - Old (V)	
Non-timber extraction (CEU)	- Honey gathering (MIE) - Bark stripping (ECO) - Rubber tapping (CAO) - Fruit gathering (FRU) - Other (AUT): specify in notes	- Active (A) - Old (V)	
Mining (MIN)	- Industrial mining (MIN IND): specify notes - Artisanal mining (MIN ART): specify notes	- Active (A) - Recent (R) - Old (V)	
Fire of human origin (FEU)	- Big bush fire (GRA FEU) - Small bush fire (PET FEU)	- Active (A) - Recent (R) - Old (V)	
Other (AUT)	- Specify in notes	- Active (A) - Recent (R) - Old (V)	

18. Transect form

The form should be filled out the same way as the recce form except that for each animal or animal sign the perpendicular distance from the centre line of the transect to the observation should be carefully measured and noted in the appropriate column.

ANNEX 2: Data reporting forms

Reporting forms are presented as spreadsheets, which should eventually link up with a database.
Note that this is not a template for a database itself and that some fields appear twice on different sheets. Also some fields can be calculated from other existing fields.
Data formats, examples, codes and corresponding values are also shown.

SHEET 1: INTRODUCTION

Country	Year	Trimester		Number of surveys	Date first survey	Date last survey	Faunal Code	Faunal Species
[text]	[number]	Code	Value	[number]	[dd/mm/yy]	[dd/mm/yy]	[text]	[text]
		1	1 January - 31 March					
		2	1 April - 30 June					
		3	1 July - 30 September					
		4	1 October - 31 December					

SHEET 2: SURVEYS (overview)

Survey_ID	Site_Name	Team_Leader	Number in Team_Monitoring	Number in Team_Porters	Start_Date	Return_Date	Total Days
[code]	[name]	[nom]	[number]	[number]	[dd/mm/yy]	[dd/mm/yy]	[number]

Next part

Total Length Recces	Total Length Transects	Total Distance Covered	Attached Map	Note
[number]	[number]	[number]	Value	[text]
			yes	
			no	

SHEET 3: Transects, recce-transects and recces (overview)

Recce_Transect_ID	RT_Type		Date_Start	Date_End	Sector 1	Sector 2	Sector 3
EXAMPLES	Code	Value	[dd/mm/yy]	[dd/mm/yy]	[text]	[text]	[text]
Mission_ID/RT01/R01	RTR	Recce Transect Recce	Start/end travel recce				
Mission_ID/RT01/T02	RTT	Recce Transect Transect	Start/end recce				
Mission_ID/RD05	RD	Travel recce	Start/end transect				
Mission_ID/RS04	RS	Simple recce					
Mission_ID/LT07	LT	Line transect					
etc.			Specify				

Next part

GPS_Start_Long	GPS_Start_Lat	GPS_End_Long	GPS_End_Lat	Init_Compass_Bearing	Length of unit	Reporter	Obs1_dung	Obs2_ape	Note
[in d.ddddddd]	[in d.ddddddd]	[in d.ddddddd]	[in d.ddddddd]	[in d.d]	[in km]	[text]	[text]	[text]	[text]

SHEET 4: Team activities

Survey_ID	Recce_Transect_ID	Activity_Type		Date	Time_Start	Time_End	Dist_Topo	GPS_Long	GPS_Lat	Map Grid	Note
[code]	[code]	Code	Value	[dd/mm/yy]	[hh:mm]	[hh:mm]	[in meters]	[in d.ddddddd]	[in d.ddddddd]	[code]	[text]
		RECCEDEP	Start/end travel recce								
		RECCE	Start/end recce								
		TRANS	Start/end transect								
		CAMP	Camp								
		REPOS	Rest								

		AUTRE	Specify									

SHEET 5: Observations

Survey_ID	Recce_Transect_ID	Obs_No	Date	Time	Dist_Topo	Dist_Perp	GPS_Long	GPS_Lat	Grid
[code]	[code]	[number]	[dd/mm/yy]	[hh:mm]	[in meters]	[in cm]	[d.ddddddd]	[d.ddddddd]	[code]

Next part

Species		Observation_Type		Observation_Age		Sign_Qty	Carcass_Ref_ID
Code	Value	Code	Value	Code	Value	[number]	[survey_ID/C1…]
CODE	Identification	OBS	Seen	FRA	Fresh		
		ENT	Heard	REC	Recent		
		SEN	Smelled	VIE	Old		
		EMP	Spoor	TVI	Very old		
		CRO	Dung	INC	Unknown or not reported		
		NID	Nest				
		BAU	Wallow				
		CAR BRA	Carcass, animal killed by poacher/hunter				
		CAR PRE	Carcass, animal killed by predator				
		CAR NAT	Carcass, natural death				
		CAR INC	Carcass, cause of death unknown				
		CAR NRA	Carcass, cause of death not reported				
		AUT	Other (specify in notes)				

Next part

Hum_Sign_Ind		Hum_Sign_Type		Hum_Sign_Age		Hum_Sign_Qty		Elephant poaching Indicator		Note

Code	Value	Code	Value	Code	Value	Code		Code	Value	[text]
HOM	People	HOM ARM	Armed people	A	Active, occupied	One	1	P	Primary indicator	
		HOM NAR	Non-armed people	R	Recent, recently abandoned	Some	2-5	S	Secondary indicator	
		HOM ENT	People heard	V	Old, abandoned	Many	>5	I	Unknown	
PAS	Human passage	TRA HUM	Human tracks	I	Unknown, not determined	Unknown	I			
		BRA CAS	Cut or broken branch							
		SEN GRA	Small foot path							
		SEN PET	Major foot path							
		TRA VEH	Vehicle track							
		PIS VEH	Dirt road car							
		PIS MOT	Dirt road motorcycle							
		ROU GOU	Tarmac road							
		PNT PIE	Foot bridge							
		PNT VEH	Car bridge							
		GUE PIE	River crossing foot							
		GUE VEH	River crossing car							
		ACC RIV	River access							
		VEH	Car or other vehicle							
		BAT	Boat							
		PIR MOT	Motorised canoe							
		PIR RAM	Non-motorised canoe							
		AUT	Other							
CAM PER	Permanent settlement	CAM GCH	Big game hunting camp							
		CAM PCH	Small game hunting camp							
		CAM ACH	Hunting camp undetermined							
		CAM CAP	Camp for capturing animals							
		CAM PEC	Fishing camp							
		CAM MIN	Mining camp							
		CAM AGR	Agricultural dwelling							
		IND	Camp undetermined							
CAM TEM	Temporary settlement	CAM GCH	Big game hunting camp							
		CAM PCH	Small game hunting camp							
		CAM ACH	Hunting camp							

			undetermined							
		CAM CAP	Camp for capturing animals							
		CAM PEC	Fishing camp							
		CAM MIN	Mining camp							
		CAM AGR	Agricultural dwelling							
		IND	Camp undetermined							
FUS	Gunshot	FUS AUT	Automatic							
		FUS LOU	Heavy calibre							
		FUS LEG	Light calibre							
		FUS IND	Unknown							
MUN	Ammunition	MUN AUT	Automatic							
		MUN LOU	Heavy calibre							
		MUN LEG	Light calibre							
		MUN IND	Unknown							
PIEG	Trap	CAB	Wire							
		NYL	Nylon							
		FOS	Pitfall							
		FIL CHA	Hunting line							
		PIE POI	Fish trap							
		FIL POI	Fish net							
		BAR POI	Fish dam							
		AUT	Other trap							
AGR	Agriculture	CHA PRE	Field being prepared							
		CHA	Field							
		JAC	Fallow							
		ARB	Tree plantation							
		CUL COM	Commercial or industrial plantation							
BOI	Wood extraction	FOR IND	Commercial logging							
		FOR PRO	Prospecting for commercial logging							
		COU BOI	Wood cutting							
CEU	Gathering	MIE	Honey gathering							
		ECO	Bark stripping							
		CAO	Rubber tapping							
		FRU	Fruit gathering							
		AUT	Other gathering							
MIN	Mining	MIN IND	Industrial mining							
		MIN ART	Artisanal mining							
FEU	Fire of human origin	GRA FEU	Big bush fire							
		PET FEU	Small bush fire							
AUT	Other (specifier)	AUT	Specify in notes							

SHEET 6: Habitat

Recce_Transect_ID	Obs_No	Survey ID	Recce_Transect_ID	Date	Time	Dist_Topo	Dist_Interval	GPS_Long	GPS_Lat	Grid
[ID]	[number]	[code]	[code]	[dd/mm/yy]	[hh:mm]	[in meters]	[in meters]	[d.ddddddd]	[d.ddddddd]	[code]

Next part

Altitude	Hab_Type		Hab_Substrate		Hab_Canopy		Hab_Understorey1		Hab_Understorey2
[in meters]	Code	Value	Code	Value	Code	Value	Code	Value	
	FOR	Forest	TFE	Terre Ferme	A	No or very few trees (0-1%)	HER	Herbs	
	SAV	Savannah	TIN	Seasonally inundated	B	Very open (1-15%)	ARB	Shrubs or small trees	
	AGR	Agriculture	TIE	Inundated	C	Open (15-75%)	LIA	Lianas	
					D	Closed (75-100%)	FOU	Ferns	
							PAL	Palm trees	
							BAM	Bamboo	
							AUT	Other (describe in notes)	

Next part

Hab_Visibility		Hab_VegType		Topography		Canopy_Species1	Canopy_Species2	Understorey_Spec1	Understorey_Spec2
Code	Value	Code	Value	Code	Value	[text]	[text]	[text]	[text]
TDE	Very dense, < 5m	FOR DEN	Dense humid evergreen forest	PLA	Plain				
DEN	Dense, 5-10 m	FOR SEM	Dense humid deciduous forest	FVA	Valley bottom				
OUV	Open, > 10 m	FOR MAR	Marantaceae forest	PRA	Light slope				

		FOR SEC	Dense dry forest	PLE	Steep slope				
		FOR CLA	Open forest	CRE	Ridge				
		FOR MON	Mountain forest	SOM	Hill or mountain top				
		FOR SEC	Secondary forest						
		FOR BAA	High altitude bamboo						
		FOR BAS	Low altitude bamboo						
		FOR INO	Swamp or inondated forest						
		FOR RAP	Raphia forest						
		FOR GAL	Gallery forest						
		FOR INC	Isolated forest in savannah						
		SAV BOS	Bosquet (isolated tree patch in savannah)						
		SAV HER	Grass savannah						
		SAV ARS	Shrub savannah						
		SAV ARB	Tree savannah						
		SAV ABS	Tree or bush savannah						
		SAV BOI	Woodland						
		PRA MON	Montane grassland						
		MAR HER	Herbaceous swamp						
		MAR PAP	Papyrus swamp						
		FOR MAN	Mangrove						
		PLA ALL	Floodplain						
		CUL VIV	Field						
		CUL JAC	Fallow/forest mosaic						
		CUL BRU	Burned field						
		CUL COM	Commercial or industrial plantation						
		CUL ARB	Tree plantation						
		AUT	Other (specify in notes)						

Next part

Forestry_Explot_Activity		Forestry_Explot_Age	Burned_Age		Landscape_Element		Note
Code	Value	[number]	Code	Value	Code	Value	[text]
ACT	Active		TRE	Very recent (< 3 months)	CLA SAL	Forest clearing with salt lick	
NAC	Non Active		REC	Recent (3-12 m)	FOR CLA	Forest clearing without salt lick	
			ANC	Old (>12 m)	SAV SAL	Salt lick in savannah	
			NON	Not burned	FOR SAL	Salt lick in forest	
			INC	Unknown	CHU	Waterfall	
					GRO	Cave	
					ROC	Big rock	
					ESC	Escarpment	
					RUI	Stream	
					RIV	River	
					LAC	Lake	
					MAR	Swamp, marsh	
					AUT	Other (describe in notes)	

ANNEX 3: Cybertracker as a data collection tool for MIKE.

The Cybertracker (http://www.cybertracker.co.za/) field computer unit consists of a PalmOS compatible handheld computer connected to a GPS. The user-friendly software for the handheld is designed to take data in the field quickly in a predefined, systematic way. Data can be rapidly transferred to a Windows-based PC using the "hotsync" function of the PalmOS and can be viewed with the Cybertracker Geographic Information System. Data can be exported to databases and Arcview GIS.
The cybertracker database and screen sequence on the Palm computer can be customized for particular projects.

Cybertracker was used during field sampling for the MIKE pilot project in Odzala, Congo. About two years ago Jean Marc Froment, project leader Ecofac (Conservation et Utilisation Rationelle des Ecosystèmes Forestiers d'Afrique Centrale, EU), Congo in collaboration with Louis Liebenberg, director of Cybertracker, introduced the use of this tool for patrol based monitoring in Odzala. The MIKE Pilot Project built on this experience and a test for MIKE surveys in the forest was conducted with the ECOFAC teams.

Advantages

Some of the advantages of cybertracker over writing with pen and paper are summarized below.

- It is up to three times faster to enter data in the field using the Cybertracker interface. We also found that most team members trained in systematic data collection have little problem learning to use this tool.
- Entering data following a fixed sequence with predefined lists of values is less prone to errors than entering data with pen and paper. There is also less potential for errors and confusion afterwards when data is transcribed to spreadsheets on a computer
- Transcription of data to the PC in an appropriate format (e.g. spreadsheets or database files) is much faster. In general very little or no data has to be entered manually although data should always be checked visually after it has been transferred.

Technical Problems

Making backups

Currently with the Palm Pilot III x and III xe it is not possible to backup the data on the handheld device itself. Backups can be made by transferring the data to the PC. It is not feasible to carry a PC laptop in the field so this is not a viable option. However we think that frequent backups are essential because of possible technical failure of the device, the possibility of losing the data while changing batteries, accidents (e.g. dropping the

Palm in the water), etc.
For instance when changing batteries on the Palm Pilot, the data are held in memory for a short period of time during which new batteries should be placed. On at least one occasion the reporter of a team in Odzala dropped the new batteries on the ground during replacement and when he finally found them and placed them in the device, the data was already lost.
We think that the backup problem is the most serious technical issue to be resolved before using Cybertracker on a large scale. MIKE surveys are expensive and it is very demoralizing to lose 10 days of hard work by a technical glitch.

The Handspring Visor handheld (http://www.handspring.com), which uses the same operating system and interface as the Palm Pilot, has a hardware expansion slot on the back of the handheld. Several modules are being developed to expand hardware possibilities. A backup module is already available and permits easy (one touch) backup of data and software.
The module can be inserted and taken out easily (plug and play) so that it can be stored separately from the main handheld (e.g. with a different member of the team). Therefore once this unit has been rigorously tested in the field, we would probably recommend it over the Palm Pilot.

GPS

At present, teams in Odzala use the Palm Pilot externally connected to a GPS via a serial cable. Recently three different GPS modules have been developed for the Handspring Visor (http://www.palmgear.com/hs/) which fit in the expansion slot. The unit is currently being tested in the forest and if successful one could use a single device integrating data capture and GPS functionality instead of two separate devices connected by cables (that sometimes get cut with secateurs or machete!) and vulnerable connections.
It remains to be tested however that an internal GPS module will be as powerful as a standalone handheld GPS under dense forest canopy.
Unless the GPS module shows significant superior performance we would recommend that a connection with an external GPS remains a possible option. This is not yet possible with the Visor. The technology of standalone GPS units develops fast and there is a bigger market for these than for integrated GPS Handspring modules. For instance in Odzala MIKE teams are now using a Garmin XL12 and a Garmin III+ (http://www.garmin.com/products/). The more recent Garmin III+ seems to perform slightly better under the canopy than the XL12 and has also some interesting new features, useful for the MIKE recce transects (such as easy programming of a destination point based on distance from the start point).
Several GPS units can also be connected to an external antenna, which may work better in dense vegetation.
The built in units for the Visor lack the possibility of attaching an external antenna.

With the cybertracker software version 2.1 that MIKE teams were using it was not possible to switch off GPS reading on the handheld in the field. This means that for each observation a GPS position was obligatory recorded. If a GPS position could not

be read (e.g. bad reception in very dense vegetation), the whole observation was discarded. Since this was unacceptable for MIKE we decided early in the process to switch off the GPS functionality altogether and only record data on the handheld without simultaneous GPS recording. When necessary, GPS positions were noted on paper separately and later added to the dataset. However with the current version 2.40 one can override the GPS reading so that you can record data without GPS reading.
For MIKE it is not necessary to record a position for each observation, especially on transects. GPS readings can slow down teams so we actually discourage them on transects except for habitat data. We recommend that GPS positions should be recorded for the beginning and end of each transect, habitat data and important observations on recces.
If GPS readings do not significantly slow down the teams on recces (as they did in dense marantaceae undergrowth) cybertracker and GPS could eventually replace topofil to measure distance along the path. On transects however we would always recommend the use of topofil because it marks the transect line and is much more accurate for measuring distances along the line.
GPS technology evolves fast and it can be foreseen that reception of GPS signals under the canopy will improve significantly in the near future.

Other

Other technical problems with cybertracker include:

- battery life
- breakdowns in hot, dusty or humid conditions (careful handling and storage in plastic bags required)
- fragile cable connections (see above)

Cybertracker software

Some of the planned changes in the cybertracker software that are significant for MIKE are the following:

On the handheld:

- the possibility of using more than 1 database and screen sequence on the same handheld. For Mike this would be useful. You can have for instance one sequence for recce-transects, one for travel recces, one for ancillary geographic and socio-economic data etc.
- regular updates to ensure compatibility with new versions of operating systems (Windows and Palm OS) and possibly new operating systems (Pocket PC?)

On the PC:

- direct connection to the ESRI Arcview Geographic Information System
- direct connection to other databases

Because of the importance of the use of DISTANCE software to analyse data it may be interesting to create a link between DISTANCE and CYBERTRACKER at different levels.

On the handheld, it would for instance be useful to have the possibility of creating histograms of transect data to visually check detection probabilities and provide regular feedback to data collection. We believe that this could significantly improve data collection quality and avoid problems such as spikes close to zero, heaping etc (see main report).
Another possibility would be to have a direct link between the two programs on the PC so that data collected with cybertracker can be directly analysed in distance. This would be interesting for site level analysis, but less so for a central data management and analysis unit where data should be stored in a high-end database management system and analysed by an expert on distance sampling analysis.

Customization for MIKE

MIKE needs customized cybertracker databases and screen sequences.
The following table summarizes several potential applications for the MIKE elephant population monitoring, at least in Central Africa.

Survey type	Features	Habitat	Status
Recce – Transects or Simple Linear Transect	Animal and human sign, vegetation	Forest, possibly some savannahs	ß version available and being tested (MIKE)
Travel recces		Forest	ß version available and being tested (MIKE)
Geographic and socio-economic database	Access routes (roads, rivers, paths), villages, demography, economic activities	Forest Savannah	In development with Ecofac
Vegetation (ground verification of satellite images)	Data on vegetation	Forest Savannah	Not yet developed
Mobile Patrols	Law enforcement data and law enforcement effort	Forest Savannah	In use by Ecofac (developed for Ecofac Odzala) and in several South African parks and projects
Barrier patrols	Law enforcement data and law enforcement effort	Forest Savannah	Not yet developed
Intelligence information	Data on poaching activities and elephants	Forest Savannah	Not yet developed

It is also recommended that procedures be established to transfer data efficiently between the cybertracker PC software and MIKE database management systems.

Use of Cybertracker in the global MIKE project

If the above-mentioned technical issues can be addressed and the tool has been successfully tested, cybertracker could in the long run be used for MIKE in many sites in Africa. However it remains relatively sophisticated technology that can break down rapidly when basic conditions are not met, such as the regular supply of new batteries, replacement units, access to technical support (good communication), training and access to the site by a trouble-shooter in case of technical difficulties.

It would be wise to train teams in the use of both pen and paper together with cybertracker so that they can always drop back to the most basic material. Motivated teams in Garamba and Virunga national parks in the Democratic Republic of Congo (DRC) continued patrolling and law enforcement monitoring using simple methods and basic tools (pen and paper) even during the most difficult periods of the recent war in this country.

In a country like DRC it would be probably be unrealistic and unsustainable to rely on cybertracker at the moment. Battery supply is irregular, and even bringing electronic equipment across the border can be problematic.

It will be necessary to have a trouble-shooter at the disposal of the sites that use cybertracker who can follow-up on data collection and who can be contacted in case of technical difficulties.

ANNEX 4: Analysis of data and survey design for the mike central African pilot project: first interim report

S.T. Buckland and F.M. Underwood

November 2000

In this report, we outline analyses of the recce-transect data, to compare encounter rates between recces and transects, and to estimate density of elephant dung-piles. We do not address here the conversion of estimates of dung-pile density into estimates of elephant density. We use a subset of the pilot survey data to illustrate the analyses.

We also outline the kinds of trend analysis that might be conducted, once sufficient data have accumulated.

1. Comparing recce and transect encounter rates

Wilcoxon's signed ranks test provides a simple way of comparing recce and transect encounter rates. For the subset of the pilot survey data, at each sample location, there was 1km of transect and 4kms of recce. Thus we divided the recce counts by four, to give encounter rates as number of dung-piles per km. Encounter rates were very variable, which is the reason that we favour the nonparametric signed ranks test. The test gave no significant difference in encounter rates for the subset of data analysed (test statistic = 84, p=0.104). However, to model the transect distance data, it was necessary to truncate the larger distances. Because distances are not measured on the recces, comparable truncation cannot be carried out, so that, other things being equal, we would expect the recce encounter rate to exceed the transect encounter rate on average.

We also calculated a ratio estimator as described in the appendix. If there is no difference in recce and transect encounter rates, we expect this estimated ratio to be close to one. Our estimate for the subset of data was 0.87 with standard error 0.20, which indicates that observed encounter rates were about 13% lower for the transects (after truncation) than for the recces, although this difference is not significant.

To calibrate recce counts against transect counts for these data, we would multiply the recce counts by 0.87 (see appendix).

2. Estimating dung-pile density from transect data

Standard line transect analyses can be conducted on the distances from the transect to detected dung-piles, using software Distance. In the simplest case, dung-pile density is estimated by

$$\hat{D} = \frac{n}{2\hat{m}L}$$

where n is the number of dung-piles detected,
$\hat{m}$ is the estimated strip half-width,
and L is the total length of transect line.

For the subset of data, a truncation distance of 2.5m for perpendicular distances was found to be satisfactory, and a half-normal detection function with cosine adjustments was chosen. Analyses were carried out on the pooled data (AIC=2021.8), and on data stratified by observer (AIC=2025.4), habitat (savannah/forest; AIC=2024.5) and both (AIC=2028.1). This indicates that no stratification is needed, and a pooled detection function may be estimated, as AIC is smallest for that case. The effective half-width of the strip was estimated to be 1.33m either side of the line (cv=9.6%; 95% ci (1.10, 1.60)), the estimated encounter rate was 8.5 dung-piles per km (cv=21.1%; 95% ci (5.6, 13.0)) and the dung-pile density was estimated to be 31.9 dung-piles per hectare (cv=23.2%; 95% ci (20.0, 50.9)). However, the fit of the model was poor ($c_2^2 = 9.69$; $p = 0.008$). The reason for the poor fit is that one of the observers recorded a high number of dung-piles as on the transect line; of 95 within 2.5m of the line, 23 (24%) were recorded at zero distance.

If we use data from the better observer only, the effective half-width of the strip is estimated to be 1.29m either side of the line (cv=13.8%; 95% ci (0.98, 1.70)), the estimated encounter rate is 7.6 dung-piles per km (cv=36.3%; 95% ci (3.5, 16.5)) and the dung-pile density is estimated to be 29.4 dung-piles per hectare (cv=38.8%; 95% ci (13.1, 65.6)). The fit of the model is now good ($c_2^2 = 1.59$; $p = 0.45$).

Estimated density is almost the same, whether or not the 'poor' observer is included. This is a consequence of restricting attention to the half-normal detection function, which does not fit the 'spike' in distances at zero. If the hazard-rate model is used, the spike is fitted, and density is estimated to be 34% higher (42.8 dung-piles per hectare). This shows the sensitivity that can occur when a high proportion of distances are recorded as zero; indeed, if we allow cosine adjustments to the hazard-rate fit, estimated density is 430% higher (168.9 dung-piles per hectare), although the fitted detection function is totally implausible!

3. Improving efficiency of dung-pile density estimates using recce counts

We can in principle improve precision in the density estimates from Section 2 by using both recce and transect data to estimate encounter rate. The results of Section 2 show that the coefficient of variation in encounter rate dominates the overall coefficient of variation in the density estimate. By supplementing the transect counts with the recce counts, we therefore expect to obtain appreciably more precise estimates of density. If we are prepared to assume that recce and transect encounter rates are the same, then the improvement in precision will be larger than if we calibrate recce counts using the ratio estimator of Section 1, because the ratio estimator itself has a variance. Section 3 of the appendix shows how a combined encounter rate is estimated, with or without calibration. Expressions for variance are also given there.

Note that the estimated density of Section 2 can be split into two components, corresponding to encounter rate and effective strip width:

$$\hat{D} = \frac{n}{L} \times \frac{1}{2\hat{m}}$$

We can therefore take estimates of encounter rate, using the equations of Section 3 of the Appendix, together with estimates of effective strip width from Section 2. However, there is an important proviso on this. The truncation distance for perpendicular distances should be the same for both the encounter rate and the effective strip width. In our case, we truncated at 2.5m when estimating $\hat{m}$. Because one or two detections were at very large distances, reliable modelling cannot be done without some truncation. The problem this creates is that the recce counts are untruncated. This forces the use of calibrated encounter rate estimates when using recce counts, because the encounter rate for untruncated recces cannot be assumed the same as for truncated transects. This raises the question of whether, in future, recce counts should be of dung-piles within say 2.5m of the line only (although calibration may still be found to be necessary of course).

For location *i*, we have

$$\hat{D}_i = \frac{\hat{e}_{ic}}{2\hat{m}}$$

with variance

$$\hat{\text{var}}(\hat{D}_i) = \hat{D}_i^2 \left([cv(\hat{e}_{ic})]^2 + [cv(\hat{m})]^2 \right)$$

(assuming that effective strip width is the same at all locations). If $\hat{e}_c$ is the overall estimated encounter rate, using transect and recce data, then overall density is estimated as

$$\hat{D} = \frac{\hat{e}_c}{2\hat{m}}$$

with variance

$$\hat{\text{var}}(\hat{D}) = \hat{D}^2 \left([cv(\hat{e}_c)]^2 + [cv(\hat{m})]^2 \right)$$

For the subset of data, and retaining the observations of both observers, the estimated encounter rates and densities at each of the 22 locations are shown in Tables 1 and 2 respectively. These location-specific density estimates allow an assessment of which factors are related to density, for example by plotting the density estimates against variables such as distance from human habitation. We hope to implement a more integrated spatial modelling approach later in the project.

Note from Table 2 that use of recce counts has made little difference on average to the precision of location-specific density estimates. Indeed, precision tends to be worse when calibrated recce counts are used than when recce counts are ignored; even if uncalibrated recce counts are used, precision on the location-specific density estimates shows no sign of improvement over the transect-only estimates. This surprising conclusion seems to arise because the main source of variability is between locations, and the extra effort at each location yields little if any benefit. The comparison may prove more optimistic once we have data from other areas, and spatial models to reduce spatial variation. However, once all the pilot data have been analysed, the decision of whether to use recce/transect combinations, and if so, of the proportion of total effort that is done as recces, should be reviewed.

4. Estimating trend from MIKE data

Simple trend estimates may be obtained from MIKE data in various ways. If we could assume that the effective width of search is constant across years, analyses can be of encounter rates. We consider first how analyses might proceed in that case, then consider the case that the effective width of search varies over time.

4.1 Analysing encounter rate

Suppose first that recce encounter rates are found to not differ significantly from transect encounter rates, or alternatively, the correction to recce encounter rates can be assumed constant across years. Assume further that the proportion of recce to transect remains the same over time. Then a simple trend analysis can be carried out as follows. For any given region of interest, sum the number of dung-piles detected within the region (whether from a recce or a transect) in each year. Denote the number of dung-piles in year j by n_j, $j = 1,\ldots, J$. Denote the corresponding distance covered by L_j. This is the combined distance of recce and transect within the region in year j. Then we can estimate trend within the region by fitting the following generalized additive model (gam):

$$E(n_j) = \exp\left(\log_e L_j + s(j)\right)$$

where $\log_e L_j$ is the *offset* and $s(j)$ is a *smoother* of year j. The error distribution can be assumed to be Poisson, and the option in which a dispersion parameter is estimated should be selected. The link function for the above model is the log. Software S-PLUS or GENSTAT can be used to fit this model. The amount of smoothing carried out is determined by *degrees of freedom*. One degree of freedom corresponds to a generalized linear model with log link, Poisson error distribution, year as a continuous covariate, and offset equal to $\log_e L_j$. If degrees of freedom are set equal to $J-1$ (number of years less one), then this is equivalent to fitting a generalized linear model as above, but with year as a factor at *J* levels. Intermediate values for the degrees of freedom yield intermediate levels of smoothing between these two extremes.

If recce counts must be calibrated using transect counts each year, then n_j from the above analysis should be replaced by the relevant transect count plus the adjusted recce count.

4.2 Analysing density estimates

In practice, effective width of search is likely to vary between years, due for example to changes of observer, or possibly changes in habitat. In this case, we need to analyse trends in density estimates, calculated by the methods of Section 3, or Section 2 if only transect data are to be used.

Dung-pile density in Section 2 is estimated by

$$\hat{D}_j = \frac{n_j}{2\hat{m}_j L_j}$$

where n_j is the number of dung-piles detected in year *j*,

$\hat{m}_j$ is the estimated strip half-width in year *j*,

and L_j is the total length of transect line in year *j*.

Thus we can use the same methods as in Section 4.1 to model the n_j, but with the offset $\log_e L_j$ replaced by $\log_e 2\hat{m}_j L_j$. For the adjusted estimates of Section 3, estimated density can be expressed as

$$\hat{D}'_j = \hat{\iota}_j \hat{D}_j = \frac{n_j \hat{\iota}_j}{2\hat{m}_j L_j}$$

so that the offset becomes $\log_e \frac{2\hat{m}_j L_j}{\hat{\iota}_j}$.

Alternatively, we might simply treat the adjusted density estimates $\hat{D}'_j$ as the responses, and model them using a gam with a log link function and a gamma error distribution.

4.3 Spatio-temporal modelling

The above methods allow trends in time to be estimated for any sampled region of interest. Thus the area might be stratified, for example according to accessibility, and trends estimated for each stratum. Given that spatial modelling of the survey data is planned, a more sophisticated approach is to estimate trends in time as a function of location. In its simplest form, this could be achieved by fitting a spatial model to each year's data in turn, to obtain a density estimate for any location for which a trend estimate is required. The time series of density estimates for a given location could then be analysed using a gam with a log link function and a gamma error distribution.

It may prove better to fit a single spatio-temporal model to all years' data, for example by including year as a continuous covariate in a gam. By including interactions between year and other covariates, trends can be interpreted in relation to factors such as proximity of human habitation, habitat, forestry operations, etc.

Appendix

1. The transect and recce variance estimates $\hat{\mathrm{var}}(n_{it}), \hat{\mathrm{var}}(n_{ir})$

In the following, suffix *t* indicates a quantity associated with transects, and suffix *r* indicates recces. Consider just the line transects which each have m_{it} (=5 for the example data set) replicates, each of length l_{ijt} (=200m for the sample data set). Let n_{ijt} be the number of observations on replicate *j* of the i^{th} transect.

Then the variance of n_{it} for transect *i* can be estimated by

$$\hat{\mathrm{var}}_s(n_{it}) = \frac{L_{it}}{m_{it}-1}\sum_{j=1}^{m_{it}} l_{ijt}\left(\frac{n_{ijt}}{l_{ijt}} - \frac{n_{it}}{L_{it}}\right) \text{ where } L_{it} = \sum_{j=1}^{m_{it}} l_{ijt}$$

More reliable estimates of variance can be obtained by estimating a dispersion parameter b_t by $\hat{b}_t = \frac{\sum_{i=1}^{k}\hat{\mathrm{var}}_s(n_{it})}{\sum_{i=1}^{k} n_{it}}$. We can then take $\hat{\mathrm{var}}(n_{it}) = \hat{b}_t n_{it}$ for each *i*.

Encounter rate is then estimated as $\hat{e}_{it} = \frac{n_{it}}{L_{it}}$ with estimated variance $\frac{\hat{b}_t n_{it}}{L_{it}^2}$. Overall transect encounter rate is estimated as $\hat{e}_t = \frac{n_t}{L_t}$ with estimated variance $\frac{\sum_{i=1}^{k}(\hat{e}_{it} - \hat{e}_t)^2}{k(k-1)}$ where $n_t = \sum_{i=1}^{k} n_{it}$ and $L_t = \sum_{i=1}^{k} L_{it}$. (This variance has the simple form of sample variance divided by sample size because the L_{it} are equal for all locations *i*.)

The same argument can be followed for the recce data.

2. Ratio Estimator $\hat{r}$:

A ratio estimate to adjust the encounter rate for recces is calculated as follows.

$$\hat{r} = \frac{\sum_{i=1}^{k} \hat{e}_{it}}{\sum_{i=1}^{k} \hat{e}_{ir}} \text{ with variance } \hat{\text{var}}(\hat{r}) = \frac{k}{\left(\sum_i \hat{e}_{ir}\right)^2} \frac{\sum_i (\hat{e}_{it} - \hat{r}\hat{e}_{ir})^2}{k-1} .$$

3. The combined adjusted encounter rate $\hat{e}_{ic}$:

The ratio estimate is used in the combined adjusted encounter rate as follows.

$$\hat{e}_{ic} = \frac{L_{it}\hat{e}_{it} + L_{ir}\hat{r}\hat{e}_{ir}}{L_{it} + L_{ir}} \text{ with variance } \hat{\text{var}}(\hat{e}_{ic}) = \frac{1}{(L_{it} + L_{ir})^2}\left[L_{it}^2 \hat{\text{var}}(\hat{e}_{it}) + L_{ir}^2 \hat{\text{var}}(\hat{r}\hat{e}_{ir})\right].$$

The variance of the product $\hat{\text{var}}(\hat{r}\hat{e}_{ir}) = \{\hat{r}\hat{e}_{ir}\}^2 \left(\frac{\hat{\text{var}}(\hat{r})}{\hat{r}^2} + \frac{\hat{\text{var}}(\hat{e}_{ir})}{\hat{e}_{ir}^2} \right)$ (approximately).

For the case with calibration, the overall estimate of encounter rate is simply $\hat{e}_c = \hat{e}_t$.

If $\hat{r}$ does not differ significantly from one, we might choose to assume that recce and transect encounter rates are the same (but see the comments above on comparability when transect counts are truncated), in which case the combined estimate of encounter rate would be

$\hat{e}_{ic} = \frac{L_{it}\hat{e}_{it} + L_{ir}\hat{e}_{ir}}{L_{it} + L_{ir}}$ with variance $\hat{\text{var}}(\hat{e}_{ic}) = \frac{1}{(L_{it} + L_{ir})^2}\left[L_{it}^2 \hat{\text{var}}(\hat{e}_{it}) + L_{ir}^2 \hat{\text{var}}(\hat{e}_{ir})\right]$.

Because there was the same amount of transect and recce effort at each location, overall encounter rate is now estimated as $\hat{e}_c = \frac{\sum_{i=1}^{k} \hat{e}_{ic}}{k}$ with variance $\hat{\text{var}}(\hat{e}_c) = \frac{\sum_{i=1}^{k}(\hat{e}_{ic} - \hat{e}_c)}{k(k-1)}$ (estimated from between-transect variability in encounter rate).

Location		(1) Transects only		(2) Calibrated		(3) Uncalibrated	
		ER	cv(ER)	ER	cv(ER)	ER	cv(ER)
1	01/RT1	13	*28%*	17.3	*31%*	19.6	*24%*
2	02/RT1	1	*100%*	5.7	*48%*	6.6	*43%*
3	03/RT1	17	*24%*	19.3	*29%*	21.8	*23%*
4	04/RT01	5	*45%*	7.4	*42%*	8.4	*37%*
5	04/RT02	7	*38%*	3.1	*51%*	3.4	*51%*
6	05/RT01	11	*30%*	9.6	*36%*	10.8	*32%*
7	06/RT01	31	*18%*	36.0	*25%*	40.6	*17%*
8	07/RT1	22	*21%*	12.2	*30%*	13.4	*27%*
9	08/RT1	7	*28%*	5.6	*44%*	6.2	*41%*
10	09/RT1	18	*24%*	20.9	*29%*	23.6	*22%*
11	10/RT1	8	*35%*	6.5	*41%*	7.2	*38%*
12	11/RT1	7	*38%*	4.7	*46%*	5.2	*44%*
13	12/RT1	18	*24%*	7.6	*34%*	8.2	*33%*
14	12/RT2	8	*35%*	10.6	*36%*	12.0	*31%*
15	12/RT3	2	*71%*	4.4	*53%*	5.0	*49%*
16	13/RT1	0		0.9	*114%*	1.0	*112%*
17	13/RT2	0		3.3	*62%*	3.8	*57%*
18	13/RT3	8	*35%*	8.2	*39%*	9.2	*35%*
19	15/RT1	1	*100%*	0.2	*0%*	0.2	*141%*
20	15/RT2	1	*100%*	0.4	*139%*	0.4	*143%*
21	15/RT3	1	*100%*	1.2	*91%*	1.4	*90%*
22	15/RT4	0		0.9	*114%*	1	*112%*
	Overall	**8.5**	***21%***	**8.5**	***21%***	**9.5**	***22%***

Table 1: Comparison of encounter rates (dung-piles per m), calculated using (1) line transect data only, (2) transect and calibrated recce data, and (3) transect and uncalibrated recce data. All analyses use transect data truncated at 2.5m.

Location		Transects only		Calibrated		Uncalibrated	
		D	cv(D)	D	cv(D)	D	cv(D)
1	01/RT1	49.0	*29%*	65.1	*32%*	73.7	*26%*
2	02/RT1	3.8	*100%*	21.6	*49%*	24.8	*44%*
3	03/RT1	64.1	*26%*	72.7	*31%*	82.0	*25%*
4	04/RT01	18.9	*46%*	27.9	*43%*	31.6	*38%*
5	04/RT02	26.4	*39%*	11.8	*52%*	12.8	*52%*
6	05/RT01	41.5	*32%*	36.3	*37%*	40.6	*33%*
7	06/RT01	116.9	*20%*	135.3	*27%*	152.6	*19%*
8	07/RT1	83.0	*23%*	45.8	*32%*	50.4	*29%*
9	08/RT1	26.4	*39%*	20.9	*45%*	23.3	*42%*
10	09/RT1	67.9	*25%*	78.7	*30%*	88.7	*24%*
11	10/RT1	30.2	*37%*	24.2	*42%*	27.1	*40%*
12	11/RT1	26.4	*39%*	17.6	*47%*	19.5	*45%*
13	12/RT1	67.9	*25%*	28.5	*35%*	30.8	*34%*
14	12/RT2	30.2	*37%*	39.9	*37%*	45.1	*32%*
15	12/RT3	7.5	*71%*	16.5	*53%*	18.8	*49%*
16	13/RT1	0.0		3.3	*114%*	3.8	*112%*
17	13/RT2	0.0		12.4	*63%*	14.3	*58%*
18	13/RT3	30.2	*36%*	30.8	*40%*	34.6	*36%*
19	15/RT1	3.8	*100%*	0.8	*10%*	0.8	*141%*
20	15/RT2	3.8	*100%*	1.4	*139%*	1.5	*144%*
21	15/RT3	3.8	*100%*	4.7	*91%*	5.3	*90%*
22	15/RT4	0.0		3.3	*114%*	1.0	*112%*
	Overall	31.8	*23%*	31.8	*23%*	35.7	*24%*

Table 2: Comparison of densities (dung-piles per ha), calculated using (1) line transect data only, (2) transect and calibrated recce data, and (3) transect and uncalibrated recce data. All analyses use transect data truncated at 2.5m.

ANNEX 5: Analysis of data and survey design for the mike central African pilot project. Third report – Part 1: Analysis of the pilot data

Len Thomas and Stephen T. Buckland

June 28 2001

Analysis of pilot data

Summary of data

Recce-transect data were collected at three sites: Odzala, Lope and Ituri. At each, the same protocol was used. Sample locations were assigned using a systematic grid. At each location, a 5 kilometer recce-transect was performed. This consisted of 5x200 meter transects separated by 4x1000 meter recces. For logistical reasons, our original suggestions for the layout of recce-transects were not followed (Scheme 1 and Scheme 2 from Buckland unpublished). Instead, a design was used that was easier to implement, but does not ensure that transects are located on the ideal line (Figure 1). Some locations at Ituri followed a different design (either a different type of recce-transect, or not consistently using 200m transects or 1000m recces); these have been excluded from this analysis.

In Odzala, the sampling grid was divided into three separate zones (strata), and in Lope there were 6 zones. We are not aware of any stratification of sampling effort at Ituri.

A summary of the number of locations at each site and the total counts at each location are given in Table 1, and histogrms of frequencies of counts in each 200m segment are given in Figure 2. These data are analyzed further subsequent sections.

In addition, at Ituri, there was historical data available from purposively-placed transects. These transects were first surveyed in 1993-95 and were repeated in 2000. These data are analyzed in a separate report.

Estimating dung-pile density from transect data

Transect data for all 3 pilot sites was extracted and imported into *Distance 3.5* (Thomas et al. 1998). Removing counts with unusually large distances (truncation) is standard practice in distance sampling analyses; this practice improves the reliability and precision of the detection function estimates (Buckland et al. 1993). After an initial inspection, a truncation distance of 4m was found to be satisfactory for all sites. The proportion of data truncated was 7% at Odzala, 12% at Lope and 9% at Ituri. Histograms of the truncated distance data are shown in Figure 3.

The Odzala data shows a considerable spike at 0 distance, with 11% of counts being recorded as exactly 0cm. Of the three observers, one in particular recorded a large proportion of 0 distances (18%). Such a detection function clearly indicates rounding of distances to 0, probably because dung piles that intersect the line are recorded as being on the line, rather than distance to the middle of the dung pile being mentioned.

Several strategies were investigated for obtaining reliable estimates of the detection function in the face of the spiked data. These include using key functions that do not fit the spike, combined with few adjustment terms, or grouping the data into intervals before analysis. All of these strategies produced similar results, and the final model used was half-normal key function with a cosine adjustment of order 2 on ungrouped data (shown in Figure 3).

Both the Lope and Ituri data appear to show the opposite problem – avoidance of zero distances (Figure 3). It seems likely that the observers at these sites were instructed to avoid recording zero distances, and so tended to record dung piles that were centered on the line as being further away. Despite these problems, the overall goodness-of-fit of reasonable detection function models was satisfactory (c^2 GOF for final models were 19.7 for Lope on 14 degrees of freedom (*df*) and 13.5 on 13 *df* for Ituri). For both Lope and Ituri there was little reason to favour any one of the three key functions (half-normal, uniform or hazard rate) – for example their AIC was very similar. For consistency with the above analysis, the final model chosen was half-normal with cosine adjustments. One adjustment of order 2 was used for Lope, but no adjustments produced a better fit for Ituri (probably because there were fewer data).

There was no evidence to support the stratification of detection function estimation by strata at Odzala or Lope. There was some evidence of differences between observers at Odzala, but there was too few observations for one of the observers to fit a separate detection function for each observer. It may be worthwhile investigating fitting observer as a covariable in future analyses, using the new MCDS engine in *Distance 4* (Thomas et al. 2001), although it is unlikely to add much to the analysis.

Results from the detection function modelling are summarized in Table 2.

Table 3 shows estimates of encounter rate and density for the 4 pilot sites. In this table, encounter rate has been estimated from variance in counts between locations, pooled across the sample zones for Odzala and Lope (there was only one sample zone in Ituri). The resulting estimate of encounter rate and density is not biased by ignoring the stratification because sampling intensity is the same in all zones.

The coefficient of variation in encounter rate ($CV(\hat{e}_t)$) is particularly high for Ituri. In addition to the smaller sample of locations, there is also one location, Boyea, where an unusually large number of dung piles were seen (36 compared with a mean for the other transects of 3.8). The Boyea region is known to contain very high elephant densities (J. Hart, pers. comm).

For all three sites, variance in encounter rate makes up the majority of the variance in the density estimate: 86% for Odzala, 90% for Lope and 95% for Ituri. Conversely, variance due to estimating the detection function makes up only 14%, 10% and 5% respectively of the variance of the density estimate.

Higher precision can often be obtained by estimating encounter rate separately by stratum (sampling zone). However, for both Odzala and Lope, there are some strata with relatively few locations in them, making the estimate of variance in encounter rate less reliable. Results for the stratified analysis are shown in Table 4. CVs for both Odzala and Lope are somewhat lower.

Comparing recce and transect encounter rates

In principal, the recce data that was collected at each location between the sections of transect can be used to improve the precision of the encounter rate estimates. However, encounter rates on recces and transects are unlikely to be the same: recces are more likely to follow elephant trails so have higher encounter rates; on the other hand observers move more slowly along transects so are less likely to miss dung piles. Therefore, before the recce data can be used, we must correct for any difference in encounter rate between recces and transects.

This is done by estimating a correction factor, $\hat{r}$, as the ratio of encounter rate on the transects with encounter rate on a subset of the recces. This ratio is then used to calibrate the remaining recce data. A combined adjusted encounter rate, $\hat{e}_c$ can then be calculated, using both the transect data and the calibrated recce data (see Appendix, section 2.4).

We estimated *r* using the 200m of recce data closest to each 200m section of transect, as shown in Figure 4. The ratio estimates and associated variance for each site are shown in Table 5. In all cases, the ratio is not significantly different from 1. However, the upper limits of the confidence intervals are quite wide (particularly for Ituri). Scatterplots of transect vs recce encounter rates for each site (Figure 5) confirm that the relationship is approximately linear. (The plot for Ituri also highlights the influential nature of the one sampling location with unusually high encounter rates.)

Improving efficiency of dung-pile density estimates using recce counts

The combined adjusted encounter rates (see Appendix) are shown in Table 6, together with the encounter rates calculated from the transect data alone. As expected, the adjusted estimates of encounter rate are similar to the estimate from transect data alone. For Odzala, the CV is also almost identical. For Lope and Ituri, however, the adjusted estimate is substantially more precise ($\mathrm{CV}(\hat{e}_c)$ is 25% smaller than $\mathrm{CV}(\hat{e}_t)$ at Lope and 31% smaller at Ituri).

The corresponding density estimates are shown in Table 7. The difference in CV between transect-only and combined estimates is smaller ($\mathrm{CV}(\hat{D}_c)$ is 22% more precise than $\mathrm{CV}(\hat{D}_t)$ at Lope and 28% more precise at Ituri). This is because CV of the density estimates is composed of variation due to estimating the detection function, which is unchanged, as well as variation due to encounter rate.

Discussion

Estimating dung-pile density from transect data

Accurate measurement of distances appears to be a significant problem for elephant dung data. Rounding to zero distances was observed in a previous analysis of historical transect data from Ituri, and occurs in the present dataset at Odzala. Reasonable analyses of the data were still possible through careful selection of an appropriate detection function, but opportunities for testing other assumptions and for further modelling of the detection function are lost if the measurements in the field are not accurate. Clearly, there is a need for comprehensive training of field crew. This is also illustrated by the Lope and Ituri data, where there appeared to be an avoidance of small perpendicular distances, probably because the field crew were told not to mark distances as zero. When dung is close to or across the line, one possible protocol is to measure the distance to one each side of the pile and record the average (distances on one side of the line would be recorded as negative).

There was not enough data for each observer to look for differences among observers. In the future, as more data accumulates, it may be desirable to include additional covariates, such as observer and habitat, in the detection function models, using the MCDS analysis engine in *Distance 4.0* (Thomas et al. 2001). Covariate modelling of the detection function has the potential to further reduce the variance of the estimate of effective strip width, although this is not the major component of variance in these analyses so the benefit will probably be slight.

One potential source of bias in the encounter rates recorded on transects was that the transects did not fall on the ideal survey line (Figure 1). It would appear to be impossible to implement a recce-transect survey protocol in the field where transects are exactly located on the ideal line. The potential size of this bias is unknown.

Comparing recce and transect encounter rates

The purpose of the calibration is to correct for any biases in encounter rate on recces. These biases could occur for a number of reasons: (i) recces do not follow the ideal line (Figure 1); (ii) on a small spatial scale, recces do not sample difficult-to-access habitat; (iii) observers move faster on recces and so may not observe as carefully; (iv) transect data is truncated, and the effective strip width is estimated relative to this truncation. Because the transect data were also not collected on the ideal line (Figure 1), there is no way to correct for this potential bias. If the transects were located on the line, then one could argue that it may be best to randomly sample sections of recce to use for calibration, as a way of getting a representative sample of the departure from the ideal line by recces. However, the disadvantage of this strategy is that additional variation is introduced to the ratio estimator: variation due to changes in dung-pile abundance along the line. Greater precision can be obtained by using the recce data closest to the transects for the calibration, because dung-pile density is likely to be more similar there. In either case, the calibration will correct for biases (ii) - (iv) mentioned above.

When choosing a subset of the recce data to use in the recce-transect calibration, two decisions must be made: which part of the recce these data should come from and how much data to use. We chose to use the recce data closest to each 200m section of transect, for the reason outlined above. Our justification for using the same length of recce and transect effort was that we expect that encounter rates on recce and transect are broadly similar and that variance of encounter rate is a function of encounter rate. In this case, maximum precision of $\hat{r}$ is obtained by using the same amount of data in both the numerator and denominator.

Improving efficiency of dung-pile density estimates using recce counts

At two of the three pilot sites, Lope and Ituri, the combined adjusted encounter rate e_c was more precise than the encounter rate calculated using transect data alone, e_t. At the other site, Odzala, the precision of e_c and e_t were almost identical. A number of factors affect the relative precision of the two estimates of encounter rate:

- e_c contains additional variance due to the inclusion of the ratio estimator, r. The expected variance of r will decrease as the sample size of sampling locations increases, although other factors are also important: in the pilot sites, the estimated CV for Ituri (14 locations) was similar to that of Lope (44 locations). var(r) will also decrease as the length of transect and recce used in the calibration at each site increase.

- the additional recce surveys included in e_c will decrease its variance. However, the size of this decrease will be affected by the amount of small-scale spatial correlation in encounter rate. If encounter rate on nearby sections of line are highly correlated, then the extra survey effort will be contributing little additional information and the decrease in variance will be relatively small. If, on the other hand, there is little correlation in encounter rate between nearby sections of line, then each additional section of line contributes new information, and so the variance of the overall encounter rate estimate will be decreased.

Figure 6 shows the correlation between paired 200 meter sections of recce at increasing distances apart. At Odzala, where the effect of the additional recce data was the least, the correlation between adjacent points is 0.67, only dropping to zero at between 3.5 and 4km. At Lope, where the additional recce data increased the precision of the encounter rate estimate, the correlation between adjacent points is 0.46, dropping to a low value at around 1km. Ituri shows an intermediate pattern, with an initial correlation of 0.55, but with a relatively shallow decline in correlation. The estimated correlations for Ituri are less precise, as there were fewer locations at that site.

A similar analysis could have been performed for transect data, but this would be unlikely to be useful as there are only 5 200m sections per location, with minimum distance apart of 1km.

Cost-effectiveness of alternative designs

Several different designs could be compared: (i) the pilot survey design; (ii) one contiguous transect at each sample location, with no recces; (iii) a recce-transect design with different allocations of effort to the recce and transect components. To determine the cost-effectiveness of these different designs, we need to estimate amount of time a survey would take under the design, and the variance of the resulting estimate.

There is at present relatively little information on timing from the pilot surveys. The only available data is from 6 of the sampling locations in Ituri (Table 8), where it took 67 minutes to do 1km of transect and 39 minutes to do 1km of recce (a ratio of 1.7:1). These results are quite different from those of Walsh and White (1999), from surveys in Gamba (594 minutes and 135 minutes, respectively; a ratio of 4.4:1). We conclude that there is little utility in doing detailed cost-effectiveness studies until more such data are available.

Neverthess, some guidance on alternative sample designs can be given.

A transect-only design (ii, above) will have a higher CV than the estimate for e_t because in the pilot study the transect segments were spaced 1km apart. Spatial correlation between nearby sections of line (Figure 5) will increase the variance if the transect segments were contiguous. In principal, this variance could be estimated from the pilot data. It seems likely that, if transects take only slightly more time to perform than recces, this design should be preferred. With a transect-only design, it is feasible to ensure that the ideal line is followed, negating that possible source of bias. The field protocol is simpler, as is the analysis. The overall variance, for a fixed cost, is unlikely to be much lower.

However, if transects take a great deal more time than recces (as at Gamba), the recce-transect design is likely to lead to more precise estimates of density. A number of variations of this design could be investigated further. The pilot data could be used to get an approximate estimate of the effect of varying the ratio of recce to transect effort. With more information about costs, an optimal ratio could be derived.

Conclusions

- A satisfactory analysis of dung-pile density was obtained from the transect data, although there were problems with the distribution of perpendicular distances at all sites.
- The pilot survey design meant that there is an unknown bias in the estimates, caused by transect lines not being located on the ideal line.
- Incorporating the recce data into the density estimates resulted in increased precision at Lope and Ituri, but not at Odzala. This is mainly caused by differences in the spatial correlation between encounter rates on nearby sections of line.
- The optimal layout of recees and transects will depend on the location, and the context of the survey. For the proposed extensive sample program, a consistent method should be used across all sampling locations. Based on the pilot data, it would appear that a recce-transect design is better than transects only. For the intensive programs at each MIKE site, different layouts could be used at each site, depending on the results of analyses such as these. However, based on the pilot data it seems that a recce-transect design should be preferred in the first instance.
- The pilot data can be used to refine the above recommendations by:
    - estimating the precision of a continuous, transect-only design, based on the variance of encounter rate on the transect sections, and the estimated correlation between paired sections of count at known distances. The historical transect-only data from Ituri could also be used in this respect.
    - determine which designs are most cost-effective, if sufficient information were available about the survey times at each site.
    - investigate the precision of different strategies for choosing the section of recce to use for calibration in the combined adjusted encounter rate estimate.
    - obtain approximate estimates of the variance that could be obtained from different combinations of recce and transect effort. Given information about survey times, an (approximate) optimal ratio could be derived.
    - examine the effect of stratification by sampling zone of the adjusted combined encounter rate estimate in Odzala and Lope.

References

Buckland, S.T. unpublished. MIKE Central African Pilot Project: proposals for survey design. WCS report.

Buckland, S. T., Anderson, D. R., Burnham, K. P. and Laake, J. L. 1993. Distance Sampling: Estimating Abundance of Biological Populations, Chapman and Hall, London.

Buckland, S. T. and F. M. Underwood. unpublished. Analysis of data and survey design for the MIKE central African pilot project: first interim report. WCS report. November 2000.

Thomas, L., Laake, J.L., Derry, J.F., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Strindberg, S., Hedley, S.L., Burt, M.L., Marques, F.F.C., Pollard, J.H. And Fewster, R.M. 1998. Distance 3.5. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. Available: http://www.ruwpa.st-and.ac.uk/distance/

Thomas, L., Laake, J.L., Strindberg, S., Marques, F., Borchers, D.L., Buckland, S.T., Anderson, D.R., Burnham, K.P., Hedley, S.L., and Pollard, J.H. 2001. Distance 4.0. Beta 1. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. http://www.ruwpa.st-and.ac.uk/distance/

Walsh, P.D., and L.J.T. White. What it will take to monitor forest elephant populations. Conservation Biology, 13: 1194-1202.

Appendix

In the following, suffix $_t$ indicates a quantity assocated with transects, suffix $_r$ indicates recces. In sections 6.1 and 6.2, a distinction is made between the sections of recce, $_u$, used to derive the ratio estimator, and the remaining part of the recce, $_s$, which is used to derive the combined adjusted encounter rate.
In the pilot surveys, each location is composed of 5 replicate transects, each of length l_{ijt} = 200m, giving L_{it} = 1000m. Between these are 4 replicate recces, each of length l_{ljr} = 1000m, giving L_{ir} = 4000m (Figure 1). In sections 6.1 and 6.2 the recces are divided into two parts, of length L_{iu} = 1000m and L_{is} = 3000m (Figure 2).
At each pilot site, the total number of sample locations at a pilot site is denoted k, while the total dung pile count on transects is n_t, and the total dung pile count on recces is n_r. The corresponding encounter rates are e_t and e_r.

Ratio estimator

The ratio estimate to adjust the encounter rate of recces is calculated by

$$\hat{r} = \frac{\sum_{i=1}^{k} \hat{e}_{it}}{\sum_{i=1}^{k} \hat{e}_{iu}} \text{ with variance } \hat{\mathrm{var}}(\hat{r}) = \frac{k}{\left(\sum_i \hat{e}_{iu}\right)^2} \frac{\sum_i \left(\hat{e}_{it} - \hat{r}e_{iu}\right)^2}{k-1}$$

where $\hat{e}_{it}$ is the estimated encounter rate on transects at sampling location i, and $\hat{e}_{iu}$ is the estimated encounter rate on the calibration section of the recces at sampling location i (Buckland and Underwood, unpublished).

The combined adjusted encounter rate

The ratio estimate is used in the combined adjusted encounter rate as follows.
$\hat{e}_{ic} = \frac{L_{it}\hat{e}_{it} + L_{ir}\hat{r}\hat{e}_{is}}{L_{it} + L_{is}}$ with variance $\hat{\text{var}}(\hat{e}_{ic}) = \frac{1}{(L_{it} + L_{is})^2}\left[L_{it}^2\ \hat{\text{var}}(\hat{e}_{it}) + L_{is}^2\ \hat{\text{var}}(\hat{r}\hat{e}_{is})\right]$, where the variance of the product $\hat{\text{var}}(\hat{r}\hat{e}_{ir}) = \{\hat{r}\hat{e}_{ir}\}^2\left(\frac{\hat{\text{var}}(\hat{r})}{\hat{r}^2} + \frac{\hat{\text{var}}(\hat{e}_{ir})}{\hat{e}_{ir}^{\ 2}}\right)$ (approximately; Buckland and Underwood, unpublished).

Overall encounter rate is then estimated as $\hat{e}_c = \frac{\sum_{i=1}^{k}\hat{e}_{ic}}{k}$ with variance

$\hat{\text{var}}(\hat{e}_c) = \frac{\sum_{i=1}^{k}(\hat{e}_{ic} - \hat{e}_c)^2}{k(k-1)}$. (This variance has the simple form of sample variance divided by sample size because the sample effort is the same for all locations.)

Distance sampling estimates of density

Dung-pile density at a pilot site is estimated from transect data by

$$\hat{D} = \frac{n_t}{2\hat{m}L_t}$$

where n_t is the total count of dung-piles on transects, $\hat{m}$ is the estimated effective strip width and L_t is the total line length. This can be re-expressed in terms of encounter rate as

$$\hat{D} = \frac{n_t}{L_t}\frac{1}{2\hat{m}} = \frac{\hat{e}_t}{2\hat{m}}$$

where $\hat{e}_t$ is the estimated encounter rate for transects. The variance is well approximated by

$$\hat{\text{var}}(\hat{D}) = \hat{D}^2\left([CV(\hat{e}_t)]^2 + [CV(\hat{m}_t)]^2\right)$$

(Buckland et al. 1993). When using the combined adjusted encounter rate, substitute $\hat{e}_c$ for $\hat{e}_t$ in the above.

Table 1. Summary of recce-transect data from pilot sites.

Site	k	n_t [1]	n_r	$\hat{e}_t$ [1] dung/km	$\hat{e}_r$ dung/km
Odzala	44	449	1565	10.20	8.89
Lope	44	174	478	3.95	2.72
Ituri	14 [2]	77	266	5.5	4.75

k = number of sampling locations, subscript $_t$ indicates transect, $_r$ recce, n is total count and $\hat{e}$ is encounter rate

[1] These values are before truncation. Table 2 shows counts and encounter rates for transect data after truncation at 4m.

[2] 22 locations were sampled in Ituri, but 8 did not follow the standard recce-transect sampling plan and were excluded from these analyses

Table 2. Summary of detection function modelling of transect data. Data were analyzed after truncation at 4m.

Site	n_t	$\hat{m}$ m	$CV(\hat{m})$
Odzala	418	1.85	5.9
Lope	153	2.79	7.1
Ituri	70	1.80	14.1

$\hat{m}$ is estimated effective strip width, CV is coefficient of variation

Table 3. Estimates of encounter rate and density from analysis of transect data. Encounter rates for Odzala and Lope have been pooled across sampling zones (strata). Units for density are dung piles/hectare. CI is 95% parametric confidence interval; CI_b is the 2.5th and 97.5th percentile from 999 bootstrap resamples, resampling locations.

Site	k	e_t dung/km	CV(e_t)	D dung/ha	CV(D)	95% CI(D) dung/ha	95% CI_b(D) dung/ha
Odzala	44	9.5	14.5	25.6	15.6	(18.8, 35.0)	(18.1, 38.8)
Lope	44	3.5	20.8	6.22	22.0	(4.0, 9.6)	(4.3, 9.6)
Ituri	14	5.0	50.0	13.9	52.0	(4.9, 39.4)	(4.5, 24.7)

Table 4. Estimates of encounter rate and density from transect data. Encounter rate has been stratified by sampling zone for Odzala (3 zones) and Lope (6 zones). Notation for CI as for Table 3.

Site	k_s	D dung/ha	CV(D)	95% CI(D) dung/ha	95% CI_b(D) dung/ha
Odzala	13, 22, 9	23.8	13.9	(18.0, 31.5)	(18.0, 32.9)
Lope	5, 3, 15, 4, 15, 2	4.5	21.1	(2.9, 7.0)	(3.2, 6.7)
Ituri	14	13.9	52.0	(4.9, 39.4)	(4.5, 24.7)

Table 5. Estimated ratio of encounter rate on transects with encounter rate on nearby recces (see text and Figure 3 for details of how recce data were chosen). Transect data were truncated at 4m. 95% CI is calculated assuming ratio follows a lognormal distribution. *p* is *p*-value from a Wilcoxon's signed-ranks test.

Site	$\hat{r}$	$\mathrm{CV}(\hat{r})$	$95\%\mathrm{CI}(\hat{r})$	*p*
Odzala	1.02	7.7	(0.88, 1.19)	0.99
Lope	1.29	17.4	(0.92, 1.80)	0.64
Ituri	0.93	18.7	(0.66, 1.34)	0.38

Table 6. Comparison of encounter rates calculated using transect data only and transect and calibrated recce data.

Site	$\hat{e}_t$ dung/km	$CV(\hat{e}_t)$	$\hat{e}_c$ dung/km	$CV(\hat{e}_c)$
Odzala	9.50	14.49	9.09	14.78
Lope	3.48	20.81	3.49	15.62
Ituri	5.00	50.01	4.43	34.53

Table 7. Comparison of density estimates calculated using transect data only and transect and calibrated recce data.

Site	$\hat{D}_t$ dung/ha	$CV(\hat{D}_t)$	$\hat{D}_c$ dung/ha	$CV(\hat{D}_c)$
Odzala	25.64	15.63	24.53	15.90
Lope	6.22	21.99	6.24	17.15
Ituri	13.89	51.97	12.30	37.31

Table 8. Time taken to complete 1km of recce and transect at 6 sites in Ituri.

	Time (minutes/km)	CV(Time)
Recce	39.0	10.6
Transect	67.4	21.5

Figure 1. Design used to lay out recce-transects at each location in pilot sites. From R. Beyers, pers. comm.

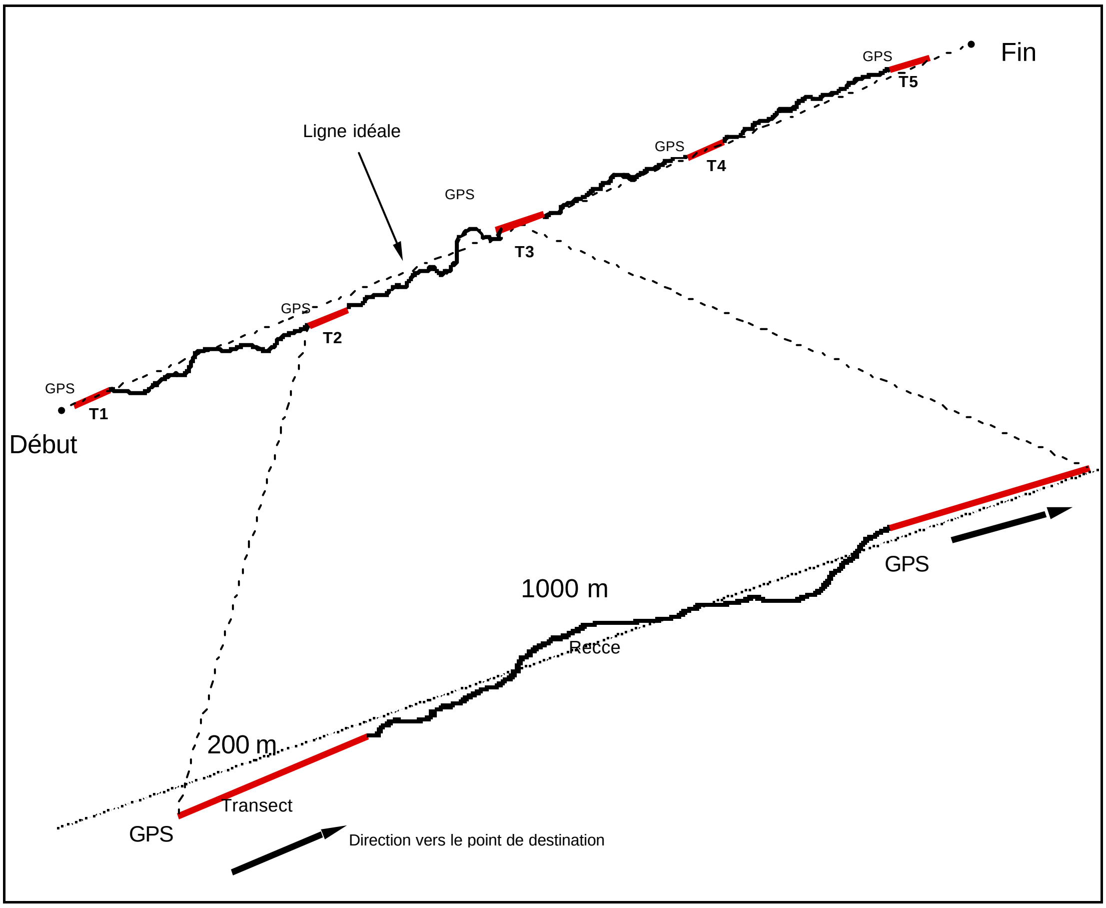
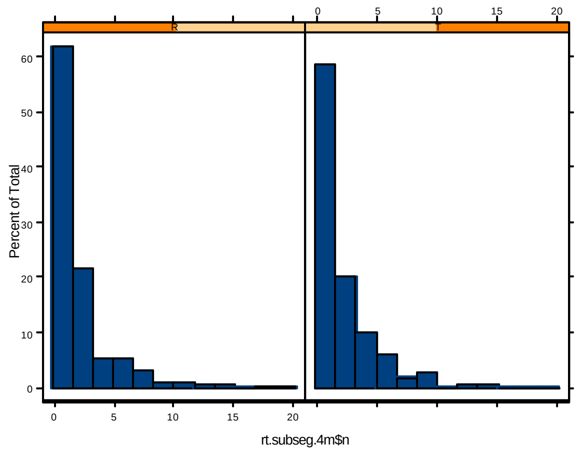
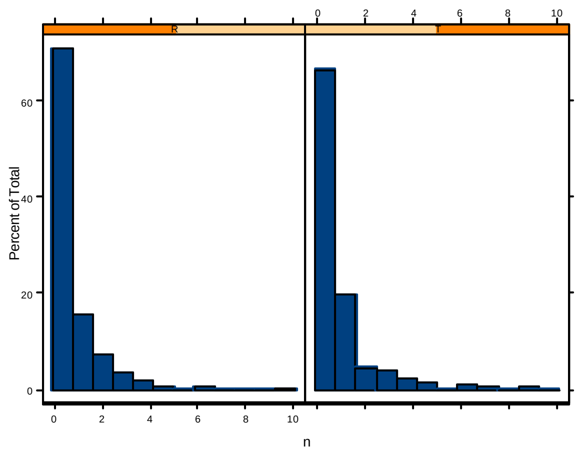


Figure 2. Histograms showing frequency of counts in each 200m segment of the lines. The left hand pane is recce data, and the right transect data.
a) Odzala



b) Lope



c) Ituri

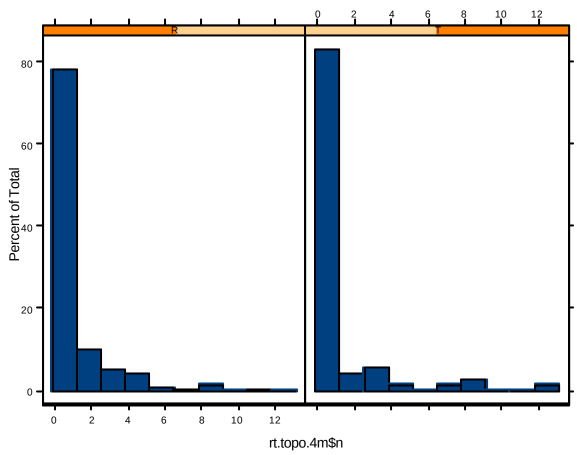
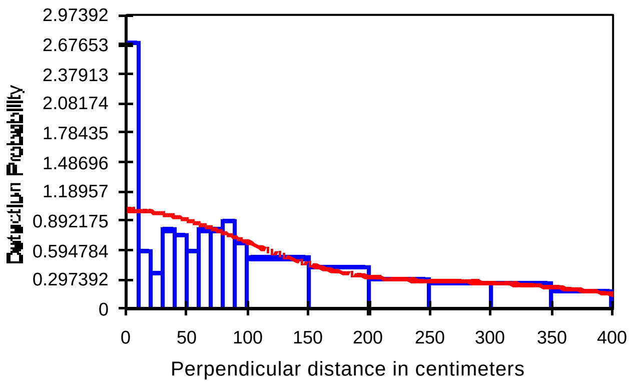
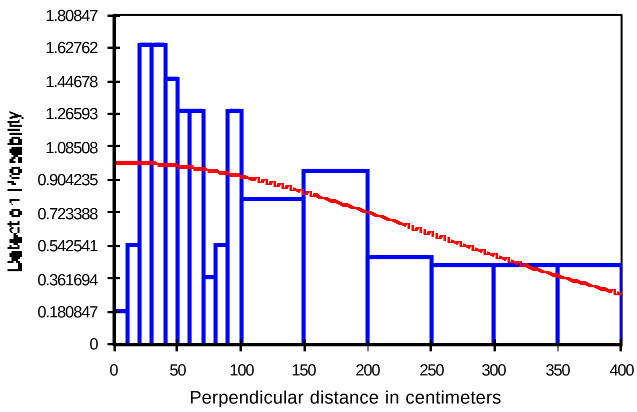


Figure 3. Fitted detection functions and scaled histograms of counts for pilot survey transect data.
a) Odzala



b) Lope



c) Ituri

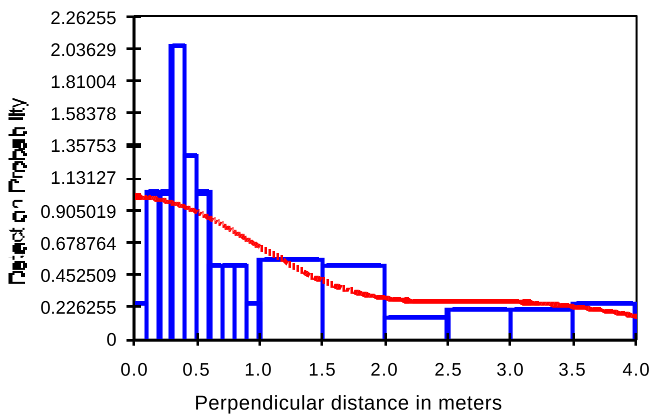


Figure 4. Arrangement of recce and transect used to compare recce and transect encounter rates. The horizontal line is the recce-transect line (which is not exactly straight in reality). Distances in meters are given along the bottom. Transects T1 - T5 (1000 meters in total) are compared with recce sections U1 - U5 (1000 meters in total), to derice the ratio estimate *r*. This is then applied to recce sections S1-S4 (3000 meters in total) to estimate expected encounter rate on these sections of recce.

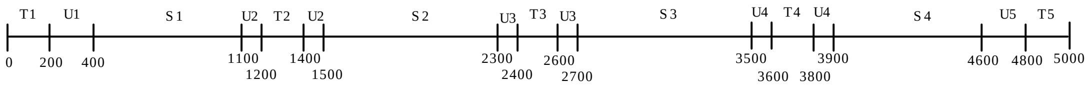
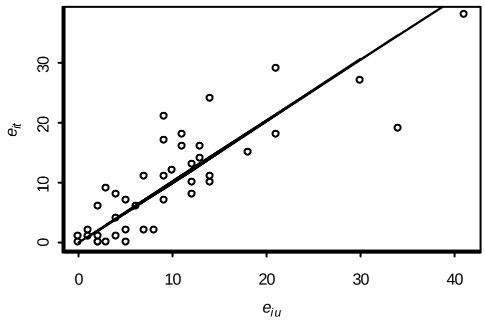
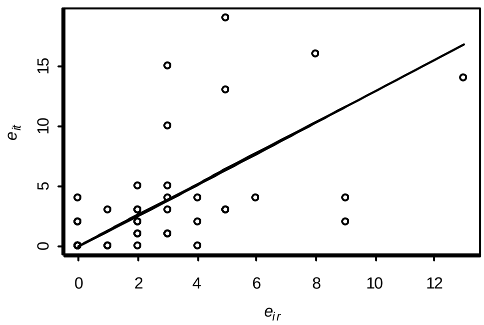


Figure 5. Scatterplot of encounter rate on transect $\hat{e}_{it}$ vs encounter rate on the segment of recce used to calcluate $\hat{r}$ r, $\hat{e}_{iu}$. The line $y = \hat{r}x$ is shown.

a) Odzala



b) Lope



c) Ituri

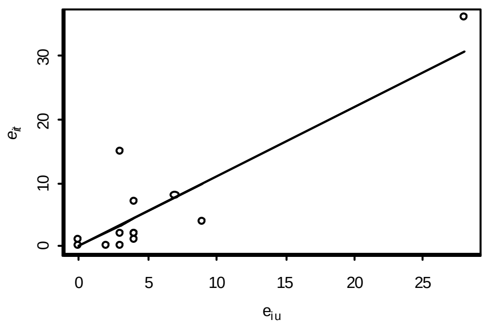
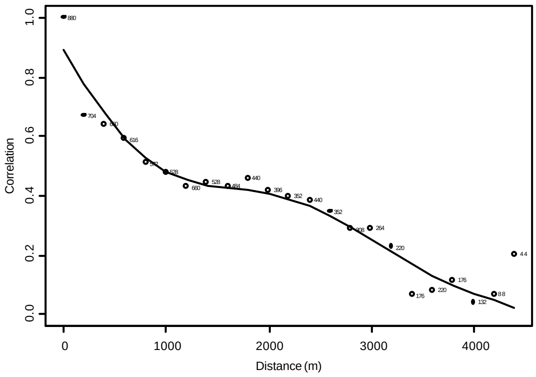
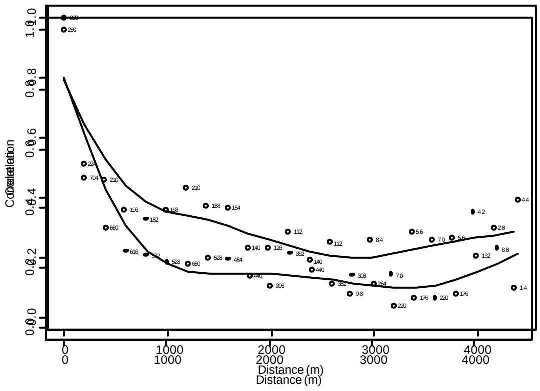


Figure 6. Correlograms of 200m sections of recce at the three pilot sites. The number of pairs used to calculate each correlation is shown beside each point. The data have been fit using a spline function, with 3 df.

a) Odzala



b) Lope



c) Ituri

ANNEX 6: Predicting variation in encounter rate at pilot sites.

Len Thomas

9 July 2001

Introduction

Our goal here is to predict the expected variation in encounter rate for different recce-transect designs.

At a pilot site, there are i=1 ... k sampling locations. Each location consists of a 5km recce-transect. This consists of 5x200m transects separated by 4x1000m recces. If the recce sections are divided up into 200m segments, then there are 25x200m segments in total. Let n_{ij} be the count on each segment j at location i (j=1 ...m, where m=25).

As shown in Figure1, there is positive correlation between counts on nearby segments of line (the figure was calculated using only the recce segments). This correlation generally decreases with increasing separation of the segments.

Derivations

Because of this correlation, successive n_{ij} on a line are not independent, and the sample variance of counts on a line, $\hat{\mathrm{var}}_i\left(n_{ij}\right)$, will be smaller than the expected variance of the counts, s^2.

The sample variance,

$$\hat{\mathrm{var}}_i\left(n_{ij}\right) = \frac{\sum_{j=1}^{m}\left(n_{ij} - \overline{n}_i\right)^2}{m-1} \qquad (1)$$

We now expand (1) and take expectations. Assuming that the expected count on a segment within a line is constant, $E\left(n_{ij_1} n_{ij_2}\right) = s_{j_1 j_2} + E\left(n_{ij}\right)^2$, where $s_{j_1 j_2}$ is the covariance in counts between segments j_1 and j_2. When $j_1 = j_2$, we have

$$E\left(n_{ij}^2\right) = s^2 + E\left(n_{ij}\right)^2. \qquad (2)$$

Let us assume that the covariance between segments is a function of the distance between segments, ie:

$$E\left(n_{ij_1} n_{ij_2}\right) = s_d^2 + E\left(n_{ij}\right)^2 \qquad (3)$$

where s_d^2 is the covariance at distance d between segments j_1 and j_2. Let the distance between two adjacent segments be D. In our case, D = 200m. Then, expanding (1), taking expectations, and substituting in (2) and (3), we obtain

$$\mathrm{var}_i\left(n_{ij}\right) = s^2 - \frac{2}{m(m-1)}\left[(m-1)s_D^2 + (m-2)s_{2D}^2 + ... + s_{[m-1]D}^2\right] \qquad (4)$$

Let the correlation between counts at distance d, $r_d = s_d^2 / s^2$. Then,

$$\mathrm{var}_i\left(n_{ij}\right) = s^2\left(1 - \frac{2}{m(m-1)}\left[(m-1)r_D + (m-2)r_{2D} + ... + r_{(m-1)D}\right]\right) \qquad (5)$$

In a continuous line of *m* segments, there are (*m* - 1) pairs of segments distance *D* apart, (*m* - 2) pairs distance 2*D* apart, etc. There are $m(m-1)/2$ pairs of segments in total. The second expression in brackets in (5) is, therefore, a weighted mean of the correlations between counts between pairs of segments a fixed distance apart, weighted by the number of comparisons at that distance.

$$\mathrm{var}_i(n_{ij}) = \sigma^2(1-\bar{\rho}) \qquad (6)$$

where $\bar{\rho} = \frac{2}{m(m-1)}\left[(m-1)\rho_D + (m-2)\rho_{2D} + ... + \rho_{(m-1)D}\right]$.

Given data consisting of counts at intervals D at a series of sampling locations, (6) can be used to estimate σ^2, by substituting estimates for the other quantities in (6). var(n_{ij}) can be estimated as the mean of the variance between counts on segments on each line:

$$\mathrm{var}(n_{ij}) = \sum_{i=1}^{k}\left(\mathrm{var}_i(n_{ij})\right)/k$$

and ρ_d can be estimated using the empirical correlations between counts (Figure 1).

We also wish to derive a formula for $\mathrm{var}\left(\sum n_{ij}\right) = \mathrm{var}(n_i)$. From the definition of the variance of a sum:

$$\begin{aligned}\mathrm{var}(n_i) &= \sum_{j=1}^{m}\sigma^2 + 2\sum_{j_1<j_2}^{m}\sigma_{j_1 j_2} \\ &= m\sigma^2 + 2\left[(m-1)\sigma_D + (m-2)\sigma_{2D} + ... + \sigma_{(m-1)D}\right] \\ &= m\sigma^2\left(1 + \frac{2}{m}\left[(m-1)\rho_D + (m-2)\rho_{2D} + ... + \rho_{(m-1)D}\right]\right) \\ &= m\sigma^2\left[1 + (m-1)\bar{\rho}\right] \qquad (7)\end{aligned}$$

where $\bar{\rho}$ is defined as for (6). Given an estimate of σ^2 from (6), and of ρ_d from the empirical autocorrelation function (Figure 1), this formula can be used to predict the variance of counts at a location for a given sampling design.

In presenting results, we usually work with encounter rates rather than variances. Conversion is easy, e.g., $\mathrm{var}(e_i) = \mathrm{var}(n_i)/l^2$ where l^2 is the length of a line.

Application

Recce data

There is another layer of complexity in the recce-transect design: that is estimation of the ratio of encounter rates in transects and recces, and from this estimation of the combined adjusted estimate of encounter rate. For the present, we sidestep this complication by considering only the recce subset of the data.

Figure 1 shows the correlation between segments of recce at given distances – called a correlogram. The data points (as opposed to the fit) are used as an estimate of ρ_d at each site. We begin by considering the recce design as implemented, i.e., 20 sections of 200m recce, with every 5th recce separated by a 200m gap (where the transects were done). Table 1 shows the estimated variance in encounter rate at a location, $\hat{\mathrm{var}}(\hat{e}_{ri})$, derived from substituting estimates for expectations into (7) and converting from var(*n*) to var(*e*). The table also shows the variance in encounter rate at a location, estimated empirically from the data as variance in encounter rate between locations:

$$\hat{\mathrm{var}}_e(\hat{e}_{ri}) = \frac{\sum (e_{ri} - e_r)^2}{k-1} \tag{8}$$

This estimate includes components of spatial variation in encounter rate between locations as well as variation between counts along the line, and should be higher than $\hat{\mathrm{var}}(\hat{e}_{ri})$. However, as can be seen from Table 1, it is not appreciably higher. The reason for this is not clear.

Table 2 shows predictions of variance in encounter rate assuming that 1, 2, 3 and 4km of continuous recce were done. As would be expected, variance decreases with increasing effort, until with 4km of continuous recce it is only slightly larger than the predicted variance ($\hat{\mathrm{var}}(\hat{e}_{ri})$) from Table 1. The predicted variance in Table 1 is smaller because there the 4km of recce effort was spaced out over 5km of line so the spatial autocorrelation was smaller.

Transect data

We now compare the precision of various transect-only designs. Since the transects in the pilot study were spaced 1000m apart, it is not possible to use them to construct a correlogram containing useful information about correlation at short distances. We have therefore used the recce correlogram (Figure 1) in the following, under the assumption that the patterns of correlation in transects should be very similar. (This could be tested to some extent by constructing the transect correlograms and comparing them to the recce ones, althout it would not be a very reliable indicator.)

The furthest-apart transects were 4800m apart, further than the furthest-apart recce segments (4400m). Therefore, to estimate $\boldsymbol{r}_{4800}$, we used the predicted value from the lines in Figure 1, extrapolated to 4800m. (Other methods could be imagined here, for example fitting a parametric curve to the correlograms.)

Table 3 gives the predicted variance, $\hat{\mathrm{var}}(\hat{e}_{ti})$, for the transect-only data (ie 5, 200m sections of transect, separated by 1000m), and the empirical variance, $\hat{\mathrm{var}}_e(\hat{e}_{ti})$, estimated from variation in transect counts between locations. Unlike the recce only data, the predicted variance is significantly smaller than the empirical one. A smaller variance is what would be expected as the predicted variance in encounter rate at a location does not contain the between-location component of variance. The reason the predicted variance is not significantly smaller for the recce data in Table 1 is unknown.

Table 4 shows predicted variance for 1, 2, 3 and 4km of continuous transects. Comparing these to the predicted variance for the pilot design 1km of transect-only data (Table 3), we see that at Odzala, the 1km of transect data in the pilot design is more precise than even 4km of continuous transect. At Lope and Ituri, the pilot design is less precise than 3km of continuous transect.

Why at Odzala should even 4km of continuous transect be less precise than 5x200m transects, separated by a 1km gap? The answer lies in the relatively high correlations between counts at distances of up to 1km at Odzala (Figure 1). Figure 2 shows $\bar{\boldsymbol{r}}$ at each site for a continuous survey of 200m to 4400m, as well as $\bar{\boldsymbol{r}}$ for recce-only design (dotted line) and transect-only pilot survey design (dashed line). At Odzala, the relatively high correlations at small distances mean that continuous lines have a high $\bar{r}$, even for 4400m of continuous surveys. Spacing segments of survey line 1km apart (dashed line), means that these high correlations are avoided and $\bar{r}$ is significantly lower. Because the correlation is lower, the higher variance one would expect from the transect-only pilot design because only 1km of line is surveyed is offset by the decrease in correlation (equation (7)).

At Lope and Ituri, the correlations between distances of 200-400 meters are somewhat higher than at larger distances, but correlation at larger distances decreases slowly, if at all (Figure 1). Therefore, the difference in correlation between the 5x200m design and the 4x1000m design are not great (Figure 2; the lines at Ituri are particularly close because the correlation at 1200m is unusually high). Therefore the gain in precision due to the distance between samples in the 5x200m design is offset by the extra survey effort in a continuous design -- in these sites at somewhere between 2000 and 3000m of survey effort.

Preliminary timing data from Ituri and Odzala indicate that 1km of continuous transect line takes between 1.7 and 2.5 times as long as 1km of recce. Under these timings, between 2.6 and 3.35km of continuous transect could be done in the same amount of time as the recce-transect design used in the pilot study. The above analysis shows that this amount of transect would yield similar or slightly greater precision than analysis of the transect-only data at Lope and Ituri, but not at Odzala.

We have previously shown that the combined recce-transect data gave better precision than transect-only data at Lope and Ituri (22% and 28% more precise, respectively: Thomas and Buckland, unpublished, Table 7). Therefore, our preliminary conclusion is that for the same amount of time, somewhat better precision will be obtained from the recce-transect design used in the pilot survey than from a continuous transect design. However other factors (such as the desire to re-locate the same transect lines each year) may cause us to favour a transect-only design over a recce-transect one.

Conclusions

- Odzala is different in character from Lope and Ituri
- At Odzala, best precision is obtained by spacing out the transect segments, as in the pilot design, by about 1km or more. This is because counts on nearby segments of line are highly correlated, and this correlation decreases significantly over the space of 1km or so. Given that the segments are spaced out, there is no great advantage to collecting and analyzing recce-data in-between the transect segments.
- At Lope and Ituri, between 2 and 3 kilometers of continuous transect gives about the same precision as the transect-only part of the pilot design (i.e. 200m of transect, separated by 1km). This is because the correlation between counts on nearby segments of transect is either lower to begin with (Lope) or doesn't decrease as much (Ituri) over the first 1km. (It is correlation at nearby segments that is the most important for determining the overall variance in counts) However, when the recce data from the pilot study are included with the transect data, the recce-transect design is more precise than 2-3 kilometers of continuous transect. Based on preliminary time data, observers could survey between 2.6 and 3.35 kilometers of continuous transect in the same time as the 5km recce-transect.

- We therefore conclude that continuous transects will give somewhat lower precision per unit time than a recce-transect design (or noncontinuous transect design at Odzala).
- However, transects offer other advantages, such as being exactly on the ideal line and being re-locatable between time periods. It may be possible to make an initial estimate of the precision lost due to not returning to exactly the same line on successive surveys, using information about variation in counts along the pilot survey lines, if the accuracy of relocating lines were known.
- Further work includes: investigating why predicted and empirical variances($\hat{\mathrm{var}}(\hat{e}_{ri})$ and $\hat{\mathrm{var}}_e(\hat{e}_{ri})$ in Table 1) were so similar for the recce-only data; incorporating information about the time taken to move through the study areas without surveying in order to compare non-continuous transect-only designs; investigating other methods for estimating correlation at distances greater than the pilot data.

References

Thomas, L. and S.T. Buckland. unpublished. Analysis of data and survey design for the MIKE central African pilot project. Third Report – Part 1: Analysis of Pilot Data.

Table 1. Estimates of variance in encounter rate at a sampling location for recce-only data, using the design implemented in the pilot survey (4km total recce effort per site, allocated as 4 blocks of 1000m, separated by 200m). $\hat{\text{var}}_e(\hat{e}_{ri})$ is estimated sampling variance in counts between locations (formula 8); $\hat{\text{var}}(\hat{e}_{ri})$ predicted variance in encounter rate at a sampling location (formula 7). See text for more details.

Site	$\hat{\text{var}}(n_{ij})$	$\bar{r}$	$\hat{s}^2$	$\hat{e}_{ri}$	$\hat{\text{var}}_e(\hat{e}_{ri})$	$\hat{\text{var}}(\hat{e}_{ri})$
Odzala	3.6734	0.4305	6.4504	8.8920	79.6174	74.0182
Lope	0.9325	0.1974	1.1619	2.7215	6.8640	6.9010
Ituri	1.8135	0.3626	2.8453	4.7500	29.5096	28.0612

Table 2. Predicted variance in encounter rate at a sampling location for recce-only data, assuming 1, 2, 3 and 4km continuous recce.

Site	1km	2km	3km	4km
Odzala	113.2487	96.8982	88.2169	79.9616
Lope	13.6975	9.9141	8.2898	7.2107
Ituri	41.7042	35.5032	31.6810	28.9234

Table 3. Estimates of variance in encounter rate at a sampling location for transect-only data, using the design implemented in the pilot survey (1km total recce effort per site, allocated as 5 blocks of 200m, separated by 1000m). Weighted mean correlation, $\bar{r}$, is calculated using the recce data, as described in the text. See Table 1 for explanation of symbols and text for more details.

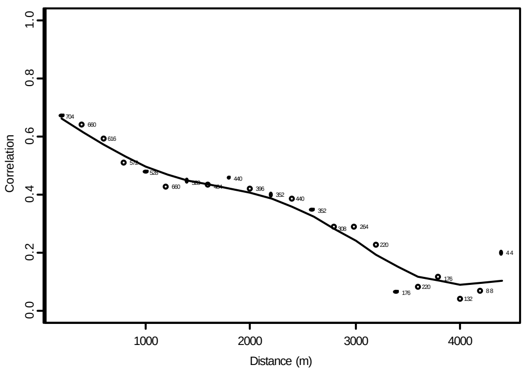
Site	$\hat{\text{var}}(n_{ij})$	$\bar{r}$	$\hat{s}^2$	$\hat{e}_{ti}$	$\hat{\text{var}}_e(\hat{e}_{ti})$	$\hat{\text{var}}(\hat{e}_{ti})$
Odzala	3.8022	0.2988	5.4227	9.5000	83.3256	59.5212
Lope	1.2295	0.1572	1.4589	3.4772	23.0560	11.8809
Ituri	2.2214	0.3452	3.3929	5.0000	87.5384	40.3949

Table 4. Predicted of variance in encounter rate at a sampling location for transect-only data, assuming 1, 2, 3 and 4km continuous transect.

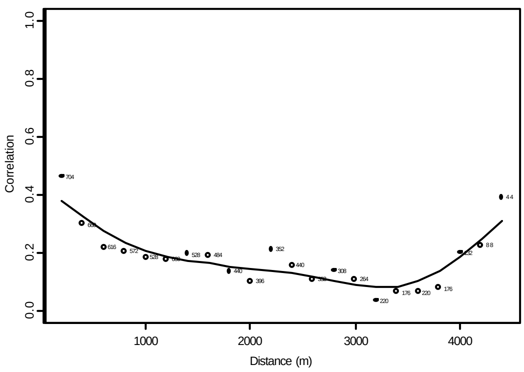
Site	1km	2km	3km	4km
Odzala	95.2043	81.459	74.1610	67.2210
Lope	17.1979	12.4476	10.4082	9.0534
Ituri	49.7310	42.3364	37.7786	34.4903

Figure 1. Correlograms of 200m sections of recce at the three pilot sites. The data have been fit using a spline function, with 3 df. (Unlike Fig 7. of Thomas and Buckland, unpublished, this figure does not contain 0 distance, and the spline is not weighted by sample size).

a) Odzala



b) Lope



c) Ituri

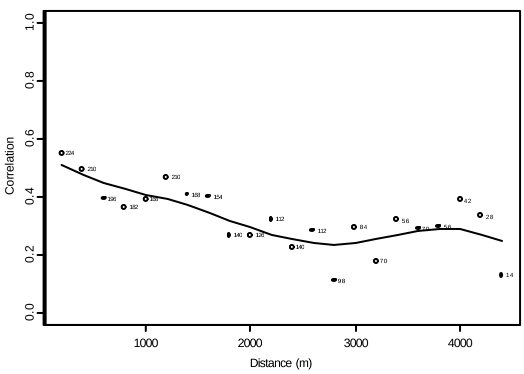
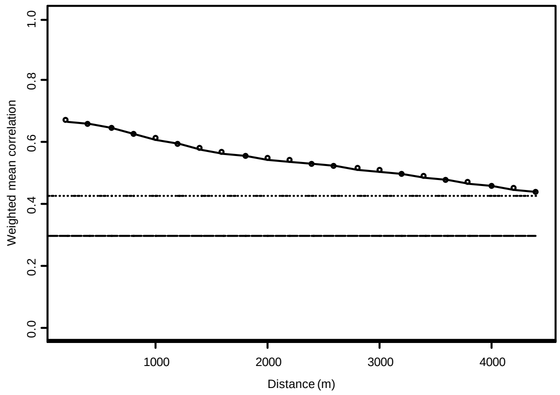
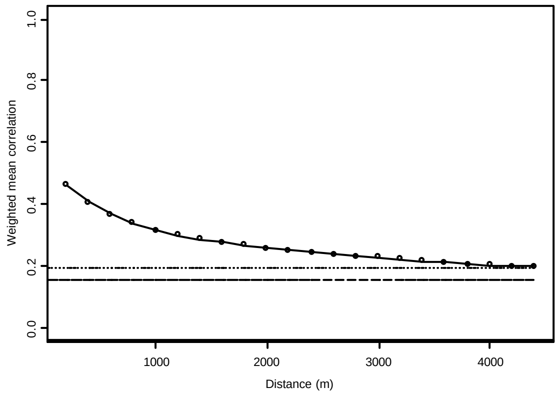


Figure 2. Weighted mean correlation vs total continuous recce length of recce at the three pilot sites. Dotted line is predicted mean correlation for recce-only data from pilot survey, and dashed line is for transect-only data.

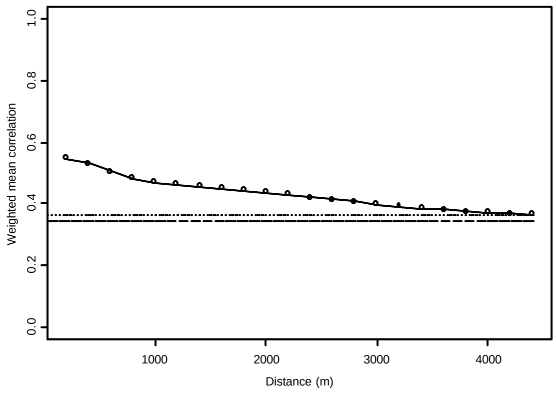
a) Odzala



b) Lope



c) Ituri


Weighted mean correlation
0.0
0.2
0.4
0.6
0.8
1.0
1000
2000
3000
4000
Distance (m)

ANNEX 7: Analysis of data and survey design for the MIKE central African pilot project: Third Report – Part 3: Analysis of Ituri transects

Len Thomas

August 21 2001

Draft

Introduction

Line transects of elephant dung were performed at Ituri in 1993-95 (hereafter referred to as 1995). As part of the MIKE pilot project, some of these transects were repeated again in 2000. This report presents results of an analysis of these data. The analysis was undertaken in late March 2001 and presented to John Hart and Rene Beyers during their visit to St Andrews April 2-4 2001. At that time only a subset of the transect data were available, so this analysis covers only that subset.

Data and analysis

At the time of this analysis, data were available from 10 locations containing transects covered in both 1995 and 2000. In many cases, not all transects at the location were covered in both time periods, but only the subset covered at both times was used (Table 1). Although transects were re-located in 2000 (John Hart, pers. comm.), the line length of the transects was often not the same in both time periods (Table 1). Total line length for transects repeated in both years was 81.1km in 1995 and 75.7km in 2000.

Data from all transects within a location were grouped for the analysis, to form independent sampling units. This gave a sample size of 10 locations.

After some exploratory analysis, it was decided to truncate the data for both years at 4m. Total observed dung piles after truncation was 183 in 1995 and 242 in 2000, giving encounter rates of 2.26 and 3.20 dung piles/km respectively.

There was evidence of considerable rounding to 0 distance in the 1995 data, while the 2000 data appeared to show some avoidance of small distances (possibly due to observers being discouraged from recording 0 distance). Various analysis strategies were investigated, including the use of key functions that avoid fitting the spike to 1995 data (e.g., half-normal or uniform with few adjustments), and grouping the data. There seemed little reason to use the same function for both years as the histograms were obviously so different. For the 1995 data, a uniform key function with 2 cosine adjustment terms was chosen, while for the 2000 data a half-normal key function with no adjustments was used. Estimated effective strip widths were 2.02 and 2.39 meters with CVs 7.6% and 5.4% respectively.

Density estimates

For the 1995 data, estimated density was 5.59 dung piles/hectare, with CV of 28.7%, 95% parametric confidence limits 2.98, 10.46 and 95% bootstrap confidence limits of

2.49, 8.19. 92.8% of the estimated variance in density came from variation in encounter rate between locations and only 7.2% from estimating the detection function.
For the 2000 data, estimated density was 6.67 dung piles/hectare, with CV of 16.6%, 95% parametric confidence limits 4.64, 9.57 and 95% bootstrap confidence limits of 4.87 and 8.95. 89.3% of the estimated variance in density came from variation in encounter rate between locations and 10.7% from estimating the detection function.
From these results, it is clear that there has been no significant change in dung density at these sampling locations between time periods. A simple approximate test for trend can be conducted using a *z*-test:

$$z = \frac{\hat{D}_1 - \hat{D}_2}{\sqrt{\hat{\text{var}}(\hat{D}_1) + \hat{\text{var}}(\hat{D}_2)}}$$

This gives $z = 0.567$, p=0.57.

Discussion

It is important not to over-interpret the results. Firstly they are based on only a subset of the data. Secondly, it is not known how transects were chosen to be repeated or not – these may be a non-random subset of the transects originally performed in 1993-95. Lastly, and possibly most importantly, the original transects were not laid out according to a randomized design. This means that it is not statistically valid to extrapolate the densities observed on the transects to the study area as a whole.
The analysis presented here could be improved in the following ways:

1. A re-analysis using all of the data.
2. The test for trend could be improved by using a paired z-test approach. This would take advantage of the fact that the same transects were used. Density on each location would be predicted from the distance analysis, and the difference in density between sampling periods calculated. A one-sample z-test could then be performed on these differences.
3. Rather than relying on the design-based estimates, which are not strictly valid in this case, it would seem better to use a model-based approach. An example of such an approach using the Odzala pilot data, is given in part 4 of this report.

Table 1. Locations covered in both survey periods.

Location code	Transect	Line length	
		1993-95	2000
ANZI	1	5	5
AKOT	3	5	4.3
	4	5	5
HTEK	1	5	3
	2	5	3.5
ISOR	1	5	5
	2	4.5	4
LEND	2A	5	5
	2B	4	5
MAGB	1	5	4.8
	2	5	5
MAWA	1	5	3.3
	2	5	3.8
MEHW	1	4.1	5
	2	2.5	5
SET	1	5	2.5
	2	5	2.5