

Geneva Forum: Towards Global Access To Health

Visualization and Analysis of Health Information: Visual exploration of spatio-temporal patterns in Epidemiological data using Geographic Information System (GIS) technologies

27th May 2008

Benjamin K. Mayala

Outline

- ❑ Introduction
- ❑ Objective
- ❑ Research Questions
- ❑ Method
- ❑ Description of Patterns
- ❑ Statistical and Visualization methods
- ❑ Visual environment
- ❑ Evaluation of Tool
- ❑ Results
- ❑ Conclusions

Introduction

□ Data

- Large amount of health data is collected
- Recorded, stored and presented in tabular format
 - Has disadvantages

□ Use of data

□ Epidemiologists

- Questions: who, what, *where, when*, why, how
- Lack familiarity with visualization - prefer statistics
 - Statistical methods - unable to show direction and speed of spread
 - Difficult to identify patterns
 - Where are clusters
 - Their change in Time

Visualization

- use of computer-supported, interactive, visual representations of (geospatial) data to strengthen understanding
- Purpose of visualization is insight, not pictures (or maps)
- Insight = discovery and explanation > decision making



Data Presentation

Ward	No HF	Pop 00	JAN 00	FEB 00	MAR 00	APR 00	MAY 00	JUN 00	JUL 00	AUG 00	SEP 00	OCT 00	NOV 00	DEC 00
Wangama	1	746	22	28	15	12	12	11	10	8	16	8	13	9
Iwungilo	1	1067	11	13	7	7	3	3	3	3	2	5	4	8
Kifanya	2	1702	35	58	78	154	218	107	122	226	92	82	56	43
Luponde	1	236	2	0	5	4	3	2	0	3	2	4	3	5
Matola	5	2770	92	109	88	85	72	73	59	77	70	47	57	51
Uwemba	1	4209	1	2	4	2	1	3	1	1	1	0	1	0
Yakobi	1	828	41	38	29	29	18	13	15	9	5	14	28	29
Igosi	3	3711	105	96	116	110	171	68	68	83	86	85	75	61
Imalinyi	2	2209	81	120	33	193	188	63	55	52	119	41	49	32
Idamba	1	606	0	0	0	0	0	0	0	0	0	0	0	0
Ikondo	2	1290	44	74	113	98	98	42	42	29	31	50	69	60
Kidegembye	2	1109	23	25	32	27	32	21	13	8	10	21	34	35
Lupembe	4	2473	48	71	58	64	59	62	54	74	63	45	42	42
Igongolo	1	1749	0	0	0	0	0	0	0	0	0	0	0	0
Ikuka	3	4988	138	122	136	96	157	128	129	150	90	79	95	84
Mahongole	1	1354	62	22	14	24	15	32	86	137	100	52	35	40
Makombako	9	11857	359	350	477	339	173	298	316	233	225	235	320	264
Mtwango	1	679	15	43	41	68	53	48	45	55	32	32	46	36
Mdandu	3	2161	54	39	54	92	58	54	59	48	109	56	47	61
Usuka	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Njombe Urban	6	5771	205	259	272	221	221	132	156	204	305	310	223	249
Ilembula	2	2028	69	158	136	184	157	109	61	85	70	77	78	37
Luduga	2	2863	287	221	194	301	246	97	82	112	127	88	88	133
Saja	1	1600	59	124	31	89	6	29	72	18	55	32	725	60
Wanging'ombe	2	3392	79	91	57	115	42	63	53	29	27	27	42	35



Objective

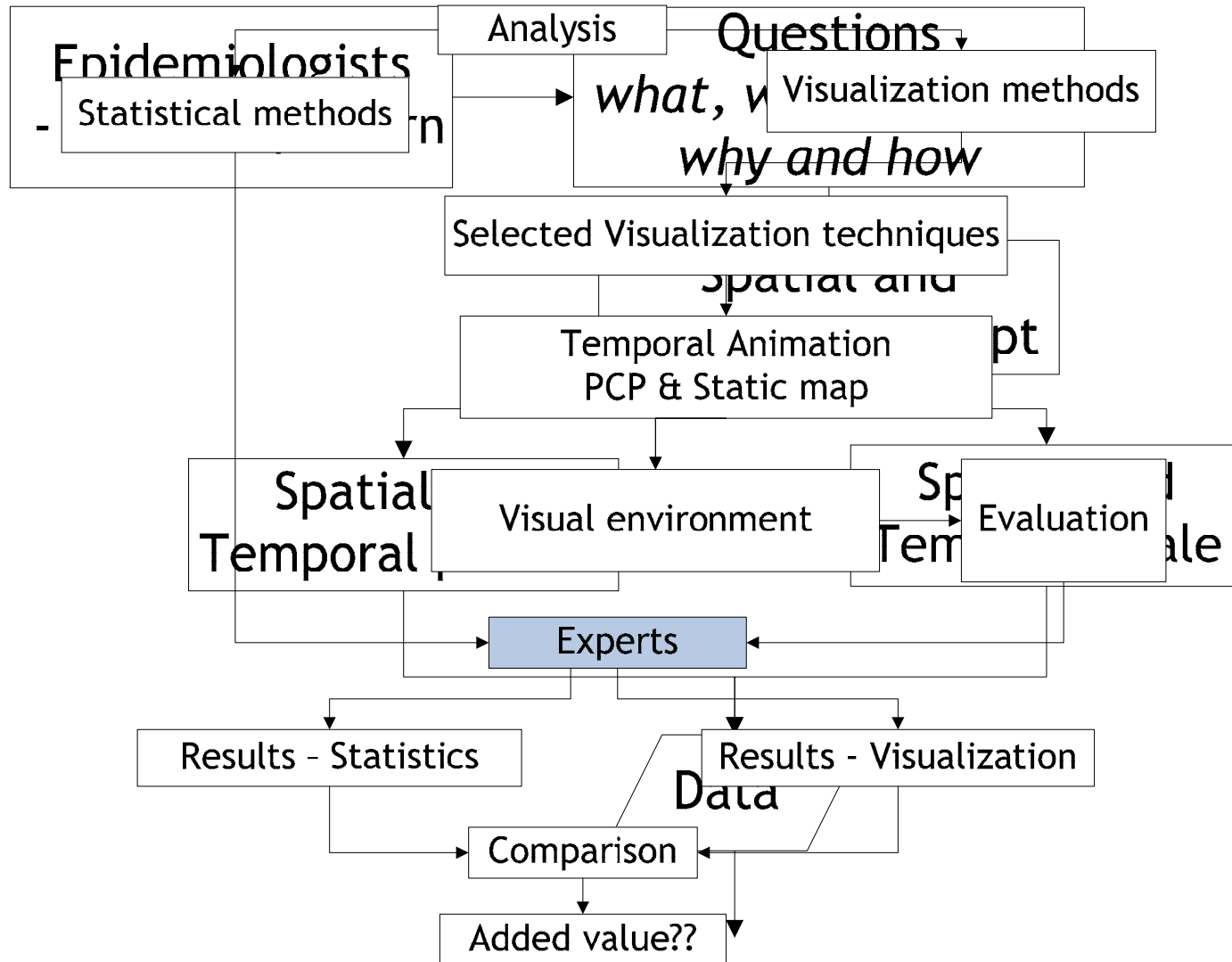
- ❑ To investigate how visualization can have added value to statistical methods in finding spatial, temporal and spatio-temporal patterns in epidemiological data.
 - ❑ Using data on Malaria reported cases in Njombe district, Tanzania - 2000 to 2004

Research Questions

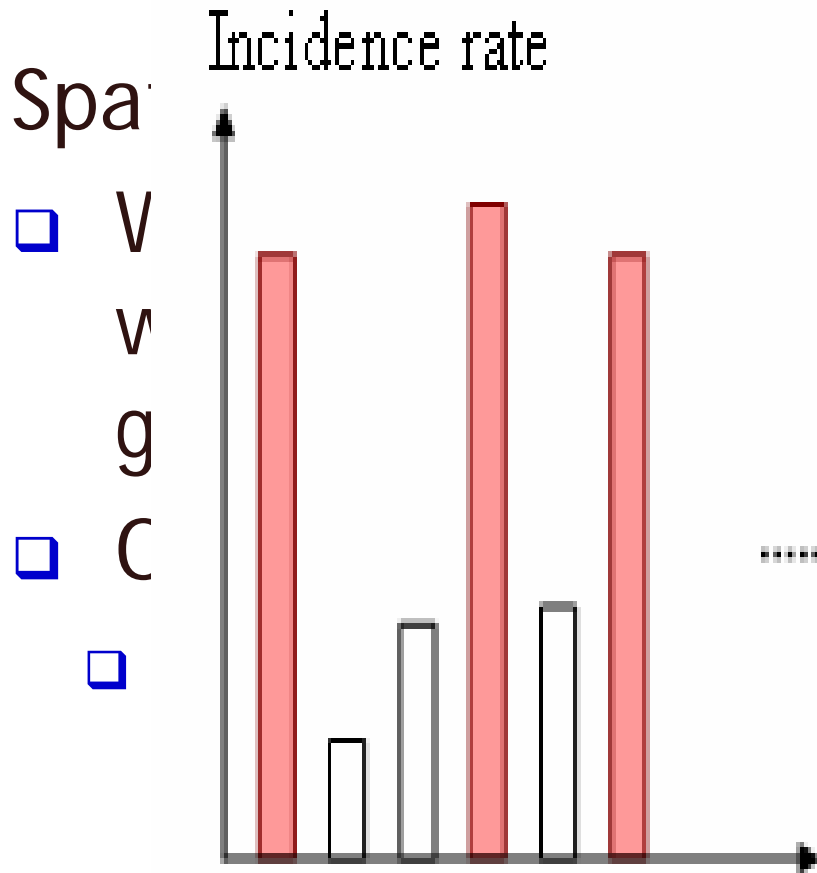
Related to:

1. Characteristics of spatial, temporal and spatio-temporal patterns
2. Availability of statistical methods to investigate these patterns in epidemiology
3. Availability visualization techniques to investigate these patterns
4. Choice of the visualization techniques
5. Possible integration of visualization and statistical methods

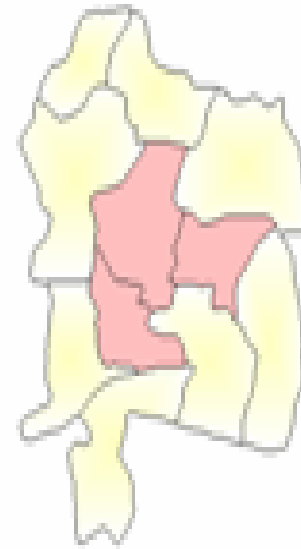
Method



Description of patterns



- V
- v
- g
- C
-



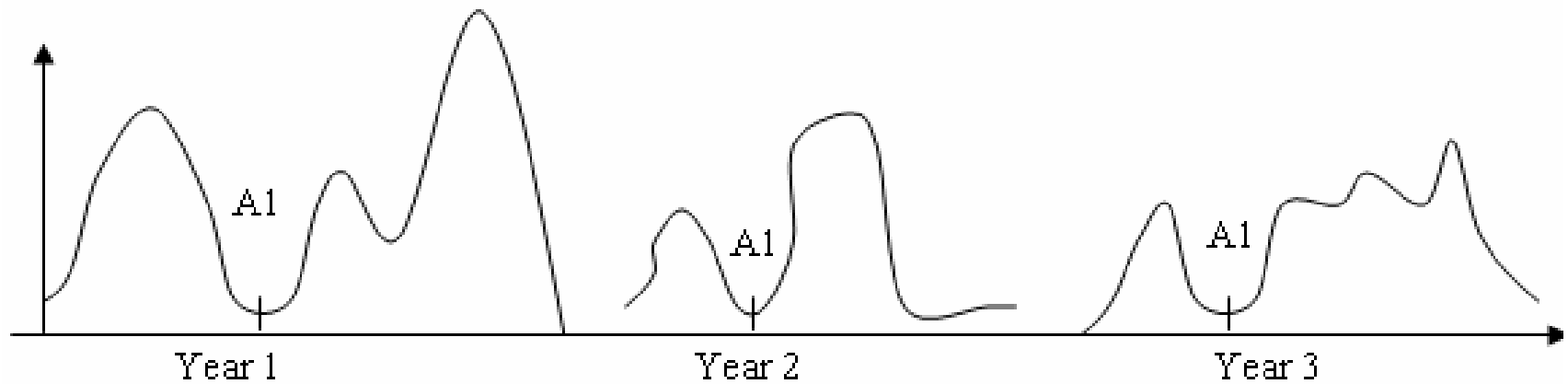
ia
atially

the
e)

Description of patterns Cont...

Temporal

Incidence rate



Year 1
Year 2
Year 3

BEHAVIOURS

Description of patterns Cont...

Spatio-temporal

- Wards with similar value (high or low) of incidence rate of diseases that are spatially grouped together in different times

Visualization methods reviewed

- ❑ Single static map
- ❑ Series of static maps
- ❑ Temporal animation
- ❑ Space-time cube (STC)
- ❑ Scatter plot matrix
- ❑ Parallel Coordinate plot (PCP)
- ❑ Start plot
- ❑ Chernoff faces
- ❑ Pixel Bar Chart (PBC)
- ❑ Tree Map

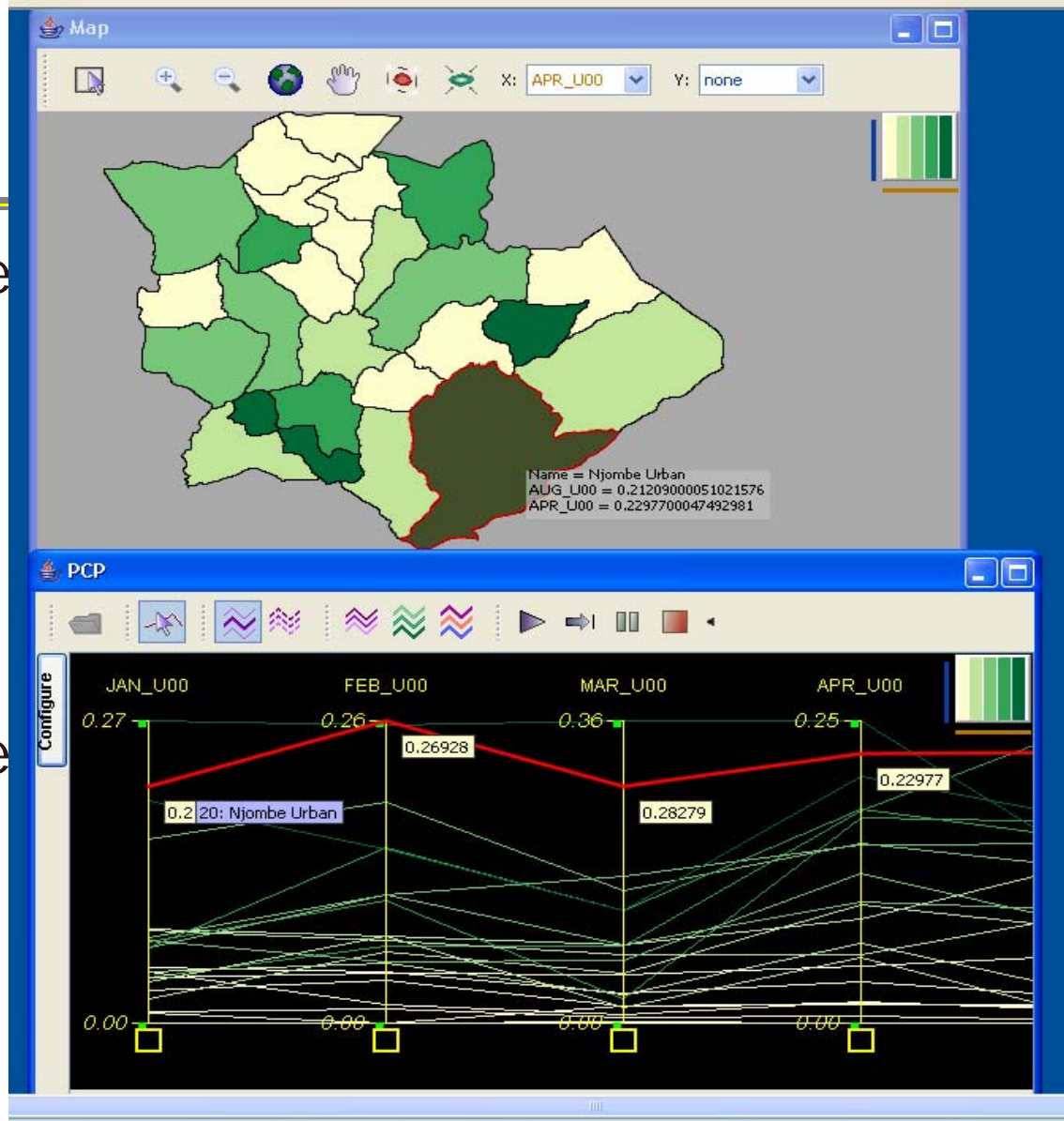
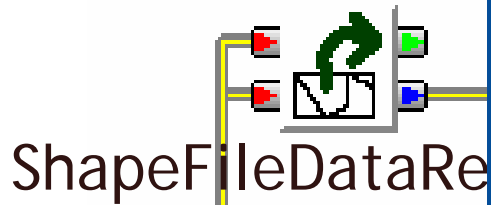
Visualization Techniques Selected

- those with potentials to provide insight into both spatial and temporal characteristics of the dataset

Temporal animation and PCP

Visual environment

■ GeoVISTA



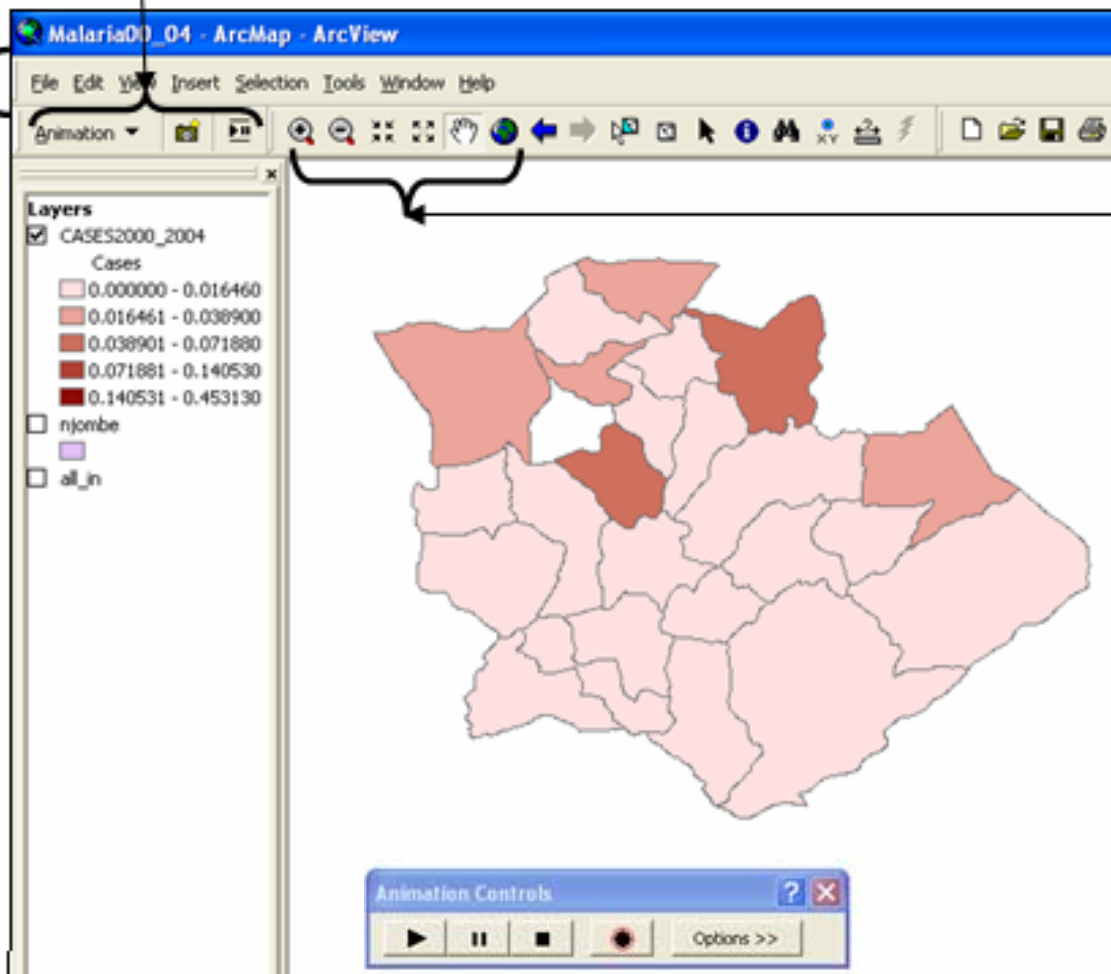
anager

Visual environment Cont...

□ ArcGIS 9.2 - Animation

Animation tool

Menu Bar



Basic tools

The screenshot displays the ArcGIS 9.2 interface. The title bar reads 'Malaria00_04 - ArcMap - ArcView'. The menu bar includes 'File', 'Edit', 'View', 'Insert', 'Selection', 'Tools', 'Window', and 'Help'. The toolbar contains various icons, with the 'Animation' tool (a play button) highlighted. The 'Layers' panel on the left shows a legend for 'CASES2000_2004' with five categories: 'Cases' (0.000000 - 0.016460), '0.016461 - 0.038900', '0.038901 - 0.071880', '0.071881 - 0.140530', and '0.140531 - 0.453130'. Other layers 'njombe' and 'all_in' are also listed. The main map area shows a map of a region with colored polygons representing the case categories. An 'Animation Controls' dialog box is open at the bottom, showing play, pause, and stop buttons, along with an 'Options >>' button.

Evaluation

- Tasks: Spatial, temporal and spatio-temporal patterns

Test procedures: 6 participants

- *3 Participants*



- *3 Participants*





NIMR

Results

Part.	Visualization					
	Spatial patterns		Temporal patterns		Spatio-temporal patterns	
	Identify	Describe	Identify	Describe	Identify	Describe
P1						
P2						
P3						
P4						
P5						
P6						

Results Cont...

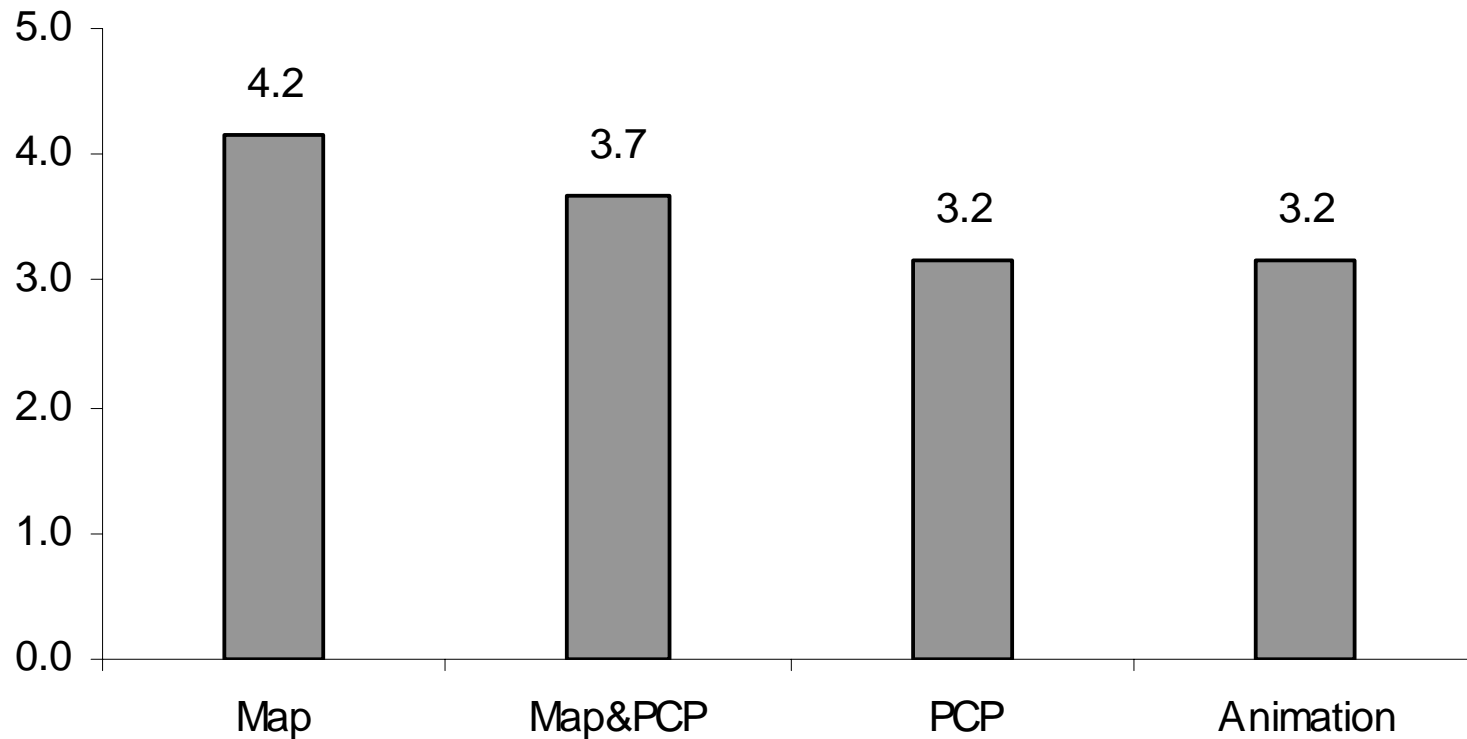
Statistics

Part.	Spatial patterns				Temporal patterns				Spatio-Temporal patterns			
	Method	Software	Identify	Describe	Method	Software	Identify	Describe	Method	Software	Identify	Describe
P1		SPSS										
		SAS*			T-test*, ANOVA*	SPSS						
P2	Monte Carlo*				Chi-Square					Excel		
										SPSS		
P3	Monte Carlo*									SatScan*		
P4						SPSS						
P5						SPSS						
P6		Matlab				Matlab				Matlab		

Results Cont...

Average score for effectiveness

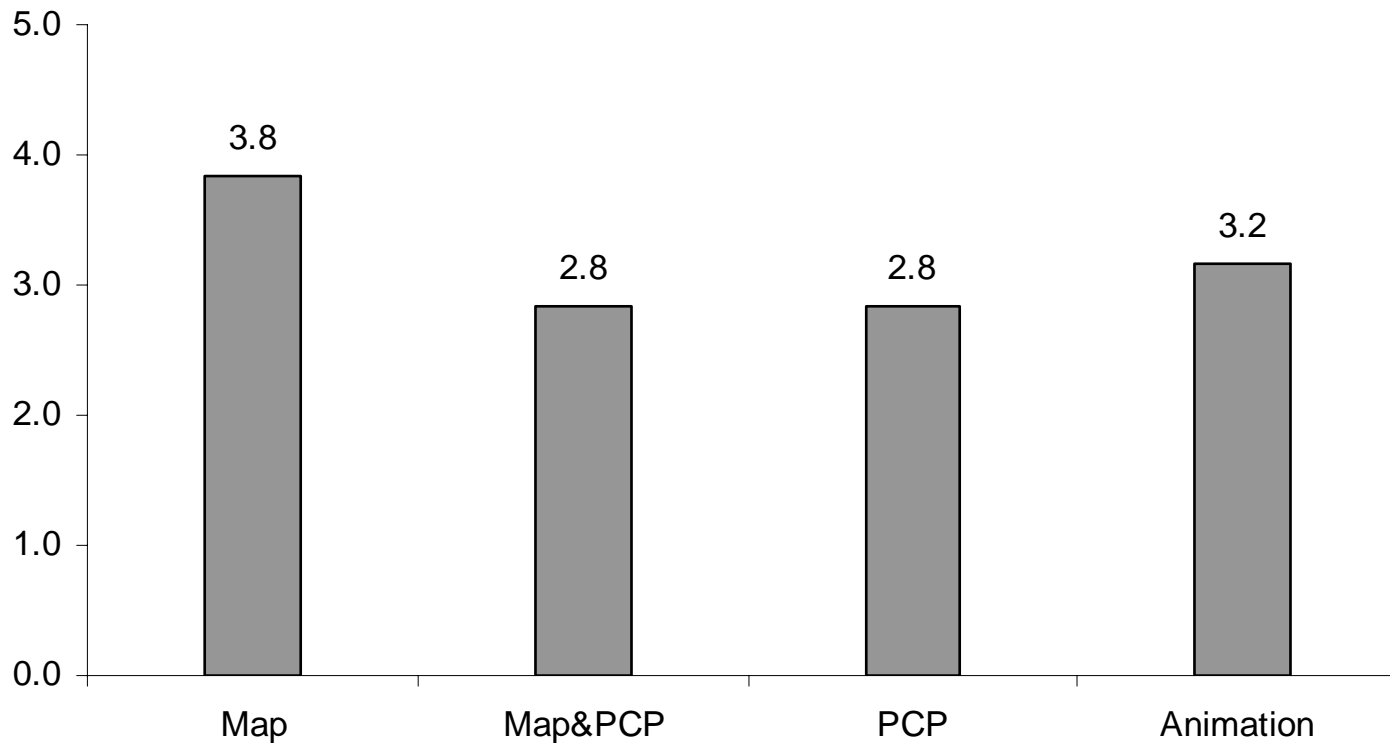
(1 = Very poor, 2 = Poor, 3 = Fair, 4 = Good, 5 = Very good)



Results Cont...

Average score for efficiency

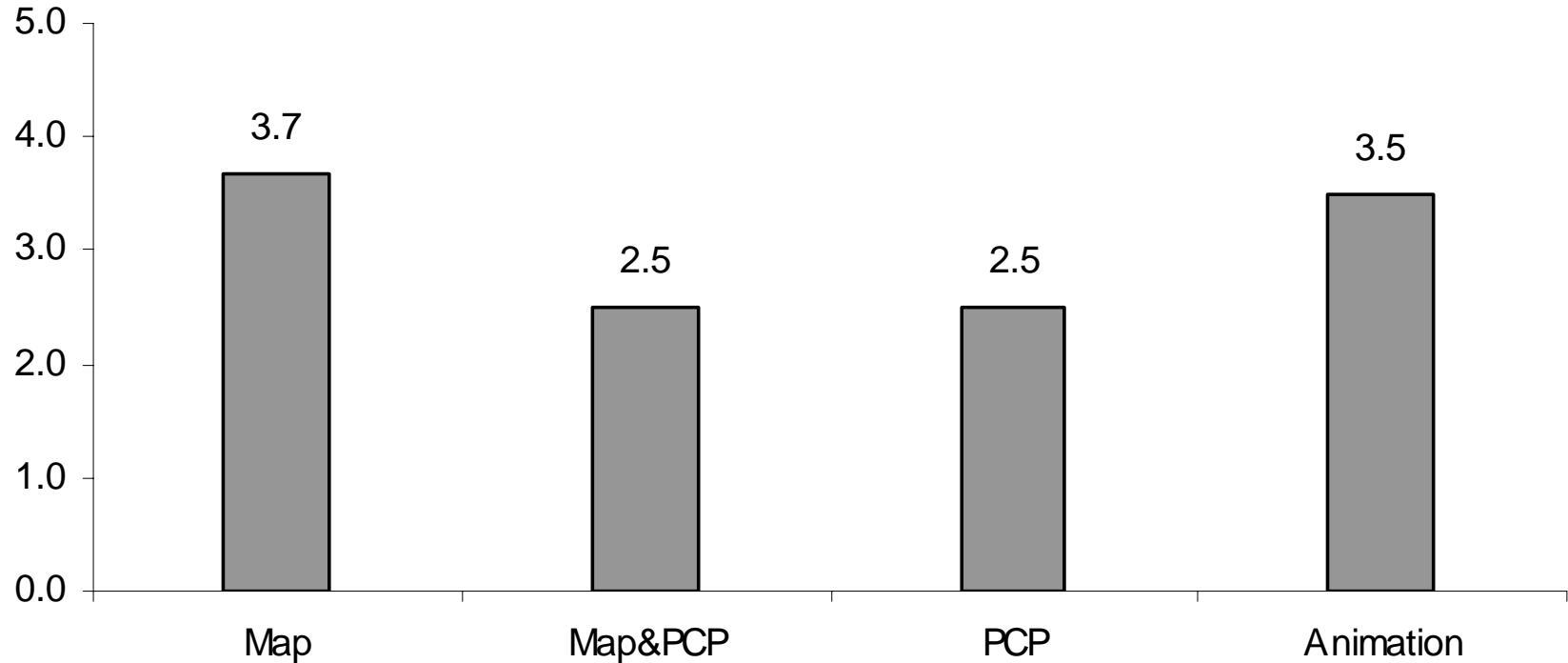
(1 = Very poor, 2 = Poor, 3 = Fair, 4 = Good, 5 = Very good)



Results Cont...

Average score for satisfaction with visualization

(1 = Very poor, 2 = Poor, 3 = Fair, 4 = Good, 5 = Very good)



Conclusion

- ❑ Visualization methods can be used first where by they give an idea of the patterns. When this idea is precise, a hypothesis can be stated and then an epidemiologist can determine whether to reject it or not, with some level of confidence, by using statistical methods.
- ❑ Statistics can also be used to measure how strong the pattern

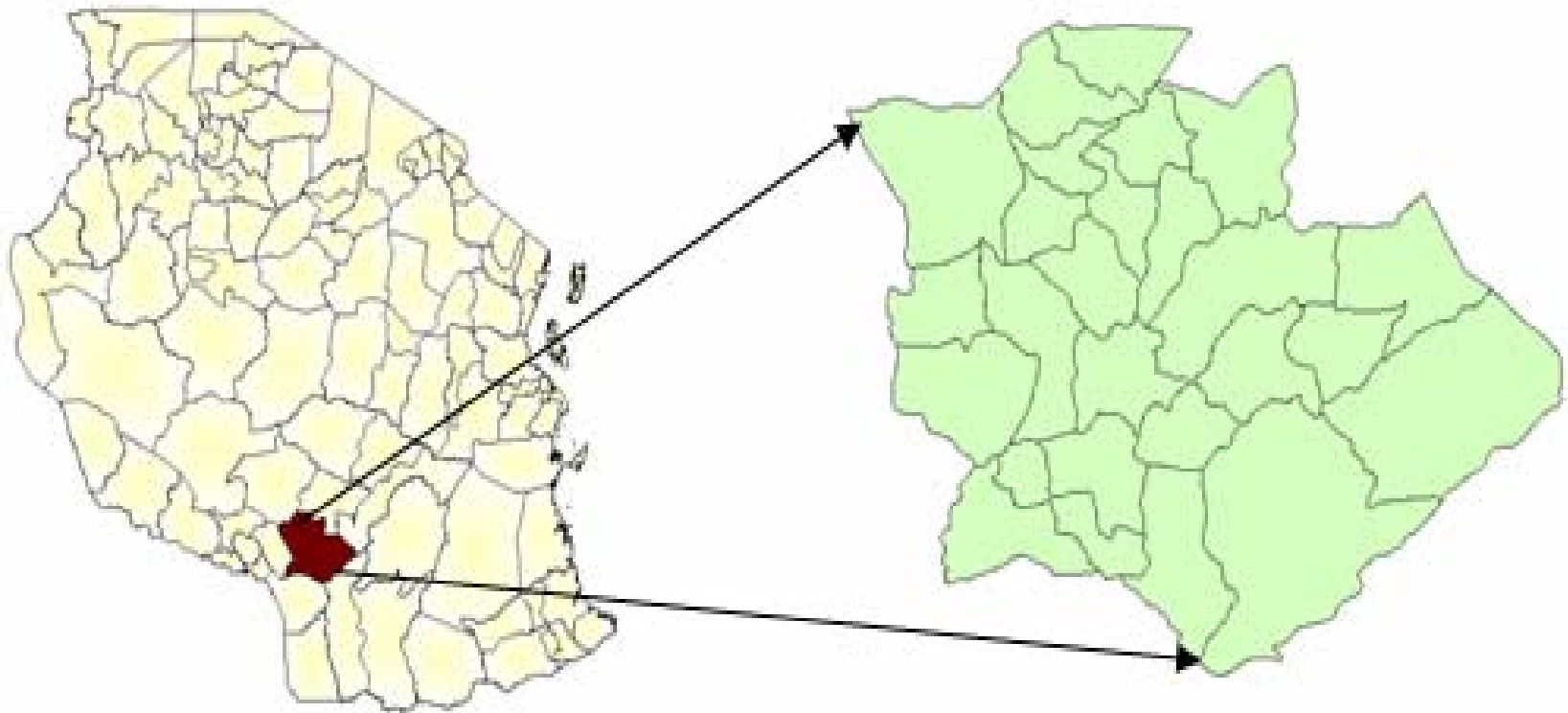
Conclusion Cont...

- ❑ To identify, describe and confirm that a patterns exists; - both methods should complement each other
 - ❑ Visualization is a good technique to start with as it will provide a clue of the patterns, particularly the spatial and spatio-temporal ones, such as areas with high values of malaria, and or months showing similar patterns
 - ❑ Then, statistics is needed to confirm what has already been observed in a visualisation

Thank you

Tasks	Task number	Visualization technique
Spatial	Selected month, and all wards <ul style="list-style-type: none"> • Identify spatial pattern • Describe patterns 	Map
Temporal	Selected one ward, all years <ul style="list-style-type: none"> • Identify temporal patterns • Describe the patterns Select one month and all wards <ul style="list-style-type: none"> • Identify temporal patterns • Describe the patterns 	Map and PCP
Spatio-temporal	Selected 1 year, and selected wards <ul style="list-style-type: none"> • Identify wards with similar curves (patterns) Selected 1 year, and all wards <ul style="list-style-type: none"> • Identify wards that are spatially grouped • Identify if grouped wards display similar • Describe the behaviour of the spatial cluster in time (e.g. the cluster remains for 3 months, then it dissolves quickly; or: the cluster gradually increases and moves towards the North) 	PCP Animation

Study Area



Limitations

- ❑ Skills on statistical methods - not all participants new the statistical methods as those described in chapter 3.
- ❑ Visualization environment of two software's - there was no a communication between the two different software's used, this forced participants to be switching from one software to the other.
- ❑ Data - the data set used for this study was for one district and at ward level only. This did not allow further investigations on factors such as elevation, geographical locations of health facilities in each ward, mosquito breeding sites etc
- ❑ Evaluation - the number of test persons, especially the targeted group (epidemiologists) was limited, only five epidemiologists were within reach during the evaluation period



NIMR

Recommendations

- ❑ Studies should be conducted that will include different visualization methods and statistical methods in one environment
- ❑ The study used data based on malaria cases only, there is need to conduct studies that will incorporate other ancillary data such as elevation, mosquito breeding sites. This will enable users to explore relationships between the number of cases and patterns. A description can be easily given and someone can get an idea of the direction of spread of disease
- ❑ The data used in this study lacks information on health facilities locations and numbers of malaria cases per facility. A further study can be conducted that will include the geographical location of health facilities; these locations can facilitate better identification and description of patterns.
- ❑ Further studies should be conducted, where by statistical methods will be selected prior to the evaluation test, as it was for visualization techniques. The methods can be introduced to the participants for the benefit of those who don't know them and for those who know them but don't use them. This can be done in the familiarization session during the evaluation