Using GIS to identify areas in need of improved access to snakebite treatment

Erik Hansson
Medical student
Lund University, Sweden

Background: Snakebites

- 1 - 5 million bites worldwide each year
  - 20 - 100 000 deaths due to snakebite each year
  - Rural, tropical areas

- 400 - 600 bites per year in Costa Rica
  - 1 - 6 deaths per year due to snakebites in Costa Rica

Snakebite deaths are preventable!

Background: Treatment

- Early intravenous administration of antivenom
- Treatment of complications

Aims

- Detect areas where there is a need of improved access to treatment.
  - Districts with high snakebite incidence (>30 bites/100,000 inhabitants/year)

  Or
  - Environmental risk factors favoring snakebites

  And
  - Long transportation time to treatment (>2-3 hours)

- Describe methods useful for similar studies in other countries.

Identification of high-incidence areas

- For *small* area data, random variation will lead to extreme rates in areas with small populations
- This gives a ”noisy” map that is hard to interpret and has a low accuracy in identifying actual high-incidence areas.

Observed incidence $\sim$ Risk

The observed incidence varies stochastically around the underlying risk.
If the unit of analysis is *small*, the random, stochastic component will become a major part of the variation in the map of observed incidence.

Example "maps"

100	500	1000	2000	1000
2000	100	1000	100	1000
100	1000	1000	1000	500
100	1000	1000	2000	500
1000	500	2000	1000	100

population

1	2	2	4	1
3	1	0	0	4
0	3	1	4	1
0	2	4	5	0
0	2	3	1	0

cases

10	4	2	2	1
1.5	10	0	0	4
0	3	1	4	2
0	2	4	2.5	0
0	4	1.5	1	0

incidence

Empirical Bayesian smoothing techniques

- Smoothing of the observed incidence towards a local or global mean.
- Borrowing of information from neighboring and similar areas to give a more stable estimation of the actual risk underlying the observed incidence.
- Especially useful if area of analysis is *small*.

What is a small area?

- In my opinion, different for different diseases, for example:
  - Snakebites:
    - Strong environmental risk factors
    - We could probably map a relevant spatial pattern also with few observed cases as the variation will to a large extent be determined by environmental factors that vary geographically.
  - Cancer:
    - Mostly weak environmental risk factors (in many cases unknown).
    - The pattern in a map of few cancer cases is likely largely determined by district differences in individual-level risk factor composition and random variation.

Empirical Bayesian smoothing using SIGEpi

- Free software developed by Pan American Health Organisation (PAHO).
- ”Suavizador espacial de tasas”, fully automated tool for spatial smoothing.

Empirical Bayesian smoothing using WinBUGS 1.4.3

- Free software for Monte Carlo sampling.
- Fits a Poisson regression model of the risk in each area with:
  - A Conditional AutoRegressive (CAR) random spatial effect – captures unmeasured spatial processes
  - Rural population percentage
  - Agricultural workforce percentage
  - Forest coverage
  - ”Suitable for Terciopelo”=<1200 m, humid, rural conditions
  - Elevation
  - Number of dry months

Empirical Bayesian smoothing using WinBUGS 1.4.3

- Smoothes towards the mean of the neighboring areas using the CAR function, and can also take into account risk factor composition.
- Can estimate the **probability** that a certain risk- or incidence threshold is exceeded, in this case 30 bites/100,000 inhabitants/year.

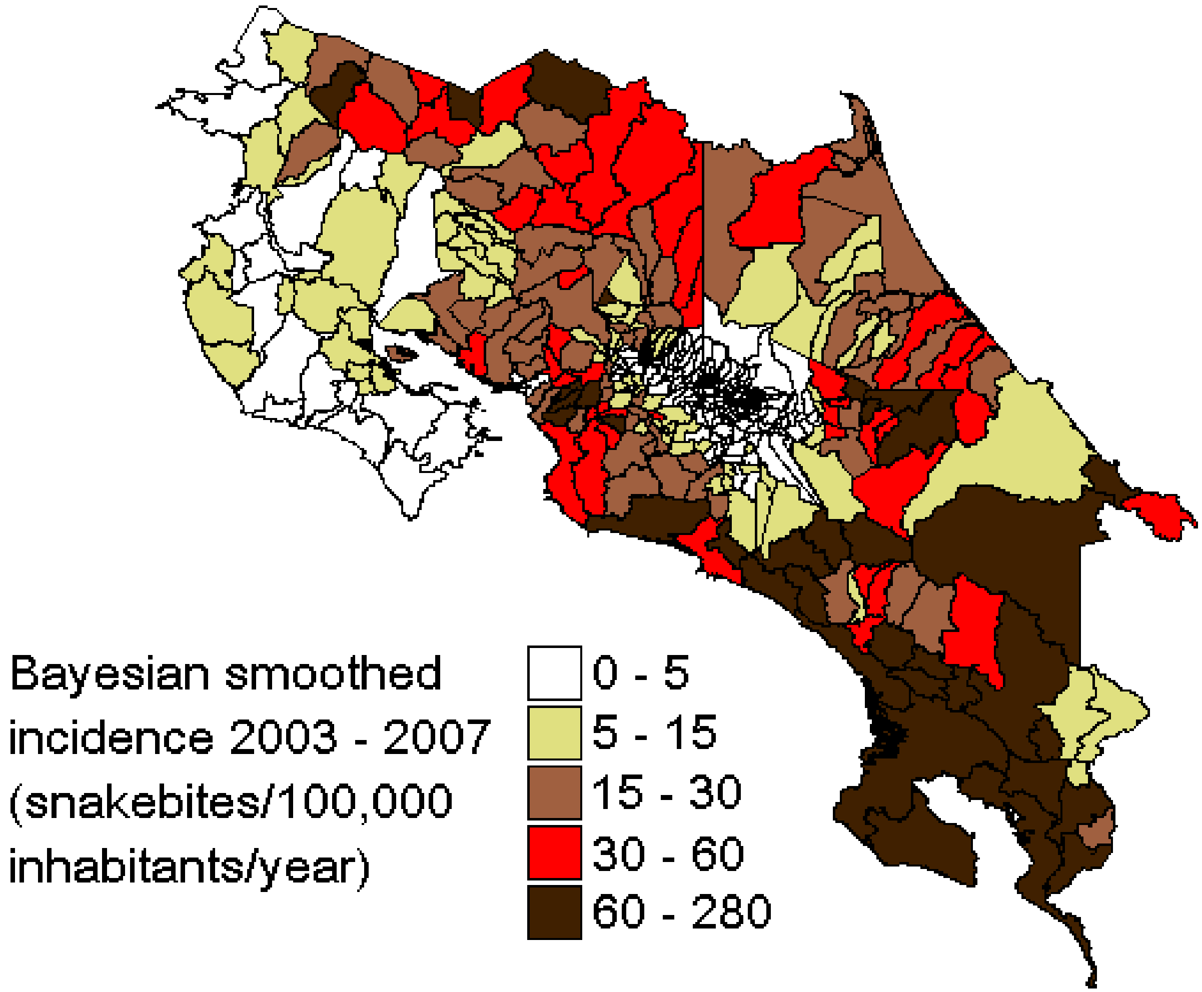
Comparison of smoothing methods

Observed period	Predicted period	Smoothing method	AUC of ROC	95 % C.I. Low	95 % C.I. High	p (better than NS)	p (better than EBS)
90-94	95-99	WinBUGs EBS	0.96	0.95	0.98	0.02*	0.14
90-94	95-99	SIGEpi EBS	0.95	0.93	0.98	0.10	
90-94	95-99	No Smoothing	0.94	0.91	0.97		
94	95-99	WinBUGs EBS	0.95	0.93	0.97	0.00**	0.00**
94	95-99	SIGEpi EBS	0.91	0.88	0.94	0.00**	
94	95-99	No Smoothing	0.82	0.76	0.87		

Comparison of smoothing methods

- More accurate identification of future high risk areas using the smoothed estimates, especially when using the CAR + risk factors and especially for a small data material:
  - Data gathered for 1 year only could have as good accuracy as data gathered during 5 years, if smoothing is used.
- Bayesian smoothing using SIGEpi is also good and more easily implemented.

Map of smoothed incidence

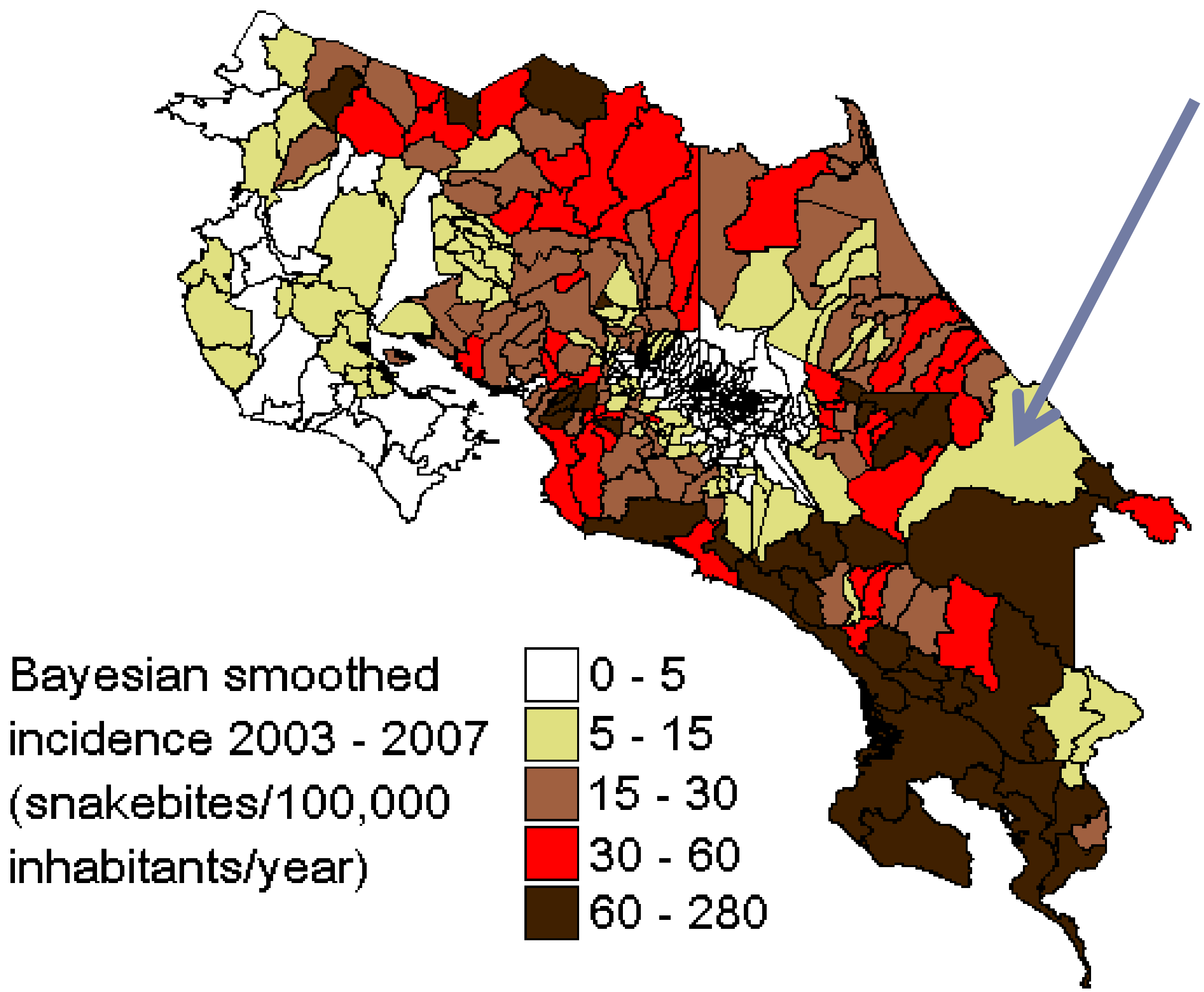


Limitations of using area-level data

- Snakebite risk is assumed equal within districts, even though there likely is big variation on sub-district level.
- For example, a high incidence among rural populations in a district with a large urban population could go unnoticed.

$$Incidence = \frac{number\ of\ bites}{rural\ population + urban\ population}$$

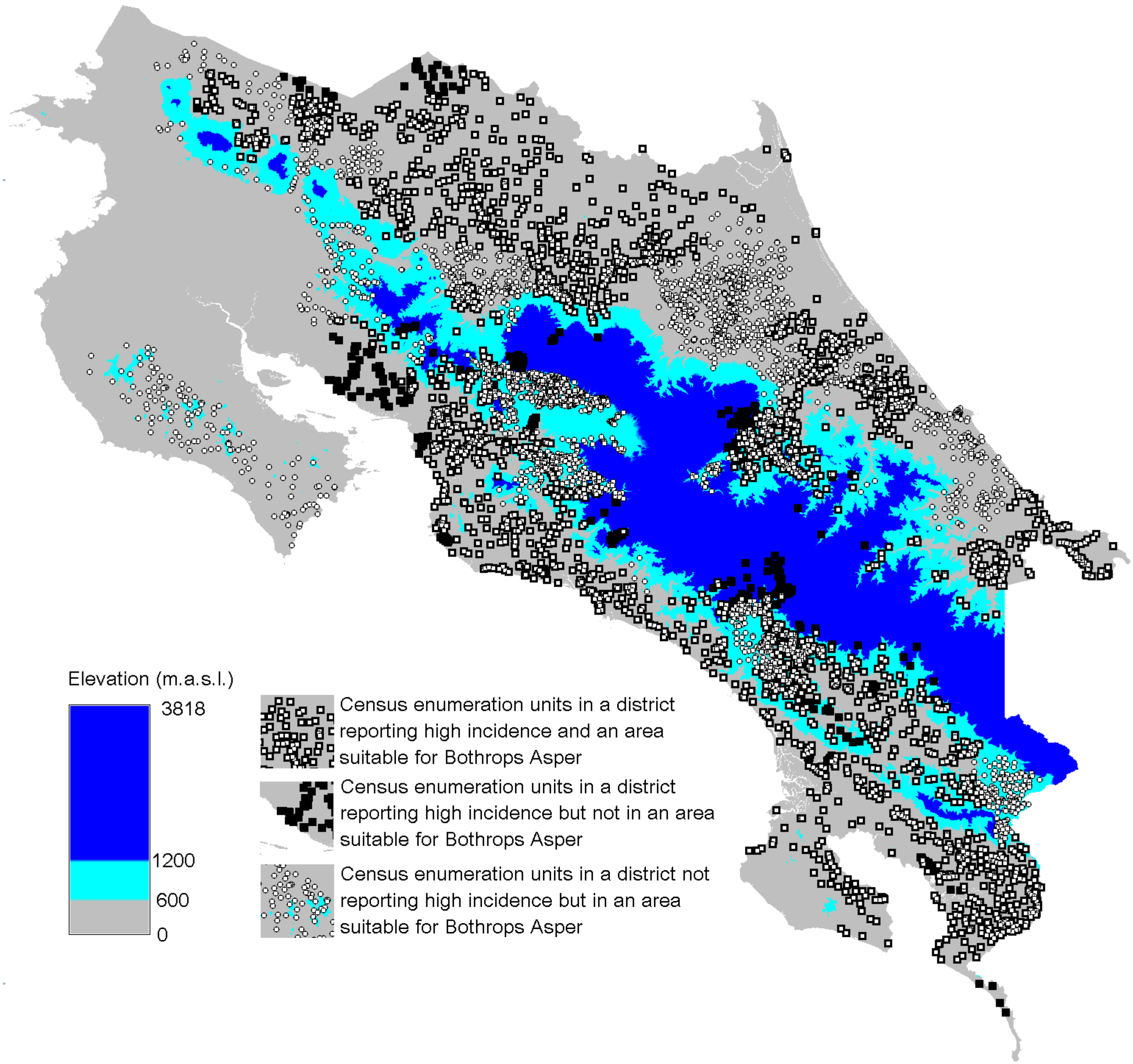
Limitation of area data



Two approaches to identify populations in high risk of snakebites

- Those living in districts reporting a high incidence (>30 per 100,000 inhabitants per year)

- Those living in areas with environmental conditions favoring snakebites:
  - Below 1200 m.a.s.l.
  - Humid conditions
  - Rural

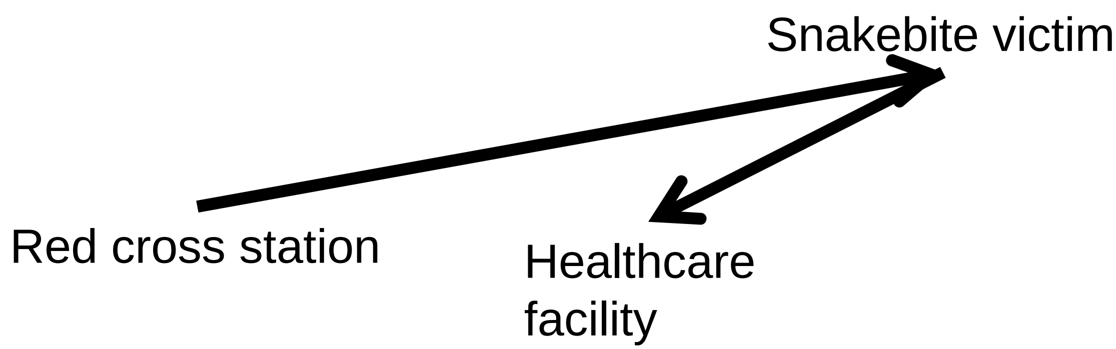

Elevation (m.a.s.l.)
3818
1200
600
0
Census enumeration units in a district reporting high incidence and an area suitable for Bothrops Asper
Census enumeration units in a district reporting high incidence but not in an area suitable for Bothrops Asper
Census enumeration units in a district not reporting high incidence but in an area suitable for Bothrops Asper

Estimating the time to treatment

- Using GRASS 6.4.1. – free GIS software with powerful raster analysis
- Road vector layer
  - Road raster layer, classified according to road type
- Raster elevation layer 30*30 m
  - Slope layer, 250*250 m
- Point vector layer of hospitals, clinics and Red Cross stations.

Assumptions

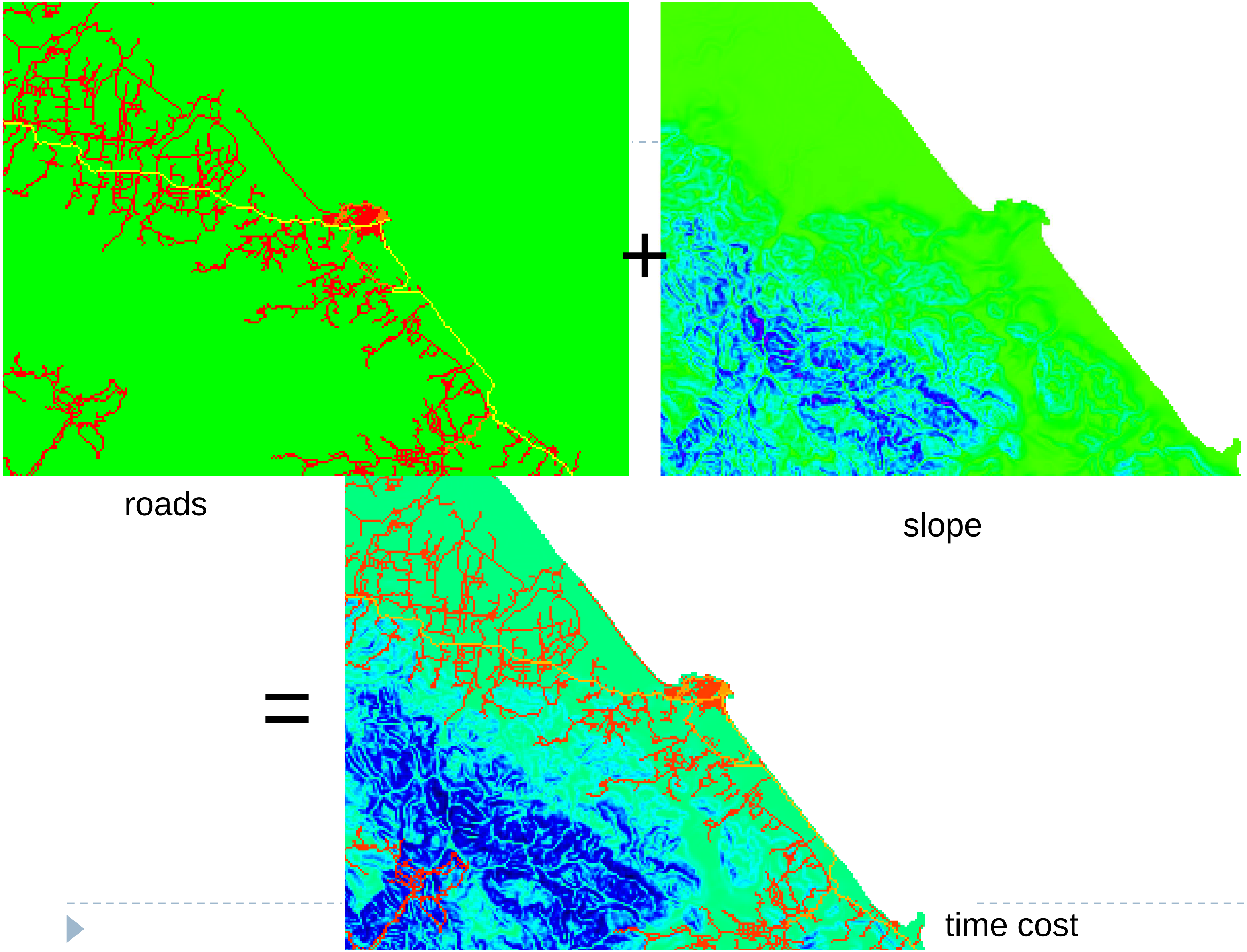
- Many snakebite victims reach healthcare using ambulance, meaning that time to treatment should be estimated as:



- Travel speeds:
  - Primary roads: 60 km/h
  - Secondary roads: 40 km/h
  - Tertiary roads: 20 km/h
  - Off-road: 3 km/h (6km/h in the raster, but will be counted twice!)

Construction of cost layer

- Overlaying road layer with the slope layer using the *r.mapcalc* tool allows adjusting speeds by physical geography
  - Steep slope become barriers to movement


roads
+
slope
=
time cost

Finding the fastest path, not the shortest using the cost layer

2	1	1	
1	3	7	10
1	2	1	1
	1	2	3

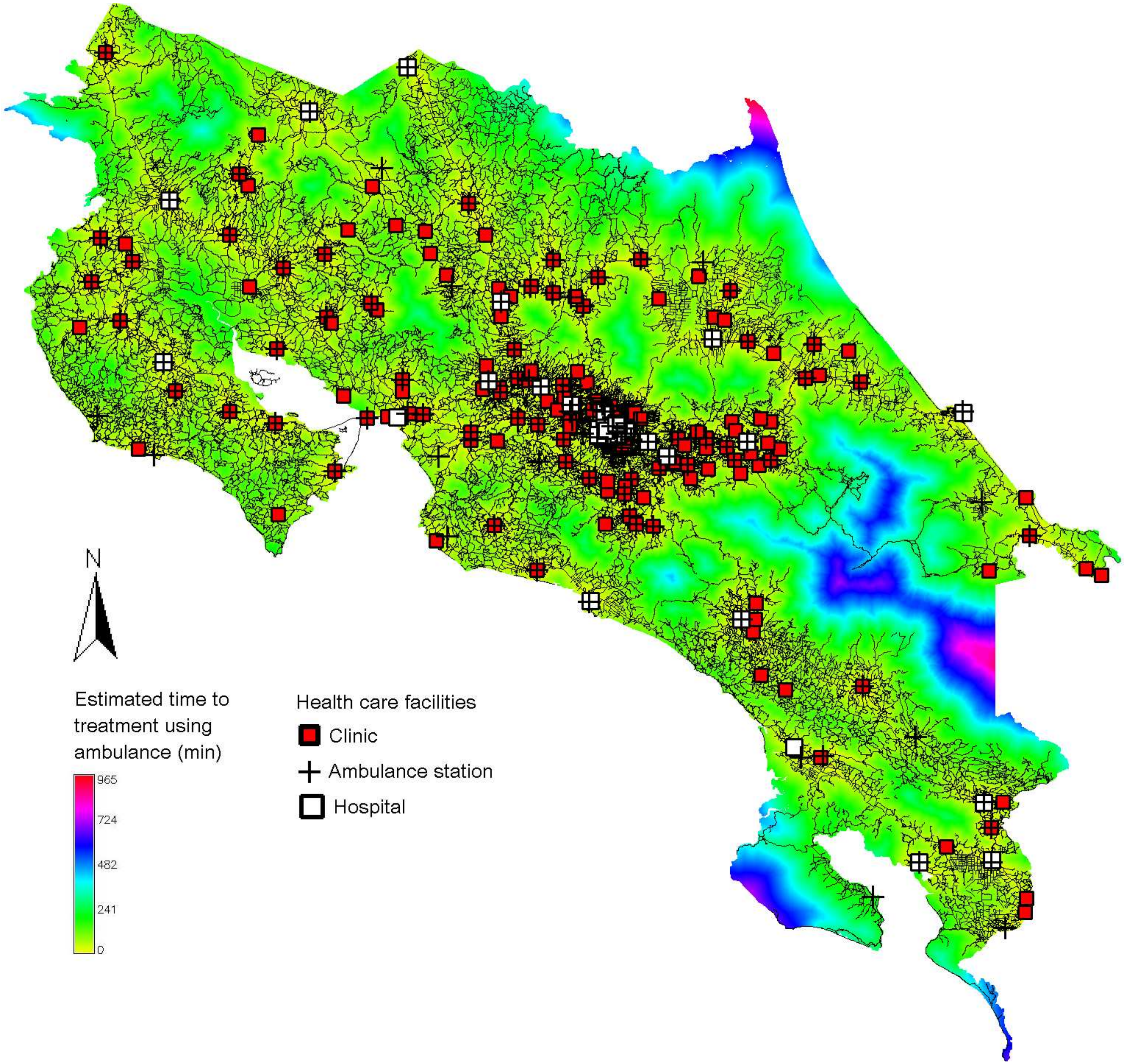
Finding the fastest path, not the shortest using the cost layer

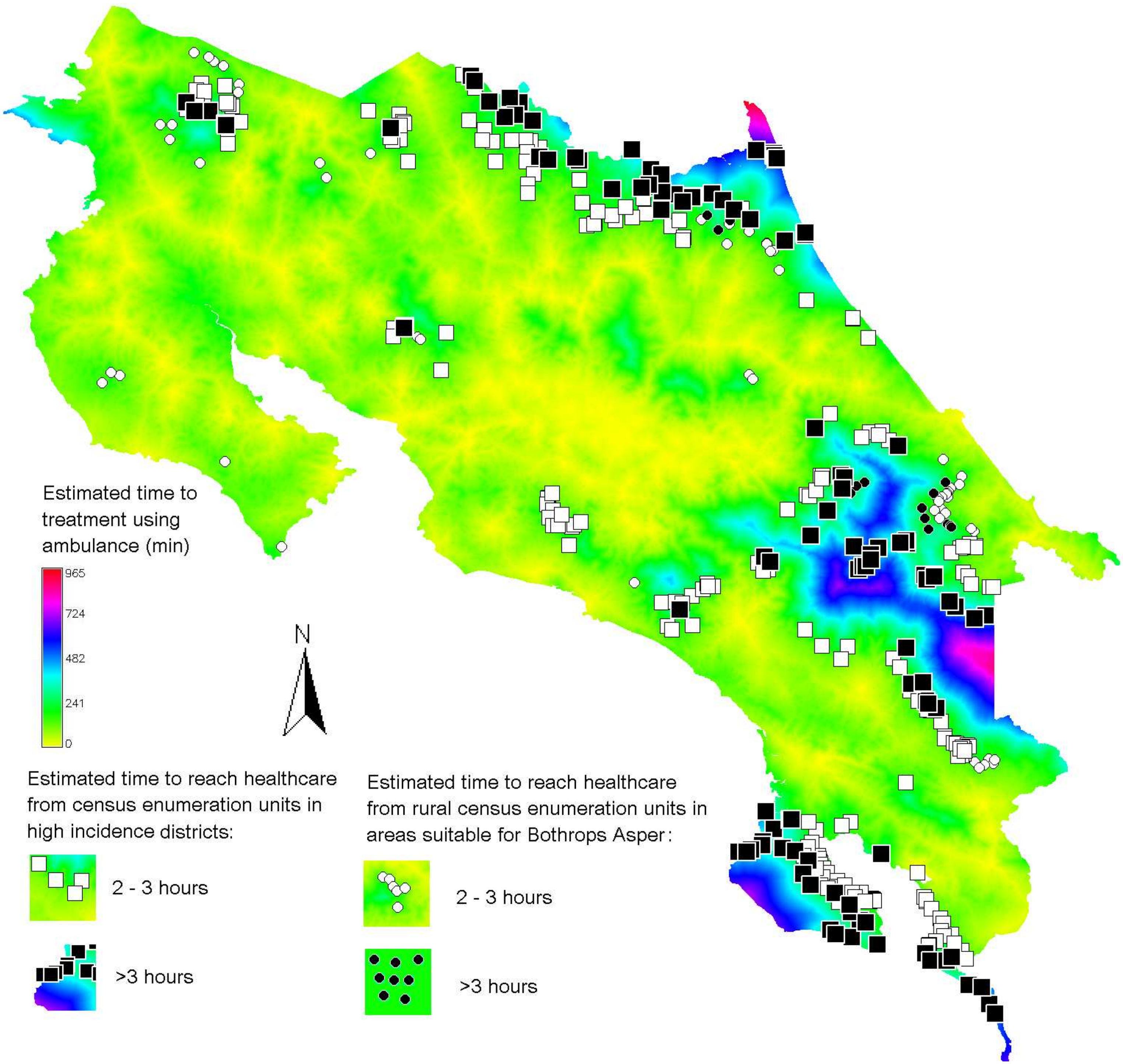
2	1	1	
1	3	7	10
1	2	1	1
1		2	3

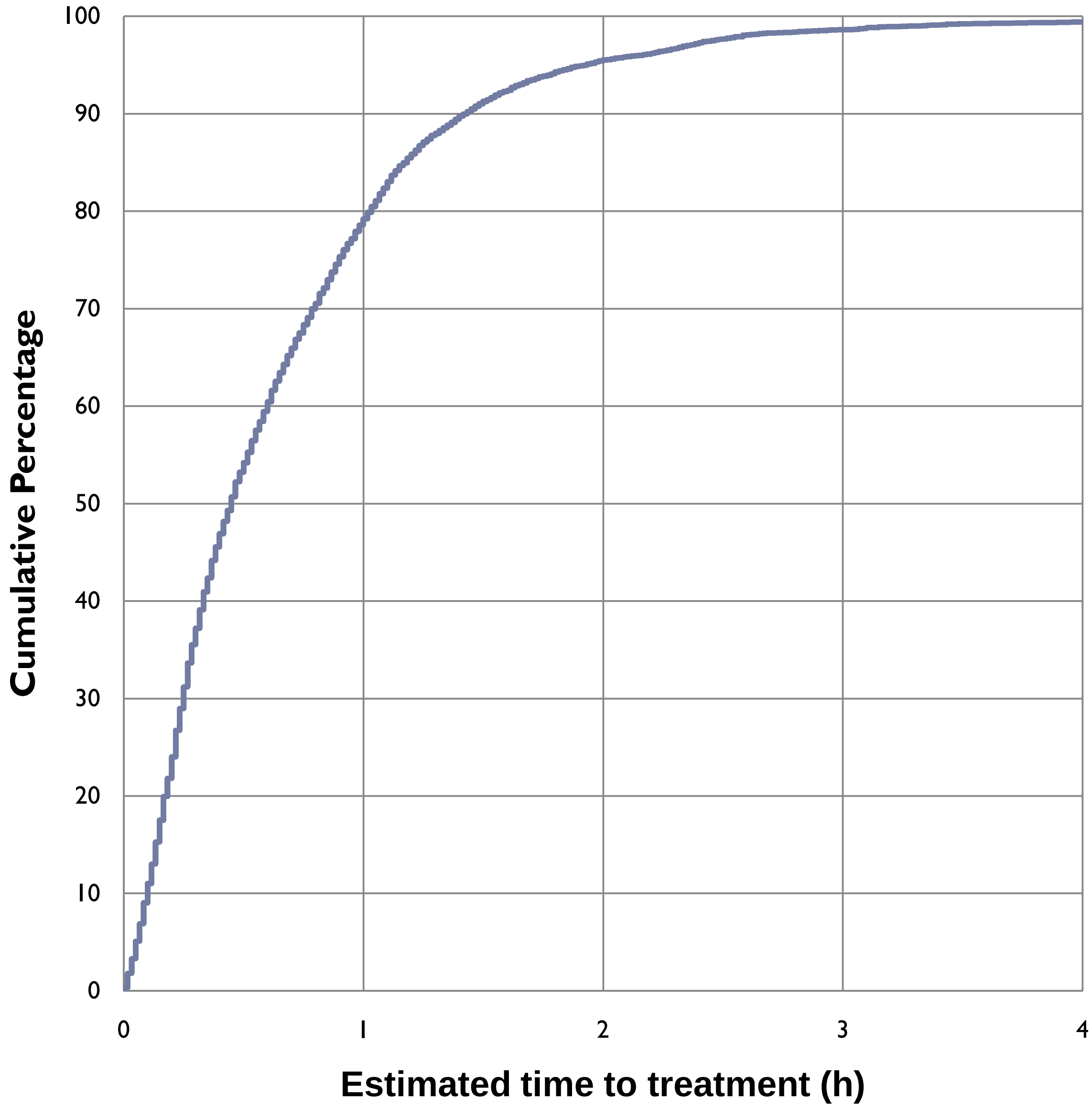
Construction of final time to treatment raster

Time from ambulance station to place of residency + Time from place of residency to hospital or clinic

= Total time to reach healthcare


N
Estimated time to treatment using ambulance (min)
965
724
482
241
0
Health care facilities
Clinic
Ambulance station
Hospital


Estimated time to treatment using ambulance (min)
965
724
482
241
0
N
Estimated time to reach healthcare from census enumeration units in high incidence districts:
2 - 3 hours
>3 hours
Estimated time to reach healthcare from rural census enumeration units in areas suitable for Bothrops Asper:
2 - 3 hours
>3 hours



Comparison with observed times

The median time to Limon hospital was 3 hours, the mean 6.8 hours

– Saborio et al. 1997

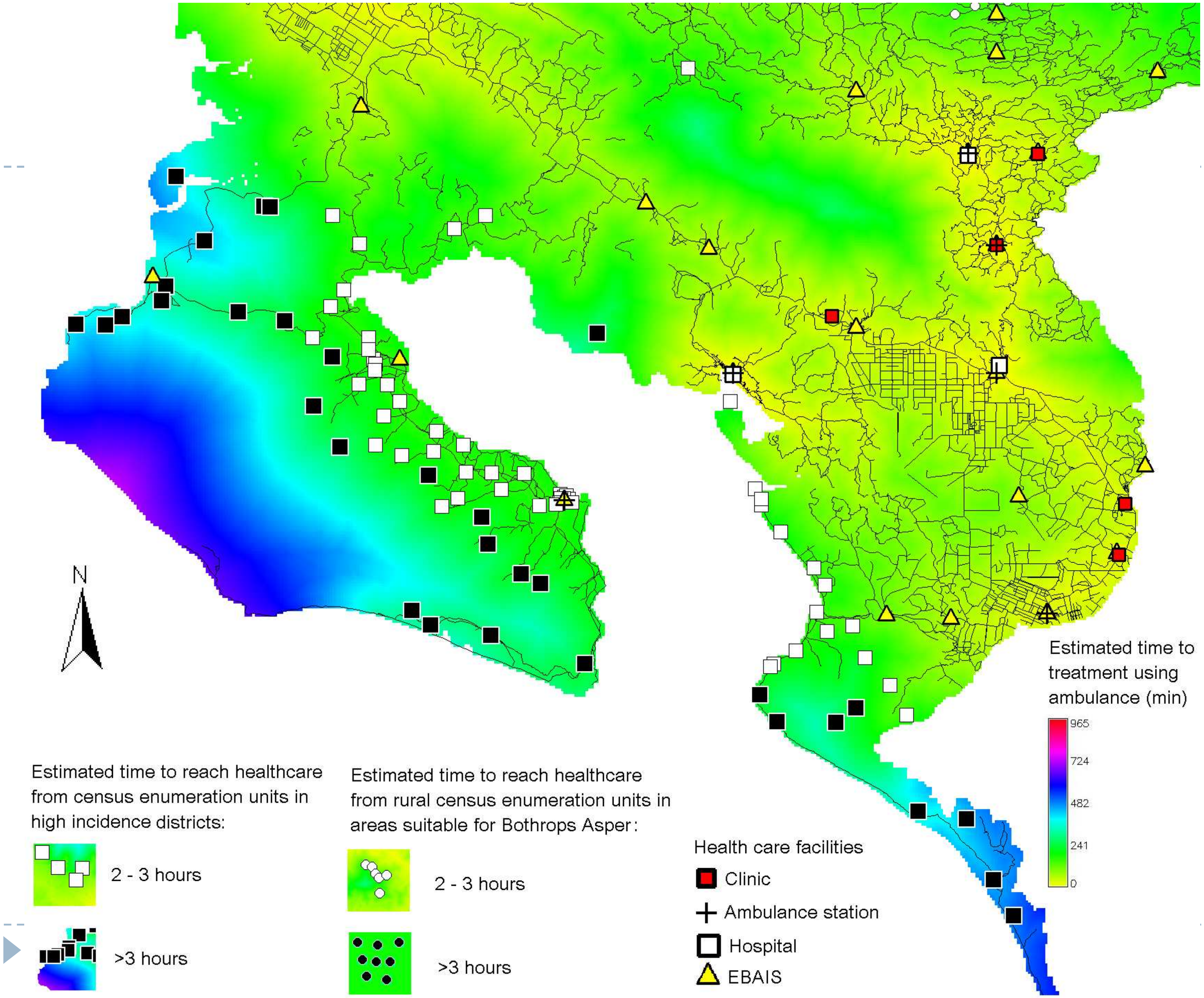
61% reached hospital (any) within 3 hours, 20% after more than 5 hours

– Arroyo et al. 1999

Very much longer than estimated times… why?

Why this discrepancy between estimated and observed times?

- Not fully comparable measurements
- Speeds set too high
- Non-spatial aspects
  - The estimated time should be considered an ”ideal” time, with only the spatial dimension of accessibility


N
Estimated time to treatment using ambulance (min)
965
724
482
241
0
Estimated time to reach healthcare from census enumeration units in high incidence districts:
2 - 3 hours
>3 hours
Estimated time to reach healthcare from rural census enumeration units in areas suitable for Bothrops Asper:
2 - 3 hours
>3 hours
Health care facilities
Clinic
Ambulance station
Hospital
EBAIS

Interpretation

- The map identify some areas with likely need of improved treatment access, but
- cannot be the only basis of decision on distribution of antivenom
  - Many possible sources of error in the data
  - Sensitive to assumptions
- Combined with the knowledge of local health care officials, it might provide a useful tool for improving access to treatment.

Conclusion

- GIS is a good tool for
  - improving interpretation of spatial data, e.g. by smoothing incidence in small areas,
  - identifying areas with environmental risk factors,
  - visualising geographical information and relationships, e.g. location of health care facilities vs. population at risk of snakebites,
  - estimating health care accessibility, e.g. the time to reach a clinic or hospital using ambulance
- but;
  - Sensitive to data errors and assumptions.
  - Risk of trusting maps too much.

Questions

- Now
- Or later to erik.hansson@med.lu.se