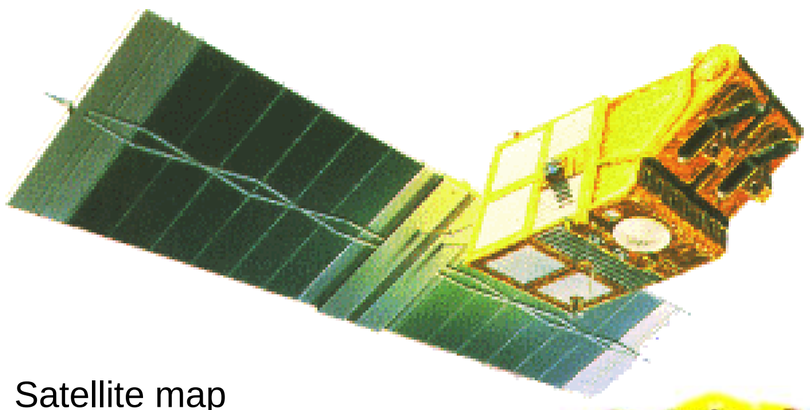
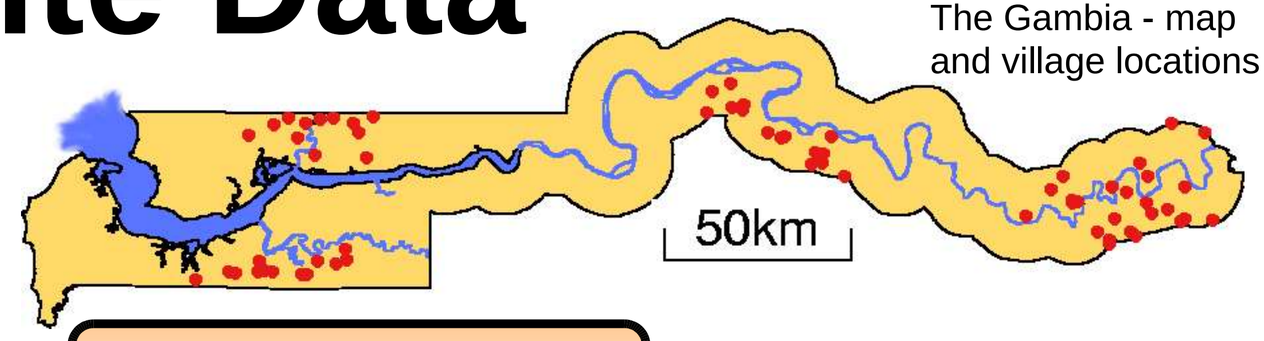
Modelling Malaria with Spatial Statistics and Satellite Data



Satellite map of 'greenness'

The Gambia - map and village locations



Satellite Data

The NDVI measurements for the mosquito season were used to classify The Gambia into ten regions using a clustering algorithm. These averaged values were then put into the model as a measure of the local greenness.

Survey Data

Children in a sample of villages throughout The Gambia were surveyed during the malaria season. They were tested for malaria parasites and data on the individual children and villages were collected. For the child, the age, sex, an bednet usage was noted, and for the village we record the location and the Public Health Centre (PHC) status.

Statistical Model

We use a standard Generalised Linear Model to model the presence of parasites as a function of the covariates of the child and the village. At this stage most of the parameters look significantly different from zero. However since the observations may be correlated due to spatial effects this significance may be overestimated.

Spatial Correction

To correct for the correlation we first quantify the correlation by fitting a variogram to the residuals from the model. From this we can compute new standard errors that consider the effect of the correlated observations.

Parameters

This table gives the parameter values and standard errors before and after the spatial correlation correction. The t-statistic is used to measure the significance of the parameter away from zero. Note how t' changes after correction, to the extent of making PHC insignificant.

Name	***Value***	***SE*** *before correction*	***t***	***SE'*** *after correction*	***t'***
NDVI	-0.40	0.09	-4.41	0.19	-2.14
NDVI-squared	0.42	0.09	4.51	0.19	2.17
UntreatedNet	-0.55	0.13	-4.18	0.20	-2.74
TreatedNet	-0.68	0.15	-4.57	0.25	-2.75
AGE	0.59	0.12	5.00	0.12	5.10
PHC	-0.43	0.12	-3.57	0.24	-1.78
significance:	p=0.01		p=0.05		

Predictions

From the model we can compute the malaria prevalence under different scenarios. For example we can consider the effects of changing bednet usage in the population:

Change	**Number of cases**
No change	728.0
Treat untreated bednets	706.0
Give non-bednet users untreated bednets	656.8
Everyone with treated nets	619.7

Conclusion

We have modelled malaria prevalence using satellite data as a covariate, and allowing for spatial correlation. Useful predictions can be made from the model.

LANCASTER UNIVERSITY

Barry Rowlingson & Peter Diggle
Dept of Mathematics and Statistics

In collaboration with: **Madeleine Thomson** ***MALSAT***
Liverpool School of Tropical Medicine