

**Investigating the spatial and temporal epidemiology
of *Schistosoma* species infection within school-aged-children
along the southern shoreline of Lake Malawi**

by

Amber Lydia Reed



A thesis submitted in partial fulfilment for the degree of Doctor of Philosophy

in the

Faculty of Health and Medicine

Lancaster Medical School

November 2024

Contents

Declaration	ii
Abstract	iii
Acknowledgements	iv
Computer code and data	iv
List of Figures	iv
List of Tables	xiv
List of Papers	xvi
List of Abbreviations and acroyms	xvii
1 Introduction	1
1.1 Thesis overview	1
1.2 Schistosomiasis	1
1.2.1 Epidemiology overview	1
1.2.2 Transmission and lifecycle	4
1.2.3 Environment and human behaviour factors	7
1.2.4 Pathogenesis and clinical manifestations	8
1.2.4.1 Acute disease	8
1.2.4.2 Established active and chronic infections	9
1.2.5 Urogenital schistosomiasis	9
1.2.6 Intestinal schistosomiasis	10

1.2.7	Co-infection	10
1.2.8	Immunity	11
1.2.9	Diagnosis	12
1.2.9.1	Direct parasitological methods	14
1.2.9.2	Indirect diagnostic methods	16
1.2.10	Treatment and preventive chemotherapy	19
1.3	Control of schistosomiasis	20
1.3.0.1	Control to target the parasite	20
1.3.0.2	Control of the snail populations	23
1.4	Study population: Lake Malawi and the Mangochi District	25
1.4.1	Geographic and demographic context	25
1.4.2	Urogenital Schistosomiasis in Mangochi District	26
1.4.3	Changing epidemiological landscape of schistosomiasis in the Mangochi District	30
1.5	Snails and their microhabitat	30
1.6	Methods overview	32
1.6.1	Modelling and statistical techniques	33
1.6.2	A basic overview of Generalised linear models/Generalised additive model/Bayesian linear models	34
1.6.3	A basic overview of deterministic and stochastic SIR/SEIR/SEIRS-type compartmental model	36
1.6.4	Background of previous schistosomiasis modelling and how it relates to the thesis	37
1.7	Aims and structure of this thesis	39
1.8	Objectives	40
2	Modelling the age-prevalence relationship in schistosomiasis: A secondary data analysis of school-aged-children in Mangochi District, Lake Malawi	42

2.1	Introduction	43
2.2	Methods	46
2.2.1	Dataset	46
2.2.2	Statistical analysis	46
2.3	Data methods	48
2.4	Results	49
2.4.1	Prevalence heatmaps	49
2.4.2	Generalised additive models	49
2.5	Discussion	50
2.6	Conclusion	57
2.7	Declaration of Competing Interest	57
2.8	Data availability	57
2.9	Acknowledgments	57
3	A geospatial analysis of local intermediate snail host distributions provides insight into schistosomiasis risk within under-sampled areas of southern Lake Malawi	58
3.1	Introduction	59
3.2	Methods	61
3.2.1	Snail abundance	61
3.2.2	Remote sensing data	63
3.2.3	Construction of 200 prediction points along the shoreline	63
3.2.4	Extraction of remote sensing data to linestring vertices	65
3.2.5	One dimensional Poisson latent Gaussian process regression	65
3.2.5.1	Bayesian Multilevel Model	65
3.2.6	Model Formulation	65
3.2.6.1	Model	65
3.2.6.2	GP prediction	66

3.2.6.3	Latent Gaussian process	67
3.3	Results	67
3.3.1	Observed data	67
3.3.2	Environmental data prediction points	67
3.3.3	Covariance function comparison	70
3.3.4	Model fit	70
3.3.5	Covariate effects	72
3.3.6	Model predictions	73
3.4	Discussion	75
3.5	Conclusion	80
4	Development of a dynamical model to enhance understanding of epidemiology of schistosomiasis in school aged children	81
4.1	Introduction	83
4.2	Methods	85
4.2.1	Dataset	85
4.2.2	Disease model formulation	85
4.2.3	Observation of prevalence	89
4.2.4	Inference	89
4.2.5	Sensitivity analysis	93
4.3	Results	93
4.3.1	Check best value of ϕ	93
4.3.1.1	Snail abundance	93
4.3.2	Model optimisation	93
4.3.3	CI Intervals	106
4.3.4	Sensitivity analysis	106
4.4	Discussion	108

4.4.1	Limitations	109
4.5	Conclusion	111
5	General Discussion	112
5.1	Chapter overviews	112
5.2	Conclusion of this thesis	115
	References	119
	Appendices	144
A	Ethics	145
B	Chapter 2	147
B.1	Raw data	147
B.2	Prevalence heatmaps	150
B.3	Generalised additive models	150
B.4	Water contact	156
C	Chapter 3	158
C.1	Construction of 200 prediction points	158
C.2	1D extracted environmental data	158
C.3	Center and scaling environment data	164
C.4	Covariance functions	167
C.5	Convergence	170
C.6	Priors and Posteriors	170
C.7	1D result	172
C.8	Bathymetric data	174
D	Chapter 4	176

D.1	Sensitivity analysis	176
D.1.1	plots	176
D.1.2	Tables	189
D.1.3	Confidence intervals	189
D.2	Log-likelihood profiles	207
E	Dissemination of research	213

Declaration

This thesis has not been submitted in support of an application for another degree or other university. It is the result of my own work and includes nothing that is the outcome of work done in collaboration except where specifically indicated. Many of the ideas in this thesis were the product of discussion with my supervisors, Professor Christopher Jewell, Dr Claudio Fronterre, Professor Russell Stothard and Dr Michelle Stanton.

Chapter 2 of this thesis has been published in the following academic publication:

Amber L. Reed, Angus M. O’Ferrall, Sekeleghe A. Kayuni, Hamish Baxter, Michelle C. Stanton, J. Russell Stothard, Christopher Jewell, Modelling the age-prevalence relationship in schistosomiasis: A secondary data analysis of school-aged-children in Mangochi District, Lake Malawi, *Parasite Epidemiology and Control*, Volume 22, 2023, e00303, ISSN 2405-6731, <https://doi.org/10.1016/j.parepi.2023.e00303>.

Chapter 3 of this thesis has been published in the following academic publication:

Amber L. Reed, Mohammad H. Al-Harbi, Makaula, Peter Makaula, Charlotte Condemine, Josie Hesketh, John Archer, Sam Jones, Sekeleghe A. Kayuni, Janelisa Musaya, Michelle C. Stanton, J. Russell Stothard, A geospatial analysis of local intermediate snail host distributions provides insight into schistosomiasis risk within under-sampled areas of southern Lake Malawi, *Parasites and Vectors*, Volume 17, page. 272, 2024, <https://doi.org/10.1186/s13071-024-06353-y>.

Amber L. Reed, BSc, MSc, MRes

Lancaster University, UK

Abstract

Schistosomiasis (Bilharzia) is a focal water-borne neglected tropical disease (NTD) caused by trematodes of the genus *Schistosoma*. In Africa schistosomes utilise freshwater intermediate snail host species of *Bulinus* and *Biomphalaria*. As with many NTDs, fine-scale data on schistosomiasis is sparse, needing application of advanced quantitative methods to better infer spatial distributions and demographic associations. This thesis applies quantitative methods to interpolate available epidemiological data from Lake Malawi to explore in greater detail the spatial and temporal epidemiology of intestinal schistosomiasis (IS) and urogenital schistosomiasis (UGS), and their snail hosts. Through applying spatial and dynamical modelling methods to a newly emerging focus of IS, concurrent within an area for UGS, peak age infection prevalence was shown to be 11 years of age for IS and IS/UGS co-infection, considerably younger than that previously reported. Using remote sensing, geostatistical analyses provided insight into snail species abundance along the shoreline, noting environmental associations, revealing substantial heterogeneities but identifying snail hotspots. To better understand transmission by replicating the previous age profiles, an age-structured SEIRS (Susceptible-Exposed-Infectious-Recovered-Susceptible) transmission model was constructed; with age-related immunity and the effect of exposure to snail populations considered, followed by an optimisation of the model parameters. From this, “snail exposure” was judged not to be simply due to proximity to “snail habitat”. These findings add quantitative insight to the Lake Malawi shoreline setting and may later support World Health Organisation guideline development in criteria for interruption of schistosomiasis transmission. Future work should consider information from longitudinal cohorts, rather than cross-sectional studies, when attempting to model and quantify fine-scale transmission heterogeneities and putative impacts of current control interventions.

Acknowledgements

I would like to thank my supervisors, Professor Christopher Jewell, Dr Claudio Fronterre, Professor Russell Stothard and Dr Michelle Stanton for their support throughout this thesis. They have provided many online Teams and in person meetings, with many hours of their time including short notice meetings, extended meetings and plenty of positive encouragement. I would like to especially thank Professor Christopher Jewell for his continuous support throughout the thesis. He has supported me in every part of the thesis, analysis, coding and writing. I would like to thank Dr Claudio Fronterre and Dr Michelle Stanton for their specialist help on geospatial aspects of the thesis, including coding and writing up. Also, I would like to thank Professor Russell Stothard for his guidance and support on the study of schistosomiasis, writing up of the thesis and the many fish and chips lunches.

Thank you to my mother, Alison Reed, for supporting me while working on this thesis despite long periods away from home with many long phone calls which I am entirely grateful for and never letting me give up despite how hard times felt. I would also like to mention, my late grandad Peter Rodney Reed, who always since I was born gave his advice, like-minded ideas and supported me, which in turn provided me with the determination and drive to pursue this thesis. I would like to thank the rest of my family including my sisters, grandma, aunties and uncle for their support throughout my PhD. I would like to mention my best friend of 27 years, Faye Harrison, who has been there throughout my PhD and most of my life. In addition, I am extremely grateful for my partner Joseph Powell for supporting me throughout this thesis, providing many well-earned rest breaks, encouragement, cooked meals and phone calls.

In addition, I am grateful for the funding provided by the Medical Research Council.

Computer code and data

All code written in this thesis is accessible on Zenodo: [Chapter 2: DOI: 10.5281/zenodo.10455702](https://doi.org/10.5281/zenodo.10455702), [Chapter 3 DOI: 10.5281/zenodo.10410622](https://doi.org/10.5281/zenodo.10410622) [Chapter 4 DOI: 10.5281/zenodo.10424064](https://doi.org/10.5281/zenodo.10424064)

List of Figures

1.1	The geographical distribution of schistosomiasis worldwide, reproduced from Weerakoon et al. under a CC-BY license [12].	3
1.2	Lifecycle of schistosomiasis, reproduced from CDC <i>et al.</i> under a CC-BY license [15].	5
1.3	Schematic diagrams showing the different diagnostic methods available for the diagnosis of <i>S. mansoni</i> and <i>S. haematobium</i> , reproduced from Christiansen, 2018 under a CC-BY license [37]	14
1.4	A map of Malawi and its surrounding countries. Northern, Central and Southern regions of Malawi shown on map, reproduced from United Nations, 2014 under a CC-BY license [74]	24
1.5	The average district prevalence (%) of <i>S. haematobium</i> from survey data collected in Malawi. Dots colours stand for prevalence (%): Blue (< 1.0), Yellow (1.0–9.9), Orange (10.0 – 19.9), Red (20.0 – 49.9), Brown (> 50.0). Red circle is approximate location of Mangochi district. Reproduced from Global Atlas of helminth infections, 2018 under a CC-BY license [80].	28
1.6	The average district prevalence (%) of <i>S. mansoni</i> from survey data collected in Malawi. Dots colours stand for prevalence (%): Blue (< 1.0), Yellow (1.0 – 9.9), Orange (10.0 – 19.9), Red (20.0 – 49.9), Brown (> 50.0). Red circle is approximate location of Mangochi district. Reproduced from Global Atlas of helminth infections, 2018 under a CC-BY license [80].	29
1.7	A diagram of the process of policy makers and modellers working together to implement improved control programmes in practice, reproduced from Behrend <i>et al.</i> , 2020 under a CC-BY license [92]	33

1.8	A compartment model representing the stages of infection; Susceptible (S), Infection (I), Recovered (R). μ represents the birth rate, λ represents the rate from S to I and γ represents the rate from I to R.	36
1.9	Compartment model reproduced from Kanyi <i>et al.</i> under a CC-BY license [111].	39
2.1	Locations of the schools sampled in the primary study, a) red markers represent a repeat of the previous collection (80 SAC sampled), green markers represent collections newly known to <i>Biomphalaria</i> intermediate host locations (60 SAC sampled) and yellow markers represent rapid mapping of the shoreline (30 SAC sampled), b) map indicating the location of Mangochi District.	45
2.2	Heatmap showing the age of the children vs school prevalence for a) <i>S. mansoni</i> [T+] b) <i>S. haematobium</i> c) co-infection [T+]. Order of schools on heatmap was by highest to lowest prevalence ranking for <i>S. mansoni</i> and showed that there was considerable heterogeneity between the schools. Further, <i>S. haematobium</i> shows a similar pattern of prevalence among SAC to co-infection.	51
2.3	Thin plate spline functions of the log odds ratio of <i>Schistosoma</i> infection by age in SAC for (a) <i>S. mansoni</i> [T+] (b) <i>S. haematobium</i> (c) coinfection [T+]. A general trend towards SAC between 9 and 12 years old having the highest odds of infection was seen in all cases, though that for UGS is not statistically significant.	53
2.4	Smoothed age-specific prevalence of <i>Schistosoma</i> association with age of SAC for each school. a) <i>S. mansoni</i> [T+] b) <i>S. haematobium</i> C) co- infection [T+]. Light Green: Chikomwe, Yellow: Chipeleka, Dark Blue: Koche, Purple: Makumba, Orange: Mchoka, Brown: MOET, Red: Mtengeza, Black: Ndembo, Light Blue: Samama, Pink: St Augustine 2, Dark Green: St Martins, Mauve: Sungusya.	54
3.1	Primary dataset collected data a) Map of Malawi in dark blue. Red crossed: study area; black line: prediction points. Parasitological surveys: b) Primary school locations along the shoreline. Malacological surveys: c) observed <i>Biomphalaria</i> snails; d) observed <i>Bulinus</i> snails.	62
3.2	Raster plot of extracted covariate data. a) Rainfall (mm), b) Daytime LST ($^{\circ}C$), c) Evapotranspiration d) NDVI. e) Soil types for southern part of Lake Malawi shoreline, adapted from Dijkshoorn <i>et al.</i> 2016 [157]. Black line shows the shoreline template where covariate values were extracted from.	64

3.3	Scatter plot of absolute snails numbers observed at sampling points versus distance along the shoreline in km. a) <i>Biomphalaria</i> sp; b) <i>Bulinus</i> spp..	68
3.4	Environmental data values extracted for each prediction point. a) LST($^{\circ}C$) b) rainfall (mm) c) evapotranspiration d) Normalised Difference Vegetation Index (NDVI) e) Luvisolic (LV) f) Planosolic (PL). e and f are compared with Gleysolic(GL) soil type. Gap in shoreline is due to the removal of CM soil type.	69
3.5	Trace plot of Bayesian log-linear Gaussian Process fitted model in Stan a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	71
3.6	Posterior plot for each species a) <i>Biomphalaria</i> sp.; b) <i>Bulinus</i> spp. Red shaded area represents the 80% credible intervals (CrI) and the extent of the curve is the 95% CrI.	72
3.7	2D mean GP prediction of number of snails $\log(\hat{\mu}_i)$ along the shoreline (km) a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp. Legend: Blue to red stands for exponential of mean GP number of snails. Dot size represents the standard deviation of the posterior predictive distribution at each vertex.	74
3.8	A collection of location photographs representative of the variation of the Southern part of Lake Malawi. Pictures taken during field work studies carried out during August 2022 showing the east side of southern part of the Lake Malawi shoreline (unpublished). Pictures taken by Alexandra Juhasz.	76
4.1	SEIRS transmission compartment model with SAC age ranging from 6 to 15. Transmission parameters are discussed in the text.	86
4.2	Logistic age-specific loss of immunity function, $1 - \rho(\text{age})$ versus age. For $\kappa = 0.5 \text{ days}^{-1}$, $C = 11 \text{ years}$, age ranging between 6 and 15 years.	91
4.3	Distance (10km) of school from lake shoreline versus snail abundance distance. Black line: $\phi = 4.48km$, red line: $\phi = 10^6$	94
4.4	Multi- β_s with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	96
4.5	Single- β with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	97

4.6	Multi- β_s with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black) against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	98
4.7	Single- β with no space effect model optimisation prevalence prediction, $\hat{\pi}_{as}$ (black line) and observed prevalence (black) against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	99
4.8	Confidence intervals for parameter estimates for <i>Biomphalaria</i> sp. models with SAC prevalence at age 6 set as α set to zero a) Multi- β_s space, <i>Biomphalaria</i> sp. b) Multi- β_s no space, <i>Biomphalaria</i> sp. c) Single- β space d) Single- β no space . . .	104
4.9	Confidence intervals for parameter estimates for <i>Bulinus</i> spp. models with SAC prevalence at age 6 set as α set to zero a) Multi- β_s space b) Multi- β_s no space c) Single- β space d) Single- β no space	105
4.10	Contour plot showing non-identifiability issue between γ and β parameters	107
A.1	Lancaster University Ethics approval 2020	146
B.1	Raw data plot showing the age of the children vs school prevalence for a) <i>S. mansoni</i> [T+], b) co-infection [T+] and c) <i>S. haematobium</i> . Order of schools on heatmap was by highest to lowest prevalence.	148
B.2	Raw data plot showing the age of the children vs school prevalence for a) <i>S. mansoni</i> [T-], b) co-infection [T-] and c) Order of schools on heatmap was by highest to lowest prevalence.	149
B.3	Heatmap showing the age of the children vs school prevalence for a) <i>S. mansoni</i> [T-] and b) co-infection [T-]. Order of schools on heatmap was by highest to lowest prevalence.	150
B.4	Smooth age term plot for the GAM of <i>Schistosoma</i> association with age of SAC for a) <i>S. mansoni</i> [T-] and b) co-infection [T-].	152
B.5	Gam of <i>Schistosoma</i> association with age of SAC for each school. Invlogit of predicted fitted values versus age, a) <i>S. mansoni</i> [T-] and b) co-infection. Light Green: Chikomwe, Yellow: Chipelekera, Dark Blue: Koche, Purple: Makumba, Orange: Mchoka, Brown: Moet Red: Mtengeza Black: Ndem-bo, Light Blue: Samama , Pink: St Augustine 2, Dark Green: St Martins, Mauve: Sungusya	153

B.6	Probability of the being positive with <i>Schistosoma</i> versus the average residuals, a) <i>S. mansoni</i> [T+], b) <i>S. haematobium</i> and c) co-infection [T+].	154
B.7	Probability of the being positive with <i>Schistosoma</i> versus the average residuals a) <i>S. mansoni</i> [T-] and b) co-infection [T-].	155
B.8	Water contact rate calculated using the questionnaire answers from the SAC collected in the primary study. Age of SAC versus their school A) Enter Lake rate B) Swim Lake rate C) Drink late rate D) Water Contact rate.	156
B.9	Water contact rate versus prevalence of <i>Schistosoma</i> infection for each school. Legend: Different symbols represent each school	157
C.1	Flow diagram showing the stages for constructing the 200 predictions along the shoreline. a) A 2-D linestring was drawn by hand following the shoreline as shown by Google Satellite imagery, b) the linestring was re-sampled to 4000 equally spaced vertices and resampled them to 200 equally intervals (red dots), c) observed sampling site locations (blue dots), d) each observed sampling site location was snapped to its nearest vertex (green dots).	159
C.2	The distance along the line from the origin (northwest-most vertex) to each of the snapper observed sampling site locations for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp..	160
C.3	Scatterplot of environmental data extracted versus distance along shoreline (km) for prediction points	161
C.4	<i>Biomphalaria</i> sp. observed sample points extracted environmental data a) Rainfall (mm) b) LST (°C) c) NDVI (index) d) Evapotranspiration (index) e) Soil type . . .	162
C.5	<i>Bulinus</i> spp. observed sample points extracted environmental data. a) Rainfall (mm) b) LST (°C) c) NDVI (index) d) Evapotranspiration (index) e) Soil type . . .	163
C.6	Observed points values for each covariate centred and scaled a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	165
C.7	Histograms of prediction points values for each covariate centred and scaled. a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	166
C.8	Comparison of covariance functions for above snails mean abundance for a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp. against distance along (km).	168

C.9	Comparison of covariance functions for number of snails predicted for a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp. against distance along (km).	169
C.10	Prior against the posterior distribution a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	171
C.11	1D mean GP prediction (exponential covariance function) of number of snails $\log(\hat{\mu}_i)$ against distance along the shoreline (km) a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp. Red crosses: observed number of snails at sampling locations along the shoreline. Black faded lines : SD of GP values	173
C.12	Bathymetric water depth (m) data for the shoreline with observed snails locations plotted a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	174
C.13	2D plot of depth (m) extracted for the prediction points with 100km buffer.	175
C.14	Depth (m) against distance along (km) the shoreline with 100km buffer and fill in NAs for the shoreline.	175
D.1	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.05$ for SAC age 6. Multi- β_s with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	177
D.2	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.05$ for SAC age 6. Single- β with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	178
D.3	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.05$ for SAC age 6. Multi- β_s with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	179
D.4	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.05$ for SAC age 6. Single- β with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	180

D.5	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.10$ for SAC age 6. Multi- β_s with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	181
D.6	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.10$ for SAC age 6. Single- β with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	182
D.7	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.10$ for SAC age 6. Multi- β_s with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	183
D.8	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.10$ for SAC age 6. Single- β with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	184
D.9	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.20$ for SAC age 6. Multi- β_s with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	185
D.10	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.20$ for SAC age 6. Single- β with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	186
D.11	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.20$ for SAC age 6. Multi- β_s with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	187
D.12	Prior <i>Schistosoma</i> infection prevalence of $\alpha = 0.20$ for SAC age 6. Single- β with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) <i>Biomphalaria</i> sp. b) <i>Bulinus</i> spp.	188

D.13 Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species 189

D.14 Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Parameter estimates for single- β spatial model for each species 190

D.15 Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species 191

D.16 Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Parameter estimates for single- β spatial model for each species 192

D.17 Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species 193

D.18 Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Parameter estimates for single- β spatial model for each species 194

D.19 Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species 195

D.20 Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Parameter estimates for single- β spatial model for each species 196

D.21 Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species 197

D.22 Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Parameter estimates for single- β spatial model for each species 198

D.23 Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species 199

D.24 Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Parameter estimates for single- β spatial model for each species 200

D.25 Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Confidence intervals for parameter estimates for *Biomphalaria* sp. models with SAC prevalence at age 6 set as α set to 0.05 (5%) a) Multi- β_s space sp. b) Multi- β_s no space c) Single- β space d) Single- β no space 201

D.26 Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Confidence intervals for parameter estimates *Bulinus* spp. models with SAC prevalence at age 6 set as α set to 0.05 (5%) a) Multi- β_s space sp. b) Multi- β_s no space c) Single- β space d) Single- β no space 202

D.27 Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Confidence intervals for parameter estimates for *Biomphalaria* sp. models with SAC prevalence at age 6 set as α set to 0.10 (10%); a) Multi- β_s space b) Multi- β_s no space c) Multi- β_s space d) Multi- β_s no space 203

D.28 Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Confidence intervals for parameter estimates for *Bulinus* spp. models with SAC prevalence at age 6 set as α set to 0.10 (10%); a) Multi- β_s space sp. b) Multi- β_s no space c) Single- β space d) Single- β no space 204

D.29 Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Confidence intervals for parameter estimates for *Biomphalaria* sp. models with SAC prevalence at age 6 set as α set to 0.20 (20%): a) Multi- β_s space sp. b) Multi- β_s no space c) Single- β space d) Single- β no space 205

D.30 Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Confidence intervals for parameter estimates for *Bulinus* spp. models with SAC prevalence at age 6 set as α set to 0.05 (5%): a) Multi- β_s space sp. b) Multi- β_s no space c) Single- β space d) Single- β no space 206

D.31 *Biomphalaria* sp. Multi- β with space effect model profile likelihood against fit of model. 207

D.32 *Bulinus* sp. Multi- β with space effect model profile likelihood against fit of model . 208

D.33 Single- β with space effect model profile likelihood against fit of model a) *Biomphalaria* sp. b) *Bulinus* spp. 209

D.34 *Biomphalaria* sp. Multi- β with no space effect model profile likelihood against fit of model. 210

D.35 *Bulinus* sp. Multi- β with no space effect model profile likelihood against fit of model 211

D.36 Single- β with no space effect model profile likelihood against fit of model a) *Biomphalaria* sp. b) *Bulinus* spp. 212

D.1 Poster Presented at British Parasitology Conference (BSP) conference 2021. Part A: Intestinal and urogenital schistosomiasis dual-infection focus: Investigating age-infection relationships for school-aged children along shoreline of Lake Malawi 214

D.2 Poster presented at BSP conference 2021. Part B: Intestinal and urogenital schistosomiasis co-infection focus: Investigating age-infection relationships for school-aged children along shoreline of Lake Malawi 215

D.3 Post presented at BSP conference 2022. A geospatial analysis of local intermediate snail host distributions provides insight into intestinal and urogenital schistosomiasis within under-sampled areas of Lake Malawi 215

List of Tables

1.1	A summary of the <i>Schistosoma</i> species mainly infecting humans within Africa and their associated intermediate snail hosts and geographical distribution [14].	4
1.2	Infection intensity [49]	16
1.3	Treatment strategy recommended for schistosomiasis with preventive chemotherapy [63]	22
2.1	Summary of prevalence of <i>S. mansoni</i> [T+], <i>S. haematobium</i> and co-infection [T+]	52
2.2	Coefficients for the GAM with age smoothing term adjusted for school	55
3.1	Estimated parameter values, mean and CrI	73
4.1	Parameters used for the 4 models fitted to <i>Bulinus</i> spp. and <i>Biomphalaria</i> sp. data: single vs school-specific β , with and without the spatial term	93
4.2	Parameter estimates for multi- β_s spatial model for each species	100
4.3	Parameter estimates for single- β spatial model for each species	101
4.4	Parameter estimates for multi- β_s no spatial model for each species	102
4.5	Parameter estimates for single- β no space model outcome for each species	103
B.1	Summary of prevalence of <i>S. mansoni</i> [T-] and co-infection [T-]	151
B.2	GAM with smooth term age adjusted for school	151

List of Papers

This thesis contains the following appended papers.

Paper 1: *Modelling the age-prevalence relationship in schistosomiasis: A secondary data analysis of school-aged-children in Mangochi District, Lake Malawi.* **Amber L. Reed**, Angus M. O’Ferrall, Sekeleghe A. Kayuni, Hamish Baxter, Michelle C. Stanton, J. Russell Stothard, Christopher Jewell

- Published in *Parasite Epidemiology and Control*, doi: <https://doi.org/10.1016/j.parepi.2023.e00303>
- Contribution: Corresponding author, co-design of statistical analyses, implementation and interpretation of analysis, drafting of manuscript and use of co-authors feedback.

Paper 2: *A geospatial analysis of local intermediate snail host distributions provides insight into intestinal and urogenital schistosomiasis within under-sampled areas of southern Lake Malawi.*

Amber L. Reed, Mohammad H. Al-Harbi, Peter Makaula, Charlotte Condemine, Josie Hesketh, John Acher, Sam Jones, Sekeleghe A. Kayuni, Janelisa Musaya, Michelle C. Stanton, J. Russell Stothard, Claudio Fronterre, Christopher Jewell

- Published in *Parasites and Vectors*, doi: <https://doi.org/10.1186/s13071-024-06353-y>
- Contribution: co-design of statistical analyses, implementation and interpretation of analysis, drafting of manuscript and use of co-authors feedback.

Paper 3: *Development of a dynamical model to enhance understanding of epidemiology of schistosomiasis in school aged children* **Amber L. Reed**, J. Russell Stothard, Claudio Fronterre, Christopher Jewell

- Manuscript circulated to co-authors in preparation for submission
- Contribution: co-design of statistical analyses, implementation and interpretation of analysis, drafting of manuscript and use of co-authors feedback.

List of Abbreviations and acroyms

AIC	Akaike Information Criterion
ART	Artemisinin
BMLM	Bayesian multilevel model
CM	Cambisols
CAA	Circulating Anodic Antigen
CCA	Circulating Cathodic Antigen
CI	Confidence Intervals
DoF	Degrees of Freedom
DNA	Deoxyribonucleic Acid
ELISA	Enzyme Linked Immunosorbent Assay
GAM	Generalised Additive Model (GAM)
GAM	Generalised Linear Model (GLM)
GP	Gaussian Process
GL	Gleysols
GLOBathy	GLObal Bathymetric
GPS	Global Positioning System
IgE	Immunoglobulin E
ITCZ	Inter-Tropical Converge Zone
IS	Intestinal schistosomiasis
ISRIC	International Soil Reference and Information Centre
KK	Kato-Katz
LPDAAC	Land Processes Distribution Active Archive Center

LSTM	Liverpool School of Tropical Medicine
LST	Land Surface Temperature
LV	Luvisols
MDA	Mass Drug Administration
MCMC	Markov chain Monte Carlo
MLE	Maximum Likelihood Estimation
MSE	Mean squared error
MODIS	Moderate Resolution Imaging Spectroradiometer
NSCP	National Schistosomiasis Control Programme
NDVI	Normalised Difference Vegetation Index
NTD	Neglected Tropical Disease
ODE	Ordinary differential equation
1D	One-Dimensional
PL	Planosols
PLE	Penalised Likelihood Estimation
PCR	Polymerase Chain Reaction
POC	Point-of-care
PSAC	Pre-school-aged-children
PZQ	Praziquantel
PC	Preventative chemotherapy
SAC	School-aged-children
<i>S. haematobium</i>	<i>Schistosoma haematobium</i>
<i>S. mansoni</i>	<i>Schistosoma mansoni</i>
SOTER	Soil Terrain Database for Malawi

SSA	Sub-Saharan Africa
SEIRS	Susceptible-Exposed-Infected-Recovered-Susceptible
SIR	Susceptible-Infected-Recovered
TAMSAT	Tropical Applications of Meteorology using SATellite data and ground-based observations
2D	Two-Dimensional
RNA	Ribonucleic Acid
UGS	Urogenital schistosomiasis
WASH	Water, sanitation, hygiene interventions
WHO	World Health Organisation

Chapter 1

Introduction

1.1 Thesis overview

The overall focus of my thesis concerns spatial and temporal epidemiology of schistosomiasis along the shoreline of southern Malawi, Mangochi District. My approach involves developing and applying different types of statistical models and analyses. This is to understand schistosomiasis transmission as my goal is to better reveal those infection and disease determinants that might be amenable to future intervention, and hence aid in improving control.

1.2 Schistosomiasis

1.2.1 Epidemiology overview

Despite intensive successful control programmes, schistosomiasis remains an extensive public health concern. The disease affects over 240 million people worldwide, its burden disproportionately affects individuals in sub-Saharan Africa (SSA) where over 90% of infected individuals reside [1] (Figure 1.1). Schistosomiasis, also known as Bilharzia, is a Neglected Tropical Disease (NTD), caused by blood fluke worms (trematodes parasites) belonging to the genus *Schistosoma* [2]. Moreover, schistosomiasis is a water-borne disease found mostly in tropical and sub-tropical areas. It is often endemic in poor communities with no access to safe drinking and washing water and inadequate sanitation and education [1, 3, 4]. It can cause mortality or lifelong severe morbidity within children, adolescents, and young adults, who are most vulnerable to the disease [3, 5]. This then can negatively impact children's physical development, affect their school performance

and consequently, impair the social and economic development of endemic areas [3]. Often the infection distribution of schistosomiasis is considered focal where there are localised communities with the highest risk to infection. Therefore, identifying and assessing these high-risk areas can help improve control method implantation and reduce the application of wasteful resources [6].

Schistosomiasis is considered endemic in over 76 other countries worldwide, with a further 700 million people at risk of *Schistosoma* infection [4, 7]. Although several species of schistosome exist, the two dominant schistosomes responsible for human infection are *Schistosoma haematobium* and *Schistosoma mansoni* (Figure 1.1). These two species partition into two separate forms of clinical disease, these being either urogenital schistosomiasis (UGS) or intestinal schistosomiasis (IS), respectively. As schistosomes require obligatory development inside particular species of freshwater snails, the geographical disease transmission of the two forms of *Schistosoma* are typically restricted by various freshwater snail host habits along the shoreline of Lake Malawi. Of note, *S. mansoni* is dependent on freshwater snails belonging to the genus *Biomphalaria*, whereas *S. haematobium* is dependent on freshwater snails belonging to genus *Bulinus* [2] (Table 1.1). These two snail genera have very different evolutionary histories but are both fully aquatic and cannot survive drying out. Of note, there are several species within each genus and not all species transmit schistosomes [8].

Schistosoma haematobium causes UGS, which was named based on its main symptom hematuria (blood in urine) and can lead to bladder cancer in some cases [9], as this parasite is a known carcinogen. In SSA, the World Health Organisation (WHO) has reported that UGS is responsible for approximately 150,000 deaths due to kidney failure and a further 13,300 deaths by bladder cancer each year. By contrast, IS is caused primarily by *S. mansoni* [7], for which the WHO reports approximately 130,000 deaths per year due to upper and/or lower gastrointestinal hemorrhage [10]. In the southern region of Malawi, it is well known that schistosomiasis is endemic with *S. haematobium* [3], however, only recently have studies reported more cases of *S. mansoni* occurring in this region [11].

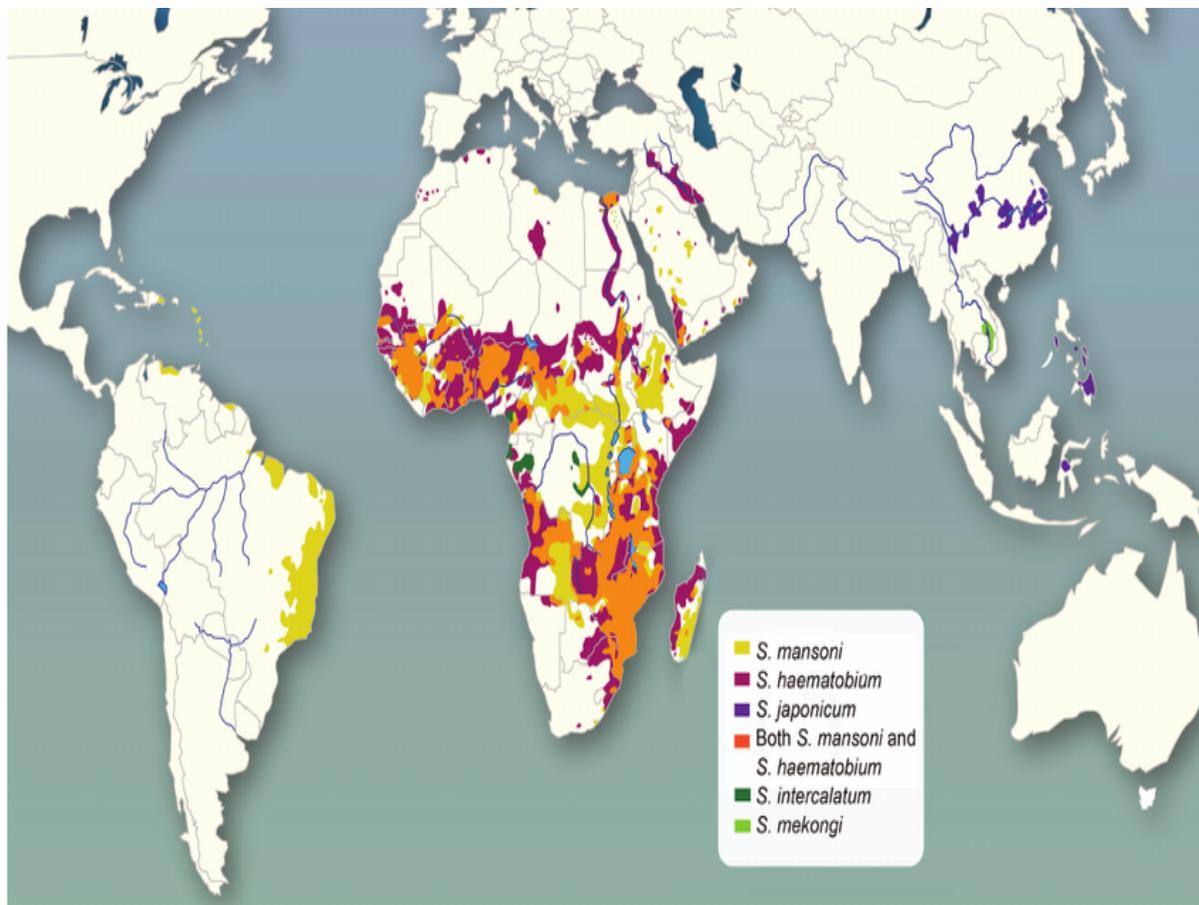


Figure 1.1: The geographical distribution of schistosomiasis worldwide, reproduced from Weerakoon et al. under a CC-BY license [12].

The main recommended control and treatment by the WHO for schistosomiasis is preventive chemotherapy (PC) with Mass Drug Administration (MDA) of the drug praziquantel (PZQ). PZQ was developed in the 1970's, and provides a safe and cost-effective single dose treatment [3, 5, 7]. Despite the success of this recommended control measure, frequent re-infection of individuals occurs leading to regional endemicity [3, 5, 13]. To tackle this, screening programmes have been introduced, which often lead to retreatment being required to reduce the transmission of the disease.

Table 1.1: A summary of the *Schistosoma* species mainly infecting humans within Africa and their associated intermediate snail hosts and geographical distribution [14].

<i>Schistosoma</i> species	Major host snail	Intermediate host	Geographical distribution
<i>S. haematobium</i>	<i>Bulinus globosus</i> <i>Bulinus truncatus</i> <i>Bulinus nyassanus</i> *		Africa Middle East Corsica (France)
<i>S. mansoni</i>	<i>Biomphalaria pfeifferi</i> <i>Biomphalaria sudanica</i>		Africa Middle east Caribbean Brazil South America

* *B. nyassanus* is not a major host across Africa but in Lake Malawi plays an important role in deep-water habitats

1.2.2 Transmission and lifecycle

There are several stages in which the transmission of schistosomes occurs (Figure 1.2). The first stage is when an infected individual urinates or excretes into the freshwater source. The adult schistosome parasites residing in the mesenteric (*S. mansoni*) or vesicular veins (*S. haematobium*) of the already infected individual produces eggs which are then shed when excreted through faeces (*S. mansoni*) or urine (*S. haematobium*) into the freshwater source [4].

Once the *Schistosoma* ova (eggs) come into contact with the freshwater source, the ova will hatch when suitable environmental conditions occur and release miracidia (larvae). The miracidia need to swim and penetrate a specific freshwater snail (intermediate host) within 32 hours after the *Schistosoma* ova hatch and are released. Once inside the freshwater snail, the miracidia larvae multiply by shedding their outer ciliated epidermal layer and develop into a mother sporocyst. Germ balls are accumulated within the mother sporocyst, causing the sporocyst to burst and move to the snail's digestive gland. These germ balls develop into daughter sporocysts via asexual multiplication (same sex). Daughter sporocysts are the larvae cercariae which are released in large amounts into the freshwater source. As this process is carried out by asexual multiplication a large quantity of larval cercariae can be produced. This process from miracidia larvae entering an intermediate snail host to the production and release of cercariae into a freshwater source takes approximately a month. The cercariae being released from the snail are affected by light and temperature, therefore their emergence times through the day and season (wet or dry) can vary among *Schistosoma* species [16, 17].

A human becomes infected with schistosomiasis by the cercariae penetrating their skin when they

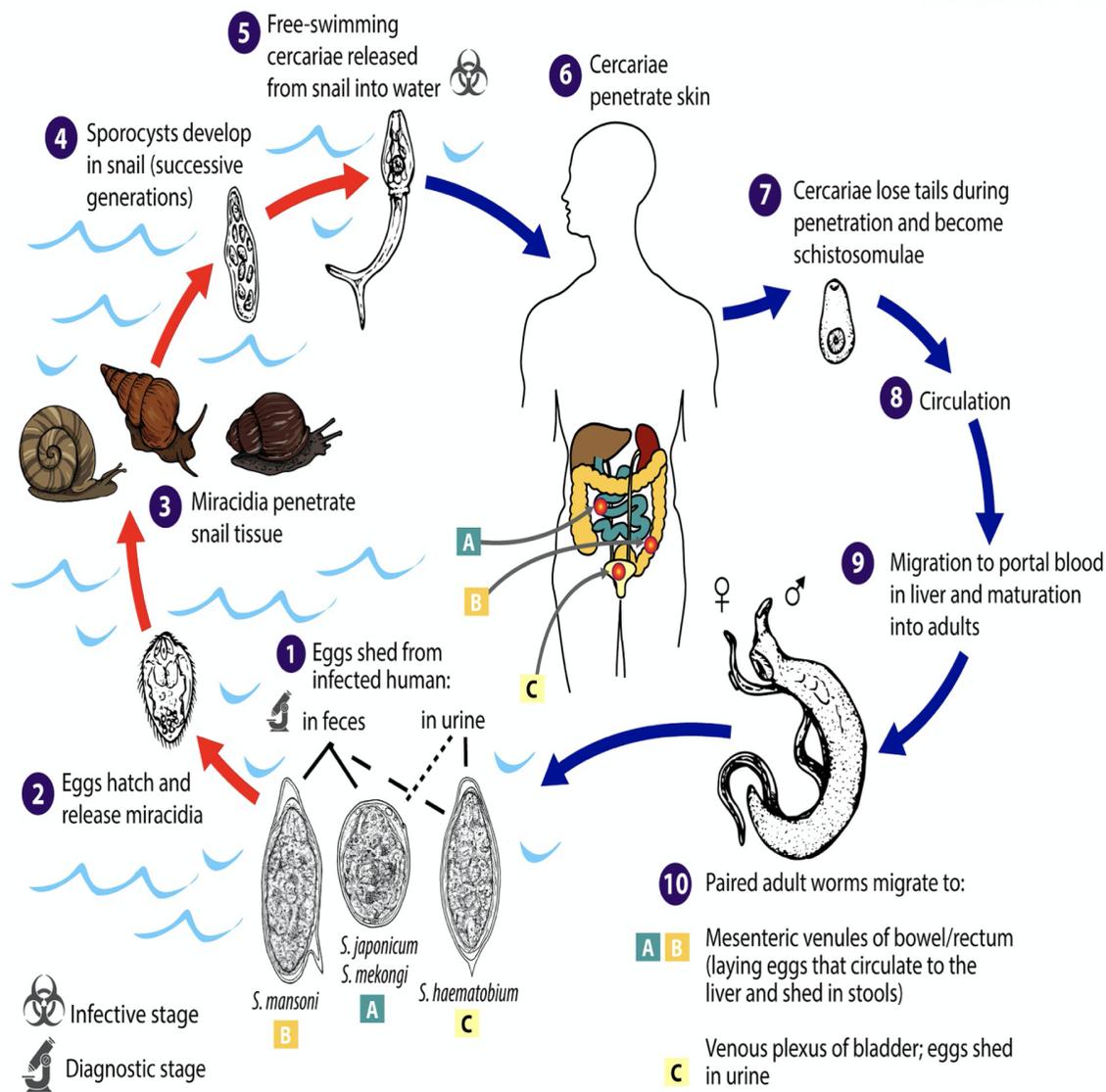


Figure 1.2: Lifecycle of schistosomiasis, reproduced from CDC *et al.* under a CC-BY license [15].

enter an infested freshwater source [4]. For the transmission to be successful, the cercariae need to quickly to penetrate the skin of a human host, as they become less infectious overtime; usually within 72 hours, or they die. The cercariae swim to a human host using its forked tails and shed them when penetrating the skin of the human host. Shedding their tails, they transform into immature schistosomulae and travel via the blood stream (venous circulation). Then 4–7 days after penetration the schistosomulae move to the lungs, followed by heart and liver, where they mature and elongate into more recognizable worms. They then travel back into the blood stream (the portal vein system) and finally reside in the veins of the gastrointestinal or genitourinary tract [4, 15, 16]. This process takes around 4–6 weeks, from the immature schistosomulae to the mature adult schistosome coming to reside in the veins and producing eggs [17].

Over approximately a month, the parasite goes from being cercariae at the skin penetration stage to being a mature adult worm. Schistosomiasis is somewhat unusual when compared to other diseases, but common to other worm-related infections is that adult worms don't multiply and increase in numbers directly. Hence, any increase in adult worm numbers is directly proportional to exposure events. More unusual in comparison to other trematode flatworms is that schistosomes are dioecious, having male and females, whereas all other trematodes are hermaphroditic. The adult schistosomes pair up with the male holding or embracing the female within its gynaecophoric canal (ventral groove). This close proximity allows the pair to live in perpetual copulation where females may produce eggs allowing the paired male to fertilize them [16, 17]. The female adult worm can produce from 200 to 2,000 eggs per day for about 5 years [4]. This leads to the first stage of the cycle again, where eggs are shed into the freshwater source through excretion or urination. In the case where eggs are produced but not shed through excretion, these eggs remain in the host tissue. Over a period of 1 to 2 weeks, the eggs will die within the tissues of the host; this can cause tissue inflammation and granuloma formation, which can lead the infected host to start displaying symptoms due to the tissue damage [4, 17]. Consequently, schistosomiasis can cause widespread severe morbidity and, in some cases, mortality, as just one pair of adult schistosomes can produce thousands of eggs per day [16].

An essential dependent for the transmission of schistosomiasis is thought to be the successful survival and fecundity of the adult schistosomes within an infected human host. This is due to the adult schistosomes being able to survive an average 3–6 years within a human, and in some reported cases up to 40 years, which is a substantially longer lifespan than within the snails (a lifespan of weeks) or free-living stages (hours). This longer lifespan gives the most chance that a infected individual is to infect another individual over this time through water contact and consequently maintain the transmission of schistosomiasis [17, 18].

Another essential dependent for the transmission of *Schistosoma* species is the location of the intermediate host snails along the shoreline [5]. There are 400 species of freshwater snails in Africa, however, only certain species of snail are able to harbour the parasite. The snail species *Biomphalaria* and *Bulinus* allow the parasite to develop inside them [14]. The life cycle of schistosomiasis cannot be completed without the miracidia penetrating a 'host specific' snail in order to develop inside and release the cercariae essential for transmission of the disease. For this reason, the transmission of schistosomiasis has a focal distribution with heterogeneity where the prevalence and intensity of *Schistosoma* infection can vary among small distances between geographical areas. Hence, transmission occurs where humans enter a water source near the specific snails [18]. Therefore, it is the location of snail populations and the successful establishment of the adult schistosomes within a human host which are the main two determinants for the transmission rate

of schistosomiasis. There are also multiple factors which can impact the transmission rate of schistosomiasis in a given location, including environmental, socio-economic and human behavioural factors within communities.

1.2.3 Environment and human behaviour factors

The environment can have an impact on the successful survival and development of schistosomiasis and snail host populations [18]. Environmental factors include climate change, water temperature of the water source, development of water dams for irrigation and the introduction of hydroelectric power [5, 19]. A community's proximity to water sources can affect the transmission rate at each geographical location. Factors including land surface and diurnal temperature and maximum rainfall have all been considered risk factors for transmission depending on the species of *Schistosoma*, and are known to affect the intermediate host geographic distribution [20]. Temperatures around 25°C increase survival of the snails, whereas above 25°C can lead to increasing snail deaths. It is believed that low temperatures have a minor effect on snail populations. The snail habitat can also be affected by vegetation, the pH and velocity of the water (flow) [18]. Seasonal changes can also have an effect on the transmission of schistosomiasis, where it is thought the wet season provides more favourable conditions for the parasite to survive and spread [21].

The socio-economic status of an area can be an indicator of the risk of schistosomiasis transmission. For example, schistosomiasis is linked to areas in poverty, that lack access to sanitation facilities and personal hygiene education. As mentioned earlier, the water contact patterns of an individual have been found to affect the *Schistosoma* infection transmission rate; this is because increasing the frequency and period of water exposure increases the chance that the individual will come across waters infested with cercariae [4]. In these low-income areas, humans often enter the contaminated water to take part in swimming, washing, farming, or wading which increases their risk of schistosomiasis transmission. High risk groups are thought to be children, fisherman, farmers, irrigation workers and those performing other duties which require regular exposure to freshwater sources [7]. School-aged-children (SAC) are one of the high-risk groups of the disease, as they will often spend long periods in the water for enjoyment, cooling off from the heat and recreationally bathing [22]. Young children are often more stationary in water than older children and adults: this is considered to have an effect on the efficacy of cercarial penetration. Young children's behaviour is often linked to their mother's water contact patterns; as this follows the sociocultural patterns of the area and can lead to the transmission of schistosomiasis in these children [23].

The spread of health education among communities on presentation, control, transmission, and

risk factors has been assessed as vital in reducing the number of cases of schistosomiasis [24]. The construction of infrastructure such as bridges has also been found to have a role in reducing the transmission of *Schistosoma* infection, by reducing the number of people entering the infested water sources in order to cross the river. The construction of toilets in schools is also thought to be beneficial in reducing the transmission of schistosomiasis [19]. However, developing infrastructure can cause to migration to urban areas and increased human population movement which has been found to contribute to the disease being introduced into new geographical areas. This is of concern, as mentioned before, as travellers and tourists from non-endemic areas are more at risk of severe symptomatic infection [4].

1.2.4 Pathogenesis and clinical manifestations

Schistosomiasis has three stages of clinical disease progression: early acute, established active infection and chronic manifestations. These are described in the following sections.

1.2.4.1 Acute disease

The skin penetration from cercariae often causes the first common signs of being infected with schistosomiasis; this is known as cercarial dermatitis [4]. This is the host immune system giving an innate immune response and can cause raised lesions which can occur after 3 hours, and a rash can appear within 7 days post-infection. This latter occurs more often in travellers and migrants visiting endemic areas. However, often the host does not display symptoms until a few weeks later, as symptoms are thought to be caused not by the adult worms themselves but by the eggs they produce, laid by the female worms. These are often not excreted and remain in the intestines or liver (IS) or in the bladder and urogenital system (UGS). The adult worms do not cause an immune response in the host, as the worms are thought to adapt by acquiring host antigens and masking their own antigens [25].

Travellers or immigrants who travel to endemic areas, often display acute schistosomiasis also known as Katayama syndrome. They are exposed to the disease at a later stage in life than expected for individuals living in endemic areas. Clinical symptoms can take weeks or even months to appear after the individuals are infected; this is due to the time taken from the cercariae penetrating the skin to the *Schistosoma* worm maturing and the production of eggs which causes symptoms [17]. Clinical symptoms can include fatigue, malaise, fever, abdominal pain, diarrhoea, headache, myalgia (pain in muscle), eosinophilia, and a non-productive cough which can last up to 2–10 weeks [5, 25]. Most recover, however, in some cases further symptoms can occur from persistent

disease such as a generalised rash, abdominal pain, diarrhoea, weight loss, enlargement of the liver (hepatomegaly) and breathing difficulties (dyspnoea) [25]. Often individuals who do not live in *Schistosoma* endemic areas are more likely to display these acute and severe symptoms. One possibility explored in research is that often the mother of the individual has had previous chronic infections due to living in an endemic area which can lead to an immune response in the baby. This is where the T and B lymphocytes are primed in the womb, leading to the individual immune system becoming ready to produce antibodies against the *Schistosoma* antigens [17].

1.2.4.2 Established active and chronic infections

Individuals living in endemic areas where they have had repeated exposure to *Schistosoma* infection often developed an established active infection instead of an acute infection. An established active infection is considered when mature adult worms are present and are producing eggs which are excreted via urine or faeces. Often the eggs get lodged within the tissues causing an inflammatory response including a granulomatous response triggered by the host and involving eosinophils, macrophages, and lymphocytes [25]. This is due to the eggs secreting glycoprotein antigens, which are produced to allow the eggs to travel through the blood vessels to suitable locations (e.g. the urinary tract or intestine).

The associated morbidity (severity of symptoms) of an individual is thought to be related to the intensity of *Schistosoma* infection and the egg-induced inflammation [25]. The chronic manifestations of *Schistosoma* infection are related with ongoing local inflammation occurring over time due to the unshredded schistosome eggs remaining trapped within the host tissues and the immune response's inability to remove the infection leading to granuloma formation [4, 17]. In children living in endemic areas, established active infections are often treatable by repeated praziquantel treatment, which kills the adult worms [25]. Furthermore, chronic *Schistosoma* infection can lead individuals to be more vulnerable to other diseases, for instance malaria. For those with HIV the progression of disease is known to be fast tracked [26].

1.2.5 Urogenital schistosomiasis

In UGS (*S. haematobium*), the granuloma formation within the tissues induced by the glycoprotein antigens can cause polyp lesions, abnormal tissue growth and ulcers to form at the opening of the ureter and bladder which can be observed using ultrasonography. This process can lead to urinary frequency, burning sensation during urination, pelvic pain and the main symptom of UGS, haematuria (blood in urine) [25, 27]. Haematuria is often mistaken as menstruation in girls, and

in some endemic regions, it is incorrectly believed that it is a sign of males going through puberty [28]. Chronic manifestations of UGS can be due to lack of immunoregulation of anti-schistosomes responses [29]. The long-term local inflammation can lead to blockages within the genitourinary system causing renal complications, which include renal failure or cancer of the bladder, which in some cases can be fatal [4].

Infertility and diminished libido can also occur in both males and females [4]. Infertility or reduced reproductive health in females is caused by schistosome eggs in the genital tract, which cause damage to their ovaries, fallopian tubes, vulva, vagina and cervix [30]. Whereas in males, UGS causes lesions in the prostate and seminal vesicles which if left untreated lead to cancer, swollen painful testicles, and death of local vessels, which can lead to infertility. Further, UGS in males is associated with genital organs, production of semen changes and other symptoms including haemospermia (presence of blood), weak erection, or premature ejaculation [31]. It is thought that infertility in males is rare, however, the rate of infertility is unknown [4].

1.2.6 Intestinal schistosomiasis

In IS (*S. mansoni*), the main symptom at the chronic stage is a distended abdomen and other symptoms including diarrhoea and rectal bleeding. These three symptoms are associated with lesions in the appendix (mucosal hyperplasia), formation of scar tissue (pseudopolyposis) and formation of polyps (polyposis) within the gastrointestinal tract. Chronic disease can cause gastrointestinal disease leading to hepatosplenic inflammation and, in the presence of heavy infection, liver fibrosis [27, 32]. Chronic manifestations of IS can be due to the down-regulation of the granulomatous response in the host [17]. Gastrointestinal disease can lead to symptoms including abdominal pain, enlargement of the liver (hepatomegaly) and spleen (splenomegaly). Furthermore, portal hypertension can occur leading to enlargement of the veins in the oesophagus (oesophageal varices) and abnormal fluid accumulation within the abdomen (ascites) causing a distended abdomen and haemorrhage which in some cases, can be fatal [33].

1.2.7 Co-infection

Co-infection with both UGS (*S. haematobium*) and IS (*S. mansoni*) can occur due to ongoing infections occurring at the same time and within the same area, which can increase the burden of disease on individuals in their community and often lead to chronic schistosomiasis. Further, co-infection can possibly increase the likelihood of individuals becoming vulnerable to other infectious diseases such as malaria or hepatitis C virus [26].

1.2.8 Immunity

The human host immune system interaction with a *Schistosoma* infection involves several different stages of the parasite life cycle, including the following: cercariae penetrating the skin, immature and mature adult schistosome and parasite eggs which remain in the tissues of the host if not excreted. Immunity to *Schistosoma* infection, known as resistance to reinfection, is thought to be developed over a period of 10–15 years from first infection [5]. In most individuals immunity is acquired over time; avoiding chronic disease, as the worm burden gradually declines over time, known as a partial acquired immunity against new infections. This leads to fewer eggs being produced and consequently, reduces the number of eggs remaining lodged within the host tissues, reducing inflammation and granuloma formation. Immunity can also occur over time due to immunological down-regulation, which causes the granulomas which do form to be smaller than the previous ones, which have already been developed into fibrous tissue and therefore symptoms are reduced over time [25]. A known vulnerable group is children who are thought to be much more susceptible to reinfection post treatment than their counterparts. Consequently, it is thought that as an individual ages they develop more immunity to the disease due to repeated exposure when they have been infected from an early age. Therefore, one component of immunity of schistosomiasis is that a possible age-related immunity occurs.

The age-dependent acquired immunity in endemic areas for schistosomes can be explained by the following: the long length of time it takes an individual to build immunity to *Schistosoma* infection is thought to possibly be due to more dead worms remaining within the host tissues after infection. The worms can take 3–10 years to die naturally or to die due to treatment after an individual is infected with schistosomiasis. Once dead, the worms can cause antigens to be released as waste products into the tissues of the host which stimulates the production of specific immunoglobulin E (IgE), which leads to a protective response, and hence leads to resistance to reinfection in an individual. Therefore, research has predicted that as an individual becomes more exposed, they accumulate more dead worms causing more IgE antigen production, and hence a stronger immune response against reinfection [5]. Another component of immunity is thought to be developed through the immune response to secreted egg antigens within the host tissue, for instance, the liver [34].

Despite the successful application of MDA programmes with PZQ chemotherapy treatment targeting SAC and high-risk adults within endemic areas, there remains high reinfection prevalence within these areas for schistosomiasis, as well as considerable morbidity and mortality. Moreover, there is a considerable concern for PZQ drug resistance to develop due to the continual and high use of PZQ, however, no evidence of this has been shown to date [25]. Current vaccines against

schistosomiasis for humans have yet to be developed for use, however, proposed candidates for an effective vaccine have been studied and research suggests there is strong evidence that a feasible vaccine could be produced [34].

1.2.9 Diagnosis

Diagnoses of schistosomiasis is vital not only in reducing the transmission of the disease, but also to reduce morbidity and mortality caused by the disease in addition to control methods such as MDA due to ongoing transmission in focal areas [35]. The ability to accurately diagnose an individual allows for fast treatment and can prevent or reduce the early stages of chronic infection [12]. There are multiple dependent factors able to detect a *Schistosoma* infection in an individual. These can include time of infection, treatment history, type of immunoassays used, test standard, intensity of infection and lastly the test sensitivity and specificity [36]. The type of species of *Schistosoma* infection can affect which diagnostic tests are used for diagnosis of an individual. Selection for a certain diagnostic test can also be confounded by the geographical location of testing, financial cost, availability of resources, education of health care workers and social-economic situation. This is an important consideration as schistosomiasis is often endemic in low-income areas which lack resources and funding.

Clinical assessment of an individual for a disease usually consists of a diagnostic test which can be used to determine the presence or absence of a disease and is used when an individual is displaying signs or symptoms of the disease. Screening tests can be used as a tool to identify asymptomatic individuals who do not display symptoms within a community. It is important for diagnostic tests to have the ability to produce an accurate and valid result, to be able to identify the disease and to differentiate whether the individual is positive or negative. The test measurement of its diagnostic accuracy or ability to identify which disease is usually measured by sensitivity, specificity, predictive values and likelihood ratios [12]

$$\text{Sensitivity} = Pr(\text{test positive}|\text{disease positive}) \approx \frac{\#\text{true positive}}{\#\text{true positive} + \#\text{false negative}} \quad (1.1)$$

$$\text{Specificity} = Pr(\text{test negative}|\text{disease negative}) \approx \frac{\#\text{true negative}}{\#\text{true negative} + \#\text{false positive}} \quad (1.2)$$

Sensitivity and specificity are the main two important measures for determining the accuracy of

a diagnostic test. Certain tests are considered gold standard by the research, and can be used to compare the measured accuracy of a diagnostic test [12]. The definition of sensitivity of a test is the test's ability to correctly identify patients who are actually disease positive (a true positive result), whereas specificity is the ability to identify which patients do not have the condition (a true negative result) (Equation 1.1 and 1.2). For instance, if a diagnostic test has a low sensitivity, then an infected individual could be given a negative result when truly positive, hence the individual will not be treated and continue to transmit the disease and be more likely to develop chronic morbidity due to high schistosome worms within their tissues. Diagnosis of an individual usually benefits from using a test with a combination of high sensitivity and specificity, however, it is important to also consider the time required to perform the test, the training and expertise required to carry out the test and the financial cost. In some incidences, high sensitivity for case detection is prioritised over high specificity; this is often where low transmission and low prevalence in a particular area is occurring [35]. There are a number of developed diagnostic methods to detect *Schistosoma* infection varying from microscopic detection or molecular techniques. The types of diagnostic tests can include direct to indirect parasitological methods, immunological diagnosis or lastly DNA and RNA detection (Figure 1.3).

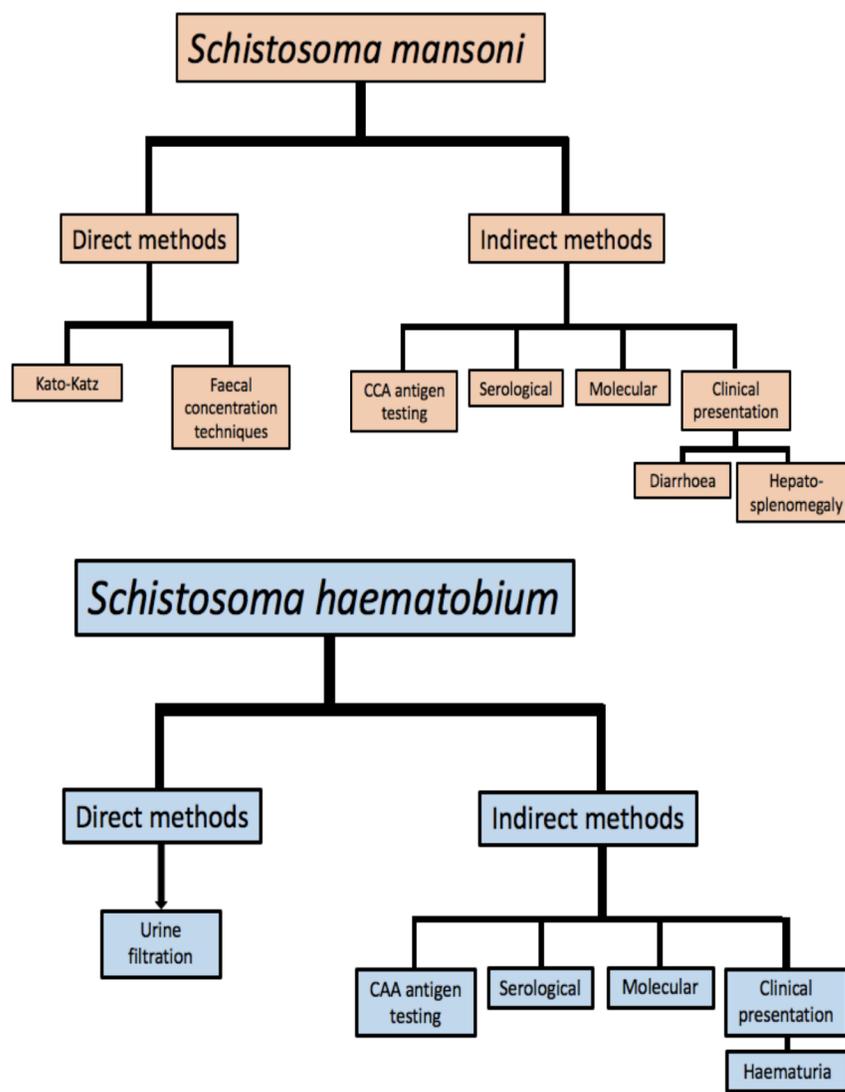


Figure 1.3: Schematic diagrams showing the different diagnostic methods available for the diagnosis of *S. mansoni* and *S. haematobium*, reproduced from Christiansen, 2018 under a CC-BY license [37]

1.2.9.1 Direct parasitological methods

i Diagnosing UGS by urine filtration

The gold standard method for diagnosis of *Schistosoma haematobium* infection in the field is known as urine filtration using a syringe followed by 100x microscopy. The test works by taking 10ml of urine syringed through a polycarbonate filter to capture eggs and then these filters are stained and observed under a compound light microscope [38]. This is the chosen

method in the field as it is easy to use, low cost and requires no specialist equipment. In higher-income areas with specialist facilities, centrifugation can be used instead of filtration to find the concentration of *S. haematobium* eggs in urine. Both these tests can often have low sensitivity; despite this the number of eggs can be identified more easily compared to stool smear testing. This is due to urine being more easily obtained, often in larger quantities and having no solid materials within it which may hinder the identification of eggs [12]. The test can often be used to grade the UGS infection's intensity by using the egg count from 10ml of filtrated urine [39]. Similarly, an important consideration is that egg excretion has been found to be highest between 10am and 2pm and can also increase the sensitivity of the test when taken between these times [40]. Urine egg detection techniques often are suitable for the diagnosis of individuals in areas with high *Schistosoma* infection levels; however, very light infections can lead to misdiagnosis [12].

ii *Diagnosing IS by Kato-Katz thick smear technique*

A method recommended by the WHO for diagnosing an individual for *S. mansoni* infection is using the microscopy-based Kato-Katz (KK) thick stool smear technique which helps visualise and quantify the individual intensity of infection. The KK technique has a high level of specificity, is easy to carry out and requires inexpensive and unskilled labour in field conditions [12]. The intensity of infection is measured based on the quantity of schistosome eggs present in the stool or urine samples; the higher intensity of *Schistosoma* infection can be a predictor for the worm burden and morbidity of an individual [41]. However, despite the high specificity the KK test often is found to have a lack of sensitivity [42]. For instance, when single samples are taken and in particular when there is a low-intensity infection present in an individual or low prevalence in an area, the sensitivity of the test can be greatly reduced [43, 44]. Therefore, treatment for the individuals living in these areas can be inappropriately given and lead to undetected transmission in these areas [44].

Further, the KK test has other limitations; these include the fact that the amount of schistosome eggs excreted through stool or urine vary throughout the day, the adult worm production also varies, and eggs can be clumped together within stool samples [12]. This variation has been shown to affect the sensitivity of the test, particularly with low intensity infections and where previous chemotherapy treatment has taken place [45–47]. Taking repeated stool samples over a period of consecutive days can be used to improve the KK test sensitivity; however, this method is often not practical in field settings, increasing the overall cost and reducing compliance with participants in studies [12, 48].

iii *Eggs count in Stool or Urine to grade infection intensity*

In the case of UGS, eggs are detected within a urine sample to identify the presence of *S. haematobium* infection, whereas for IS the detection of eggs in a stool sample can be used to identify the presence of *S. mansoni* infection. The egg count can be used to grade infection intensity based on WHO guidance (Table 1.2) [49]. Identifying the presence of infection based on the presence of schistosome eggs, as per the parasitological method, is only worth carrying out after the individual has been infected for 4 to 6 weeks; after this time oviposition occurs which is the process of egg laying for the schistosomiasis life cycle. Hence, detecting eggs in stools and urine is not suitable for the early diagnosis of schistosomiasis before the adult worms have matured and started producing eggs.

Table 1.2: Infection intensity [49]

	Light intensity	Moderate-intensity	High intensity
<i>S. haematobium</i>	< 50 eggs/10mL		\geq 50 eggs/10mL
<i>S. mansoni</i>	1-99 epg	100-399 epg	\geq 400 epg

iv Concentration methods for diagnosis of IS

Urine and stool samples can be concentrated by using the Visser and Pitchford concentration method where debris within the sample is filtered out while retaining the schistosome eggs for visualisation under a light microscope. This method allows for a more sensitive method for diagnosis of *Schistosoma* infection compared to other traditional microscopic methods. This concentration method allows for grouping multiple faecal samples to confirm the presence or absence of *S. mansoni* ova before assessing the intensity of infection. The method can often be impractical in field settings, as often only one stool sample can be taken [50].

1.2.9.2 Indirect diagnostic methods

i Clinical presentation

Alternative indirect methods can help to diagnose an individual with UGS: this can be through observing their clinical presentation instead of using direct methods. This can include looking at changes in appearance, concentration, and content (urinalysis) of an individual's urine. Blood (hematuria) and protein (proteinuria) in urine also can give identification of UGS infection [12]. In cases where an individual has a light infection, often diagnosis based on clinical

symptoms will lead to un-diagnoses, as the test lacks specificity and has a low sensitivity [41]. Clinical community studies have used these known association indicators to apply a combination of rapid-detection reagent strips which can detect these indicators within urine and questionnaire surveys asking for demographic information to diagnose individuals in the field [12].

Another diagnostic method based on micro-haematuria in urine is urine heme dipsticks known as Point-of-care (POC) test strips (e.g., Haemastrix). Previous studies have found that this method for detecting micro-haematuria has a higher sensitivity than patient reported haematuria cases when diagnosing UGS among SAC in endemic areas [51]. These diagnostic methods can be used to monitor the effect MDA treatment programs for *S. mansoni* infection have, as higher urinary tract morbidity is associated with visible haematuria; and if control methods are working, haematuria is visualised less [52].

Similarly, for IS, clinical presentation of distended abdomen can be associated with chronic infection of *S. mansoni* infection. However, in endemic areas, often unassociated nutritional disorders can affect individuals which can lead to clinical manifestations such as distended abdomen. This can lead to individuals being presumed to have IS, which can give an incorrect clinical picture of the area unless further diagnostic tests are carried out. Furthermore, many individuals with *Schistosoma* infections for both UGS and IS can remain asymptomatic for a considerable length of time, and hence clinical presentation is often only useful as a diagnostic marker in place of chronic infections [32, 53].

ii Rapid diagnostic tests

Immunodiagnostic techniques using rapid diagnostic tests allow for detection with light or no eggs infections before worms reach fecundity and as well as early detection. Antigen detection can be used as a very effective method in diagnosing an individual for *Schistosoma* infection. Schistosomula, the adult worms, or their eggs secrete antigens, which can be detected in blood (3 weeks post-infection), urine or sputum (by use of ELISA methods) of the individual and can be used as diagnostic targets. Circulating antigens are commonly detected otherwise known as circulating cathodic antigen (CCA) and circulating anodic antigen (CAA). These two glycoprotein antigens are known to be released by the adult worms, via their waste products. CCA and CAA antigens can be used to assess the intensity of infection and for treatment purposes [12]. The levels of circulating antigen found are often higher in individuals with higher schistosome egg counts (higher intensity of infection) and, outside endemic regions, where a low false positive rate occurs, which suggests the antigen detected is specific to schistosome infection [54].

For IS caused by *S. mansoni*, the POC CCA detection in urine is thought to be a very effective,

rapid diagnostic test and more sensitive than the KK test [55]. The POC detection in urine test has been found to have high sensitivity in cases ($\geq 95\%$) with moderate to high *S. mansoni* intensity infections. Similar to the KK test, the sensitivity of the CCA test for diagnosing IS is lowered ($\geq 75\%$) in areas with low *S. mansoni* infection rates. The specificity remains around $\geq 95\%$ [12, 56]. However, the CCA test for diagnosing *S. haematobium* has a limited accuracy for diagnosing UGS and has been noted to possibly be incompatible in areas where co-endemics occur with IS and UGS infections. The CCA strip test has been improved for successful diagnosis of UGS and is effective in areas with average to high levels of infection. Therefore, the CCA test with a positive result is strongly linked to the individual's infection intensity [12].

iii Serological and molecular methods

Advances in research have been developed to tackle the reduced sensitivity often occurring with diagnostic tests in areas with low infection intensity. Novel tests include monoclonal antibody-based diagnostic assays in serum to detect *S. mansoni* infection using immunomagnetic separation and fluorescence microscopy techniques. Despite the higher sensitivity and specificity of the serological method within areas of low infection prevalence, they cannot discriminate between active and past infections [57, 58]. Therefore, serological methods are not useful in endemic areas where people have had repeated exposure to *Schistosoma* infection, however, they can be useful with travellers [17].

In addition, molecular tests can be used for high sensitivity in areas with low infection intensity. These include using polymerase chain reaction (PCR) techniques to detect the schistosome Deoxyribonucleic Acid (DNA), cell-free parasite DNA in host urine, stool or organ biopsy samples or in other body fluids (e.g., sputum) [12]. These tests allow for high sensitivity and specificity in areas with low infection intensity, however, these tests are often difficult to apply in the field as they require specialised equipment, training, and facilities [41, 59]. Therefore, these serological and molecular tests are mostly used in wealthy areas, for instance, western travel clinics for travellers returning from visiting a schistosomiasis endemic region [60].

iv Research development in diagnostics

Other novel methods, including imaging devices (mobile phone or web cams) are also being used more often as mini-microscopes which can be useful in low-income areas. In some cases, there is clinical evidence of an individual having schistosomiasis but not yet a confirmed positive result from a diagnostics test; in this instance, organ biopsies and imaging techniques may be required for use when urgent diagnosis is needed to treat complications [12].

1.2.10 Treatment and preventive chemotherapy

The recommended treatment for *Schistosoma* infection is preventive chemotherapy, using PZQ [14]. PZQ was first discovered in the 1970s, and by 1988 had been tested for safe and effective use against schistosomiasis and put on to the market. The drug PZQ can be taken orally, is considered safe and relatively low cost and has been found to be effective against a mature *Schistosoma* infection (adult worms) [5]. Recovery rates between 60% and 90% have been reported and it is known to reduce egg excretion in cases with persistent infection [61]. The PZQ treatment is thought to be so effective that it has been found to work within an hour of consumption and a single dose is usually considered sufficient to kill all adult worms. However, there is some debate about whether higher or numerous doses are required [62]. The WHO recommends a certain dosage of the PZQ drug to ensure the best efficacy, rate of cure and egg reduction rate [14]. In the field, often PZQ dose height poles are used to help healthcare workers treat children [23]. SAC infected with *Schistosoma* infection benefit significantly from treatment with PZQ in curing infection or reducing worm burden [63]. The full drug mechanism is still currently being researched.

Despite PZQ being a well-tolerated drug, there are common side effects, including headache, gastrointestinal upset, dizziness and in some rare cases more serious adverse effects can occur [61]. In addition, there are disadvantages of the drug PZQ for instance, not being able to target the other stages of the life cycle of *Schistosoma* infection and low efficacy against immature schistosomes. Therefore, PZQ is most effective six weeks after an individual is infected and once the adult worms have matured and resided in their final location [64]. Moreover, it has been reported that high treatment use reduces the effectiveness of the PZQ treatment. This has been found to be due to the *Schistosoma* often developing reduced sensitivity to the drug, as well as schistosome building up a tolerance to the drug over repeated treatments. Evidence of drug resistance occurring with the drug PZQ has yet to be shown in the field conditions despite a study in 1994 demonstrating PZQ resistance to *S. mansoni* infection when tested on mice [65, 66].

Another drug administered for the treatment of *Schistosoma* infection is artemisinin (ART). This drug is used only as a preventive form of treatment and mainly used in high-risk groups; for instance, travellers and migrants who are entering endemic areas momentarily. ART works by protecting an individual for a 3-week period and has the ability to kill all immature schistosomula if consumed every 14 days. It is thought the drug ART works in producing a toxic effect to the schistosomula although the full mechanism is still not understudied. Treatment using both ART and PZQ, as a combined therapy has been studied to identify if there is an increased number of individuals cured from *Schistosoma* infection within endemic areas. In addition, corticosteroids can be used to treat acute schistosomiasis by reducing the inflammation within the infected individuals

tissues and reducing or stopping the formulation of granulomata in response to schistosome eggs. Treatment for advanced disease is most often the drug PZQ as it can stop or prevent complications such as pulmonary hypertension and stop the scarring of the tissues occurring (fibrosis) [5].

1.3 Control of schistosomiasis

1.3.0.1 Control to target the parasite

Implementation of control strategies are restricted by the resources available, the financial funding, the education of health workers and political support. New guidelines and plans for the Neglected Tropical Disease (NTD) roadmap have been formed for 2021–2030, created by the WHO and their partners. Previously, in the 2020 roadmap, 75% of SAC children were covered with preventive chemotherapy. The new roadmap plan (2021–2030) for schistosomiasis, outlines extending chemotherapy treatment to all populations who require it and providing any medicines that are needed [67]. Furthermore, the WHO had previously proposed a goal of < 5% prevalence of heavy infections occurring in 2020 to keep the disease under control; by 2025, < 1% prevalence of heavy infections across treatment locations and by 2030, across all endemic locations. Therefore, it is targeted that by 2030 schistosomiasis will no longer be considered a public health concern (defined as < 1% proportion of heavy intensity schistosomiasis infections) [63].

The new guidelines set by the WHO are evidence-based recommendations that try to interrupt transmission, leading to elimination of schistosomiasis as a public health concern and are designed to eliminate the morbidity burden from the disease. For successful control and possibly elimination of the disease, extensive monitoring of the disease and applications of multiple different control interventions are required. Despite several extensive implementations of control programmes to date, the burden of schistosomiasis is still very high, with the disease still endemic in many areas. The reason for the lack of successful control is thought to be due to the lack of accurate diagnostics for case detection and the inability to screen the communities in schistosomiasis endemic areas [12]. Since 2006, WHO guidelines school-based and community-based preventive chemotherapy programmes have been scaled-up, reducing schistosomiasis-associated morbidity, however, despite this, in certain areas *Schistosoma* infection has remained high, with persistent hotspots for transmission. The new guidelines suggest a wider approach targeting not only SAC but also pre-school-aged-children (PSAC), adults and pregnant women. They also recommend targeting not only high-risk communities but also low and moderate risk ones [68].

The new WHO guidelines have six new revised recommendations for managing schistosomiasis

including the expansion of preventive chemotherapy eligibility from the predominant group of SAC to all age groups (2 years and older), lowering the prevalence threshold for annual MDA treatment, and increasing the frequency of treatment [68]. The first three recommendations are a control treatment strategy for schistosomiasis based on the current prevalence of infection in endemic areas which depend on MDA programmes and monitoring of SAC infections (Table 1.3) [63]. The fourth recommendation is targeted to who and when people are treated; health facilities provide access to treatment with praziquantel to control morbidity in all infected individuals regardless of age, including infected pregnant (excluding the first trimester) and lactating women and PSAC aged 2 years. Further, the decision to administer treatment in children under 2 years of age should be based on testing and clinical judgement. The fifth recommendation encompasses water, sanitation, hygiene (WASH) interventions, the control of *Schistosoma* infection's intermediate snail host, and changes to human behaviour; for instance, using more appropriate places than the lake to dispose of urine or faeces.

Finally, the sixth recommendation is in communities where no autochthonous transmission has occurred in 5 years: testing for *Schistosoma* infection in humans, the intermediate host snail and non-human mammals is required to have higher sensitivity and specific tests, to ensure low-intensity infections are still identified. A two-stage verification process has been suggested by the WHO, where a highly sensitive test is carried out first before a high specific test is carried to establish the *Schistosoma* infection status. The success of these six recommendations is largely limited by community distribution of the MDA treatment, funding and resources, MDA effectiveness and lack of education for communities [63].

Table 1.3: Treatment strategy recommended for schistosomiasis with preventive chemotherapy [63]

Category	Prevalence <i>Schistosoma</i> infection	Recommendation
High-risk community	$\geq 50\%$	Lack of an appropriate response to annual preventive chemotherapy, despite adequate treatment coverage ($\geq 75\%$), WHO suggests consideration of biannual (twice yearly) instead of annual preventive chemotherapy.
Moderate-risk community	$\geq 10\%$	Annual PZQ treatment with single dose at $\geq 75\%$ in all age groups above 2 years of age including adults, pregnant women after first trimester and lactating women
Low-risk community	$< 10\%$	In cases, where there has been a programme of regular preventive chemotherapy, to continue the intervention at the same or reduced frequency towards interruption of transmission; or (ii) where there has not been a programme of regular preventive chemotherapy, to use a clinical approach of test and-treat, instead of preventive chemotherapy targeting a population.

1.3.0.2 Control of the snail populations

The fifth recommendation mentioned targeting the *Schistosoma* infection's intermediate snail hosts to reduce transmission. Freshwater snails are essential for schistosomiasis transmission as they are the intermediate host of the parasite, and a single snail can release thousands of cercariae into the freshwater. Therefore, control strategies to manage snail populations has been created [63, 69]. The use of chemicals to kill the intermediate host snails in freshwater sources have been proposed as a control method; however, this can lead to other species of animals living in the water being harmed, and repeated treatment with chemicals is required over a long period of time to avoid the snails repopulating the water sources. In addition, since livestock can be infected with schistosomiasis, runoff from pastureland may occur, which can allow schistosomes to contaminate otherwise disease-free water sources [70]. Other control methods include modifying the habitat to reduce snail breeding sites, or using biological control such as molluscivorous fish or crustacean-eating birds in the area to reduce the number of snails. The problem with these methods is that they can be expensive, and it can be difficult to sustain their effect on the predators [69, 71, 72]. Genetic modification has been explored allowing for detection of latent schistosomiasis infection and identification of the resistance gene in the snail host, however more research needs to be completed to better understand the genetic structure of the snails [73].

1.4 Study population: Lake Malawi and the Mangochi District

1.4.1 Geographic and demographic context

Malawi, based in southeast Africa, is surrounded by land, and shares its borders with Tanzania, Zambia and Mozambique (Figure 1.4). Malawi is split into three regions called Northern, Central and Southern. The geographical landscape of Malawi varies depending on the region, with mountains in the Southern region and high plains in the North and Central regions. Malawi has diverse features ranging from tropical rainforests, open grasslands, and scrubland at high altitude, to woodland containing a large variety of fauna and flora. The seasons in Malawi consist of wet and hot in November to April and cold and dry in May to October where the temperature varies between around 14 to 32°C. Malawi is a low-income country where malnutrition, poverty and disease are rife throughout the country which leads to reduced life expectancy, lack of education, and hindered economic development [74]. Malawi's population is around 17.6 million and is growing at a rapid pace (~ 2.9% per year) as there is a high fertility rate due to a lack of family planning resources [75]. Two fifths of the population are young individuals ranging between 10 and 29 years, which in theory is good for the economic success of the country as the young contribute most to economic development. However economic growth has been hindered by a mixture of factors such as low-income levels, lack of resources, a growing overall population, and sever morbidity and mortality rates.

Malawi's east side is mainly covered by Lake Malawi, which is 600km long, 75km wide at the widest point and is second deepest lake in Africa [76]. It is widely known as a natural resource for irrigation for agriculture, water supply, fishing industries and tourism [74]. The lake is the habitat of the intermediate snail host, which allows schistosomiasis to occur. *Schistosoma haematobium* and *S. mansoni* have varying transmission throughout Malawi (Figures 1.5 and 1.6) also known as having a heterogeneous distribution of transmission and this can lead to higher transmission in focal areas across Malawi. There are visible geographical and seasonal variation of snail distributions due to the changing ecology of Lake Malawi, which in turn, can affect the survival of the intermediate host snails and cercariae. For instance, the degree of the wind and wave exposure changes throughout the year in Lake Malawi; During the wet and hot season in December to April and part of dry season April to July, the transmission of schistosomiasis is thought to be low whereas the transmission is higher during the cold and dry season in May to October, peaking in October [3].

The southern part of Lake Malawi in Mangochi District has had a reported increase in schistosomiasis transmission since the 1980s. This is thought to be due to overfishing which has reduced the

number of molluscivorous fish and allowed the number of snails to increase in quantity. A ban on fishing within 100m of the shoreline of Lake Malawi National Park has been enforced in the southern part of Malawi. However, this is often ignored as seine or gill nets have been observed despite the ban [76]. Mangochi District has a population of over 60,000 individuals and the area is known to be endemic for UGS caused by *S. haematobium*, with recent surveys at schools suggesting a prevalence of 0 to 26%.

National Schistosomiasis Control Programmes (NSCP) (2011–2016) were created following the success of the Malawian schistosomiasis control schemes starting in the 1960s [77]. NSCP has always focused on the distribution of chemotherapy and education within schools and communities to try reducing the transmission of schistosomiasis and more recently the approach has changed to focus on prevention, surveillance and control alongside other NTDs. Epidemiological mapping of *S. haematobium* and *S. mansoni* has allowed NSCP to predict that 40 to 50% of Malawi's population are at risk of schistosomiasis. However, these estimates are thought not to be up to date as the data is based on previous high risk school surveys [3]. In addition, National Schistosomiasis and Soil-Transmitted Helminths Control Programme run by the Ministry of Health has been created [78]. This program was integrated into the School Health and Nutrition Programme, an initiative led by the Ministry of Education and the Ministry of Health to aid MDA among SAC. In 2015, the NTD Master plan for 2015–2020 was formed for Malawi, with the target to transform Malawi into a nation free from NTDs by 2020, including schistosomiasis. Furthermore, the master plan encouraged an integrated approach among the most prevalent endemic NTDs in Malawi to accelerate the implementation of NTDs prevention and control programmes.

Malawi has achieved high MDA coverage in the targeted districts, however, the prevalence of the disease remains high in most districts. As mentioned before, the WHO roadmap 2021–2030 for schistosomiasis and other NTDs targets elimination as a public health concern (defined as < 1% proportion of heavy intensity schistosomiasis infections) through a combination of control methods including more targeted MDA, WASH, environment and snail control [63, 67]. It is unlikely that the goal of reducing the burden of schistosomiasis and STH to levels of no public health importance in Malawi by 2025 will be reached. This is due to the country's reliance on MDA programmes, its lack of community engagement, and the failure to put more focus into the other methods of interventions [79].

1.4.2 Urogenital Schistosomiasis in Mangochi District

UGS is prevalent throughout Malawi compared to IS which is thought to be limited to the central and southern highlands, Lower Shire and Likoma Island. The intermediate host *Bulinus spp.* snails

have often been found in Lake Malawi, whereas before 2017, no *Biomphalaria* sp. snails have been reported to be found in the lake.

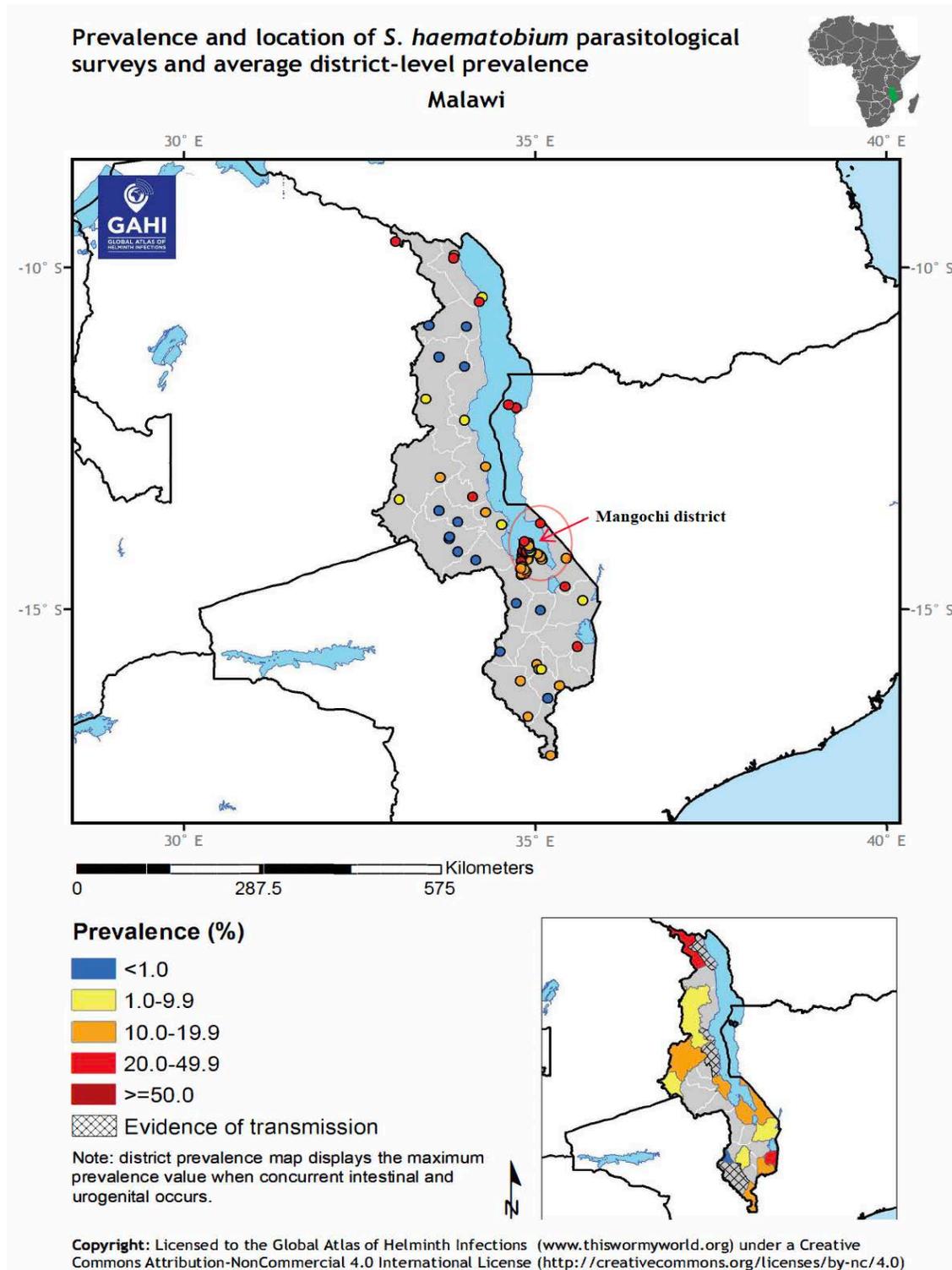


Figure 1.5: The average district prevalence (%) of *S. haematobium* from survey data collected in Malawi. Dots colours stand for prevalence (%): Blue (< 1.0), Yellow (1.0 – 9.9), Orange (10.0 – 19.9), Red (20.0 – 49.9), Brown (> 50.0). Red circle is approximate location of Mangochi district. Reproduced from Global Atlas of helminth infections, 2018 under a CC-BY license [80].

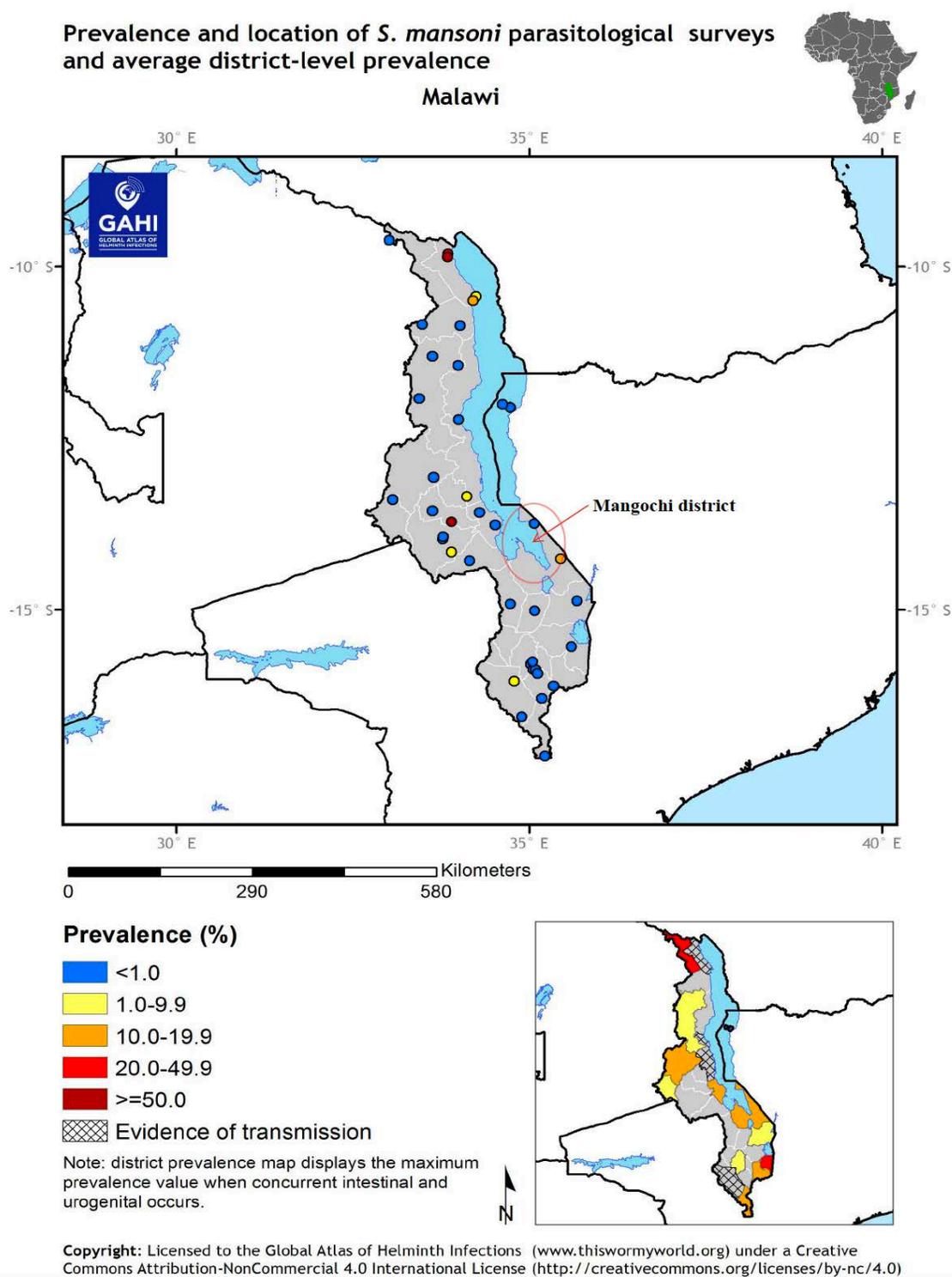


Figure 1.6: The average district prevalence (%) of *S. mansoni* from survey data collected in Malawi. Dots colours stand for prevalence (%): Blue (< 1.0), Yellow (1.0 – 9.9), Orange (10.0 – 19.9), Red (20.0 – 49.9), Brown (> 50.0). Red circle is approximate location of Mangochi district. Reproduced from Global Atlas of helminth infections, 2018 under a CC-BY license [80].

1.4.3 Changing epidemiological landscape of schistosomiasis in the Mangochi District

Schistosomiasis transmission is dependent on the presence of compatible snail species to transmit the disease. The geographical locations of the *Bulinus* and *Biomphalaria* snail species are important in order to predict the transmission areas of *S. haematobium* and *S. mansoni*. The increasing movement of human populations and their overlap with the snail habitats has allowed the spread of schistosomiasis to new locations with cercariae being sufficiently adapted to survive in new locations, and it has an opportunistic nature, which means that even short time exposure can lead to *Schistosoma* infection [81].

From 2017, *B. pfeifferi* snails have been found along the Mangochi District shorelines of Lake Malawi with IS positive diagnosis occurring within local SAC [11]. This transmission of IS indicates autochthonous transmission, which is where one individual is infected, and the transmission occurs in the same area. When *Bulinus* and *Biomphalaria* snails are present along this area they can cause a co-endemic of *S. haematobium* and *S. mansoni*. The true impact of a new species of schistosomiasis in an area (*S. mansoni*) on the prevalence of *S. haematobium* is unknown. Recently environmental and ecosystem changes along Lake Malawi have been linked to increased prevalence of UGS [82]. Furthermore, natural disasters, for example cyclone Idai in March 2019, have caused flooding which can change the ecosystems of the lake. Changes in lake water levels have been reported [11]. The numerous epidemiological changes around Lake Malawi in the Mangochi District are understudied and further investigation is required.

1.5 Snails and their microhabitat

The favoured habitat of the freshwater snails which transmit schistosomiasis varies greatly between species. This leads to the distribution of snails being often local and patchy due to their dependence on the habitat. Furthermore, often human populations tend to move about which makes it difficult to quantify which infections were due to imported infection as compared to local cases. These factors make it difficult to measure how often the snails are present and the distribution of the snails in a certain area [83].

Bulinus snails prefer shallow water with rich vegetation and in some cases can live outside water in a state of dormancy, although they can also be found in deeper waters and have links with temporary water bodies. *Bulinus globosus* are often found in rivers, streams, lakes, seasonal pools, irrigation systems, and artificial ponds. *Bulinus nyassanus* are nocturnal and stays buried within

sediment during the day, whereas *B. pfeifferi* snails are a tolerant species, often living near humans, and are linked to permanent water conditions and cannot live outside freshwater [84, 85]. *Biomphalaria pfeifferi* snails are often found to favour slow-moving streams, rivers, dams and ponds. The changing distribution and invasiveness of the *Biomphalaria species* is thought to be due to climate and human behaviour changes [85, 86]. The building of irrigation schemes, canals and hydroelectric dams can encourage the spread of the freshwater snails. Furthermore, overfishing in the waters for molluscivorous fishes *Trematocranus placodon* can lead to an abundance of freshwater snails present [85].

Environment characteristics can substantially vary between areas with short distances having larger variation [83]. Snail density can be affected by numerous factors including the presence of vegetation, rainfall, the depth and flow of the water and food supply [8]. The presence of vegetation provides the food supply and shelter for freshwater snails, as well as providing them an area to depositing their eggs [87]. *Biomphalaria pfeifferi* freshwater snails prefer *Vassisneria* plants [11]. Rainfall patterns have an impact on water levels and have a possible effect on the prevalence of schistosomiasis and the freshwater snails' distribution. A study, carried out in Ghana showed that higher overall annual rainfall was associated with an increase in the prevalence of schistosomiasis in an area and vice versa, for low overall annual rainfall [88]. More rainfall leads to an increase in ponds and high water levels and can lead to an increase in snail breeding sites due to more runoff into irrigation channels and faster flowing water increasing the chances for the parasite to make contact with the freshwater snails. However, higher water levels often result in more turbulent waters leading to faster flowing water which can agitate the snail populations and reduce the ability for cercariae to survive. Flooding has also been found to be associated with an increase in the abundance and distribution of freshwater snails; often the snails are found in new or previously eliminated areas. When drought occurs, there can be a decrease in freshwater snails and transmission of schistosomiasis. Droughts over 9 months can lead to death of the infected freshwater snails and so a decrease in transmission, whereas in droughts lasting less than 7 months, certain areas will increase in transmission due the ability for the larval to survive over this time [85].

Climate change over time is thought to contribute to the changes in freshwater snail abundance and locations [85]. Chemical characteristics such as pH (presence of hydrogen ions and acidity of the water), temperature, conductivity and saline concentration of the water could also contribute to the density of snail distribution, however, further investigation is required. Firstly, temperature of the water has been found in research to be a factor in the abundance of snails found within an area as it possibly affects the growth, survival, fecundity, distribution, and breeding conditions, and hence, an optimal temperature is required for snail development and survival to occur [83] [89]. $25^{\circ}C$ is considered optimal as when the temperature surpasses $30^{\circ}C$ the snails start to die off and $40^{\circ}C$ is

known to be fatal [90] [83]. In addition, the temperature of the water also affects other parts of schistosomiasis transmission, including the rate at which miracidia penetrate the freshwater snails, the release of cercariae into the freshwater and its relative skin penetration of the human host [85]. Secondly, the pH of the water has a possible effect on biological population dynamics and transmission of schistosomiasis. In most studies, pH is not associated with the snail abundance, however, Levitz *et al.* 2013 [91] and other studies have found that a lower pH was associated with higher snail abundance [83, 91]. *Biomphalaria* snails have been studied which has been shown to influence the pH with optimal pH ranges between 7.0 and 9.0 for snail populations, whereas there are no studies on *Bulinus* snails. Therefore, pH association with snail abundance still needs to be studied further to find the true relationship. Thirdly, high conductivity, has been suggested to have a reduction in the snail abundance found in a certain area [90]. Lastly, high saline concentration in the water has shown to have some declining effect on some species of freshwater snail, however, research so far suggests most freshwater snails are tolerant to changes in saline or do not live in saline water and more focus has been on cercariae who have a reduction in production with lower saline concentration [85].

These seasonal changes in water levels and biotic changes in the water due to climate change or human activity have altered the transmission of schistosomiasis. Furthermore, freshwater snail surveillance is considered important as it is essential for the transmission of schistosomiasis to occur. Due to the schistosome infections' ability to survive within a human for a long time, this can make it hard to determine when and where an individual was infected. This is where snail surveillance comes into play to help tackle this difficulty [83].

1.6 Methods overview

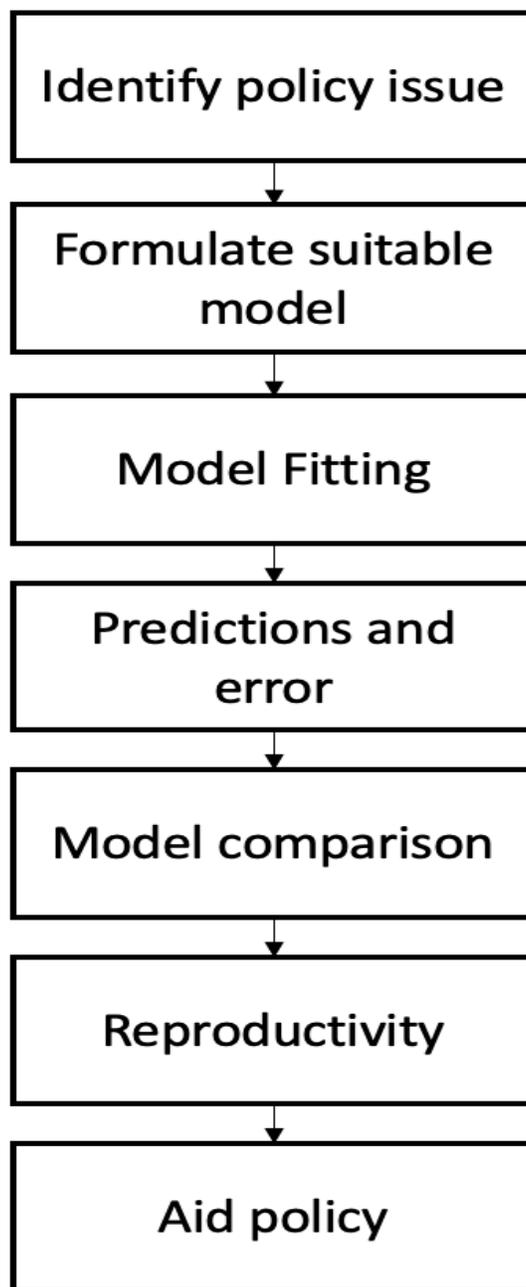


Figure 1.7: A diagram of the process of policy makers and modellers working together to implement improved control programmes in practice, reproduced from Behrend *et al*, 2020 under a CC-BY license [92]

1.6.1 Modelling and statistical techniques

Modelling is a continuous process which requires models to be updated regularly, re-fitted or modified depending on the scientific evidence and techniques available for input (Figure 1.7). Models

are never considered complete and constant collaboration between the model creators and policy makers is essential for effective implementation of control strategies. Further, a desirable part of the modelling process is the comparison between different models for the disease by various individual modelling groups [92].

Statistical models have allowed for uncertainty to be considered within estimates and can be deterministic and non-deterministic. This uncertainty can include random effects and considerable variation in data [93]. Statistical modelling methods to assess the spatial epidemiology, and hence identify areas at risk of infection, are gradually being used more often in replacement of traditional investigative methods, for instance immunology [94]. Furthermore, they allow for a reduction in required sample data, and hence save cost and time required for prevalence surveys. Statistical techniques can include logistic regression, generalised linear models (GLMs), generalised additive models (GAMs), Bayesian statistical models and spatial statistical modelling. These allow us to predict unobserved data using environmental and socio-economic predictors and infection risk based on observed data. Previously *Schistosoma* infection risk has been mapped using model-based geostatistics by a spatially continuous prevalence of infection [95]. Advances in spatial information techniques have allowed for the use of remote sensing, geographical information systems and global positioning systems (GPS). These methods can help towards resource allocation and allow for cost-effective control methods to be applied [94].

1.6.2 A basic overview of Generalised linear models/Generalised additive model/Bayesian linear models

Nelder and Wedderburn *et al.* [96], formulated GLMs. We assume y_1, y_2, \dots, y_n are independent observations from independent and identically distributed random variables Y_1, \dots, Y_n conditional on d covariates in vectors $\underline{x}_1, \dots, \underline{x}_n$. i.e.

$$y_i \sim Y_i | \underline{x}_i, i = 1, \dots, n. \quad (1.3)$$

We then assume the equation $\mathbb{E}(Y_i | \underline{x}_i)$ of $Y_i | \underline{x}_i$ is related to the function of a linear combination of covariates \underline{x}_i and coefficients $\underline{\beta}$, i.e.

$$\mathbb{E}[Y_i | \underline{x}_i] = \mu_i = g^{-1}(\eta_i), \quad (1.4)$$

where g is the “link” function and

$$\eta_i = \sum_{j=1}^d x_{ij}\beta_j, \quad (1.5)$$

is the linear predictor and x_{ij} is the j th component of \underline{x}_i . For example, suppose y_1, \dots, y_n are Poisson distributed, we have :

$$y_i \sim \text{Poisson}(\mu_i), \quad (1.6)$$

$$\mu_i = e^{\eta_i}, \quad (1.7)$$

$$\implies \log(\mu_i) = \eta_i. \quad (1.8)$$

GAMs are similar to GLMs, however, there is a non-linear relationship for the data, which provides a way of smoothing out the estimation where otherwise noisy data occurs. Hastie and Tibsh *et al.* 1986 [97], created GAMs from a combination of GLMs and additive models. Where the linear predictor (η) is replaced with an additive one, which can be represented as,

$$\eta_i = \sum_{j=1}^d x_{ij}\beta_j + s(z_i), \quad (1.9)$$

where $s(z_i)$ is considered the smooth function.

Statistical inference is usually carried out using a frequentist approach where the unknown parameters are drawn from the observed population and predicted values are calculated. Another approach used can be Bayesian, including Bayesian linear models (BLMs) where the unknown parameter values are uncertain. These unknown parameters can be considered the random variables where we use probability to make inference. Bayes theorem is behind Bayesian statistics, which is the conditional probability of the likelihood of an outcome happening again based on previous outcomes. The theorem gives the posterior distribution (the predicted outcome) using the prior distribution (the relative weighting you give to each parameter) and the observed data to enable inference [98].

Furthermore, spatial statistical models allow for spatial autocorrelation (residual spatial correlation) to be considered. Where there is residual spatial correlation, points close together are more likely to have similar risk than those far apart [99]. Model-based geostatistics can usually be ap-

plied to Bayesian statistics modelling which allows for the inference of uncertainty in parameter estimates and predictions [100].

1.6.3 A basic overview of deterministic and stochastic SIR/SEIR/SEIRS-type compartmental model

Mathematical models can be used to draw aspects of real-world complex systems given population data and input parameters. McDonald, 1911 [101] established the modern mathematical modelling approach with a set of equations to approximate continuous-time dynamics of a disease through transmission [101]. Deterministic modelling shows how output has changed given by the values inputted, first reported by Ross, Kermack and Mckendrick [102–104]. These papers suggested that the probability of infection of a susceptible individual is analogous to the number of its contacts with infected individuals and created the susceptible-infected-recovered (SIR) model as shown in Figure 1.8. Models have also incorporated over time the exposed part of transmission giving the susceptible-exposed-infected-recovered (SEIR) model and the ability of individuals to move back from recovered to susceptible, which is modelled as susceptible-exposed-infected-recovered-susceptible (SEIRS). We explain examples of this in relation to schistosomiasis in more detail in the next section (1.8.4). Frost, 1976 [105] introduced the assumption that the infection spreads from an infected individual to a susceptible individual through discrete time Markov chain events [105]. Frost’s paper set the overall basis of stochastic epidemic modelling where the range of the outcome was produced through many simulations.

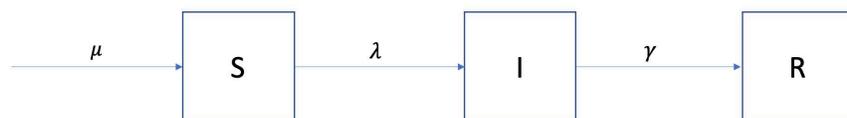


Figure 1.8: A compartment model representing the stages of infection; Susceptible (S), Infection (I), Recovered (I). μ represents the birth rate, λ represents the rate from S to I and γ represents the rate from I to R.

1.6.4 Background of previous schistosomiasis modelling and how it relates to the thesis

Both Nelson & Macdonald carried out pioneering work on transmission dynamics of *Schistosoma* infection [106–108]. Nelson developed survival tables for each stage of the *Schistosoma* infection lifecycle and was able to calculate the fertility of the adult and larva worm. The study was the first to link the relationship between population ecology and epidemiological study of infectious diseases. Macdonald, 1965 [108], was based on differential equations to show the changes over time for worm burden and studied the impact separate sexes on the parasite population growth and decay. Anderson and May, 1982 [109], expanded on Macdonald’s work and used partial differential equations to model changes over time. In addition, it studied the age of the human population worm burden, using for instance, egg production and acquired immunity based on previous exposure [109].

More recently, Anderson and May, reviewed transmission models and control of *Schistosoma* infection by Mass Drug Administration (MDA), assessing control programmes, individual based stochastic models to study both infection and drug compliance, fitting models to observed patterns using statistical approaches to find the likelihood of outcomes and to understand the dynamics of transmission. Furthermore, the spatial element of heterogeneity of transmission of schistosomiasis has also been studied [110]. Many mathematical models have been created over time to study the transmission dynamics, and control of the spread of schistosomiasis, which in turn has aided decision makers in applying appropriate control methods and helped work towards WHO 2030 goals in eliminating schistosomiasis [111, 112]. The World Health Organisation (WHO) has supplied new guidelines to target elimination of schistosomiasis by interrupting transmission [63]. Mathematical models have been used to study the direct transmission of SAC over time, and hence, explore their transmission dynamics [113–115]. Intrinsic factors such as age and their related exposure have a possible importance in interrupting transmission when moving towards elimination of the disease [116]. However, mathematical models often have many limitations and assumptions so can be limited in the ability to produce a realistic picture of the transmission dynamics.

Mathematical models of the transmission dynamics usually include humans, snails and the free-living stages of miracidia and cercaria [111, 114, 117, 118]. Other transmission models include humans, cattle and snails, however, this thesis focuses only on human and snail populations [119]. Kanyi *et al*, 2021 [111] uses all four types, adding treatment and preventive intervention water satiation and hygiene (WASH) to their mathematical models as shown in Figure 1.9 [111]. The study subdivides the human population into susceptible (S_h), exposed (E_h) and treated group (T_h), where the treatment group is in replacement to recovered group (R_h). In addition, they subdivided

the snail population into susceptible (S_s), exposed (E_s) and infected (I_s) snails and have a compartment for the free-living stages of miracidia (N_m) and cercaria (N_c). Lopez *et al.* 2024 [120] proposed if there was a vaccination available and added vaccination (V_h) in replacement to the recovered group (R_h). This study moved away from looking at the control and treatment strategies of a population and instead aimed to look at human behaviour and voluntary vaccines [120]. However, in this thesis we do not have data on whether the snails were infected or not. In addition, the thesis uses the concept that the children are a small fraction of the total population. Consequently, in this thesis we are not considering all forms of transmission occurring. As there is no reason given the epidemic assumption we have that the infection should come and go. We treat it as a constant reservoir of infection. If we were modelling the whole population, then you mostly likely would not be able to do this but as we have other adults and children which would be causing infection that we do not consider in our model, we consider them a constant reservoir of infection.

In relation to this thesis, NTD data is often highly sparse, and quantitative methods that allow us to smooth data and interpolate gaps in the data provide a useful tool in these situations, allowing us to build on the mathematical models previously researched. The main statistical methods used in this thesis are GAMs with thin-plate splines which provide a method of interpolating across noisy time-series data, capturing non-linear relationships in data in a smooth way. Then the geospatial analysis with Gaussian process approach again allows us to smooth over observations that are necessarily sparse in terms of space. Moreover, an ordinary differential equation (ODE) mathematical modelling approach allows us to smooth over periods of time in which we do not observe prevalence and also capture assumptions we might have about the disease case-generating process. These methods allow us to draw conclusions from our given sparse data to improve policy makers' decisions to implement control methods appropriately and effectively.

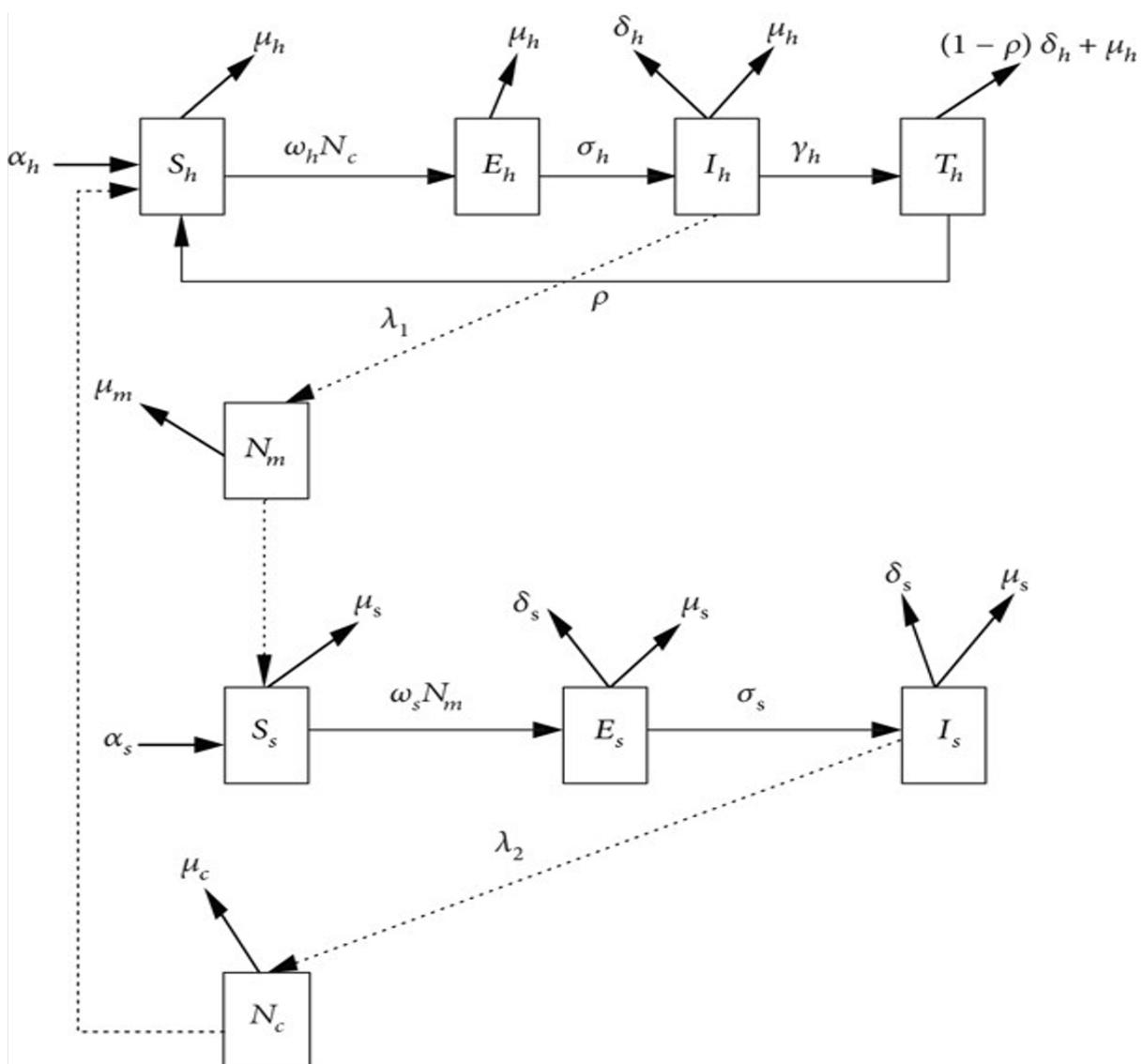


Figure 1.9: Compartment model reproduced from Kanyi *et al.* under a CC-BY license [111].

1.7 Aims and structure of this thesis

The aim of the thesis is to primarily focus on aiding the appropriate control of schistosomiasis along the shoreline of Lake Malawi in Mangochi district and help towards reaching the World Health Organisation's (WHO) neglected tropical disease roadmap for 2030 [67]. The project involves a secondary analysis of parasitological surveys and malacological data to explore the spatial and temporal epidemiology of *Schistosoma* spp. infection within SAC, who are considered one of the most vulnerable groups to *Schistosoma* infection.

To our knowledge, this is the first study to analyse the IS infection age-relationship with SAC along

the southern part of shoreline of Lake Malawi in the Mangochi District. Furthermore, a geospatial model will be used to identify where the SAC comes in contact with the intermediate snails and the abundance of freshwater snails along the shoreline, and also help to predict the un-sampled parts of the shoreline. A dynamical model was used to reconstruct the age-prevalence profiles with age related immunity allowing us to understand *Schistosoma* infection dynamics over time. Consequently, aiding control method applications and providing insight to interpreting *Schistosoma* infections to support the WHO roadmap. Therefore, the study is aimed at helping policy makers know where to target control programmes in Malawi and to limit the funding and resources required.

1.8 Objectives

This thesis aims to achieve five main Research Objectives in order to aid schistosomiasis programmes and inform policy:

Research Objective 1: Summarise current knowledge of transmission of Urogenital and Intestinal schistosomiasis within SSA and along the shoreline of Lake Malawi

This Research Objective aimed to provide a solid base to this thesis through Chapter 1: Literature Review, which reflects the most relevant information for the chapters that follow.

Research Objective 2. Investigating age-prevalence relationships for SAC along shoreline of Lake Malawi

For this Research Objective, a Generalised Additive Model (GAM) was used to predict the relationship between age of SAC and *Schistosoma* spp. infection prevalence using pilot parasitology data collected along the shoreline of Lake Malawi in Mangochi district. The GAM was used to investigate SAC age-profile for IS and UGS infection and co-infection at an individual and grouped school level. Results are shown in Chapter 2: SAC age-profiling of IS and UGS infection

Research Objective 3. Identifying the locations along the shoreline of Lake Malawi where children are most likely to be exposed to infection

For this Research Objective, a Bayesian multilevel model (MLM) with a Gaussian Process (GP)

was developed to analyse and predict the occurrence of snails along the lake shoreline. The main body of the research is reflected in Chapter 3: Predicting spatial epidemiology of *Schistosoma* spp. infection using malacological data. Further, this research objective was used to show whether there is heterogeneity in transmission along the shoreline of Lake Malawi and to identify persistent hotspots for transmission despite annual MDA programmes.

Research Objective 4. What are the main determinants of infection within SAC?

For this Research objective, a dynamical model was developed using SEIRS model encompassing Chapter 2 and Chapter 3 analysis output data and this follows Chapter 4: Development of a dynamical model to aid schistosomiasis control application within SAC to be produced.

Research Objective 5. Does school distance from the shoreline have an effect on exposure?

For this Research Objective, the dynamical model developed in Chapter 4 also measures whether the school distance from the shoreline has an effect on exposure to snail abundance.

Finally, the General Discussion chapter (Chapter 5) brings together and summarises all the chapters in this thesis. Key findings are used to draw conclusions and propose informed modifications to, as well as support for, parts of the WHO guidelines for the control and elimination of schistosomiasis.

Chapter 2

Modelling the age-prevalence relationship in schistosomiasis: A secondary data analysis of school-aged-children in Mangochi District, Lake Malawi

Amber L. Reed^{1,2}, Angus M. O’Ferrall², Sekeleghe A. Kayuni^{2,3}, Hamish Baxter², Michelle C. Stanton⁴, J. Russell Stothard², Christopher Jewell⁵.

Author Affiliations

1. Lancaster Medical School, Lancaster University, Bailrigg House, Bailrigg, Lancaster, LA1 4YE, UK
2. Tropical Disease Biology, Liverpool School of Tropical Medicine, Pembroke Pl, Liverpool L3 5QA, UK
3. MASM Medi Clinics Limited, Medical Aid Society of Malawi (MASM), P.O. Box 31659, Lilongwe 3. Malawi
4. Vector Biology, Liverpool School of Tropical Medicine, Pembroke Pl, Liverpool L3 5QA, UK
5. Mathematics and Statistics, Lancaster University, Bailrigg House, Bailrigg, Lancaster, LA1 4YE, UK

Keywords: *Schistosoma mansoni*, *Schistosoma haematobium*, Co-infection, Generalised additive models, Age profiling, School-aged-children

Abstract

Schistosomiasis is an aquatic snail borne parasitic disease, with IS and UGS caused by *Schistosoma mansoni* and *S. haematobium* infections, respectively. SAC are a known vulnerable group and can also suffer from co-infections. Along the shoreline of Lake Malawi a newly emerging outbreak of IS is occurring with increasing UGS co-infection rates. Age-prevalence (co)infection profiles are not fully understood. To shed light on these (co)infection trends by *Schistosoma* species and by age of child, we conducted a secondary data analysis of primary epidemiological data collected from SAC in Mangochi District, Lake Malawi, as published previously. Available diagnostic data by child, were converted into binary response infection profiles for 520 children, aged 6–15, across 12 sampled schools. Generalised additive models were then fitted to mono- and dual-infections. These were used to identify consistent population trends, finding the prevalence of IS significantly increased [$p = 8.45 \times 10^{-4}$] up to 11 years of age then decreased thereafter. A similar age-prevalence association was observed for co-infection [$p = 7.81 \times 10^{-3}$]. By contrast, no clear age-infection pattern for UGS was found [$p = 0.114$]. Peak prevalence of *Schistosoma* infection typically occurs around adolescence; however, in this newly established IS outbreak with rising prevalence of UGS co-infections, the peak appears to occur earlier, around the age of 11 years. As the outbreak of IS fulminates, further temporal analysis of the age-relationship with *Schistosoma* infection is justified. This should refer to age-prevalence models which could better reveal newly emerging transmission trends and *Schistosoma* species dynamics. Dynamical modelling of infections, alongside malacological niche mapping, should be considered to guide future primary data collection and intervention programmes.

2.1 Introduction

SAC are known to be one of the most vulnerable groups for schistosomiasis, which can lead to severe morbidity, and in some cases mortality. Standard infection and transmission rates in SAC are 3–4 times higher than in adults [17]. Children are thought to be first infected soon after birth upon freshwater contact(s), with prevalence increasing with cumulative parasite exposure(s) up to adolescence [63]. Over time, ongoing inflammation within the tissues, from accumulating trapped eggs, can lead SAC to suffer from malnutrition, anaemia, and neurological and developmental delays [121]. Furthermore, acute and chronic infection with urogenital schistosomiasis (UGS) and/or intestinal schistosomiasis (IS) can lead to debilitating symptoms and signs such as stunting,

but whether chronic co-infections are truly synergistic is equivocal [122].

To counter schistosomiasis, WHO recommend preventive chemotherapy by mass drug administration (MDA) with the anthelmintic PZQ. MDA treatment programmes can avert and reverse some of these disease manifestations as well as diminish transmission. However, PZQ is only effective against adult worms, leaving immature (drug tolerant) worms to remain within the body [17]. Since MDA does not guard against reinfection, SAC often reacquire infection upon subsequent water contact, with persistent “hotspots” occurring [3, 5, 23, 121]. As a consequence of ongoing persistent schistosomiasis infection among SAC, children are often absent from school, and have delayed learning affecting their ability to work as they enter adulthood [63]. This further hinders the socio-economic advances of a geographical area, a known risk factor for schistosomiasis [121].

A decrease in prevalence of infection is known to occur after young adolescence, which is typical of community age-prevalence relationships [63, 123]. This is thought to be due to the development of partial immunity over time given repeated exposure, as well as decreased contact with water or more enigmatic changes in skin texture, for example [17, 124, 125]. Prevalence among SAC and the wider population varies considerably between geographical areas, often with localised rates each community [17]. There are many factors that influence the transmission rates in a specific area, such as demographic and environmental factors, MDA, and snail-schistosome ecology [121].

Consequently, prevalence data can be noisy but pooling across schools allows for inferences to be extracted. The heterogeneity of transmission in a geographical area within a community influences the age at which prevalence and intensity of infection are at their highest in SAC, leading to some SAC being burdened more than others [53]. The identification of areas with high prevalence and intensity of infection is essential to allow for more appropriate application of control interventions, such as MDA [126].

The southern part of the Lake Malawi shoreline in Mangochi District has been reported to have increasing schistosomiasis infection rates since the 1980s, with known UGS endemicity in the region [76]. Al-Harbi *et al.* 2019 [11] and Kayuni *et al.* 2020 [127] reported the emergence and an outbreak of IS since 2017 in this region, in part due to the newly detected presence of *Biomphalaria*, a keystone snail intermediate host for *Schistosoma mansoni*. They suggested better inspection of age-infection dynamics is needed before intensification of current control methods is advised. Similarly, with the recent endorsement of urine-CCA testing for prevalence mapping of IS [56], closer inspection of infection data by age would further underpin its guiding role. To our knowledge, however, there are no studies that have analysed the age-prevalence relationship of IS within SAC in the context of a newly emerging focus of infection set against a background of UGS.

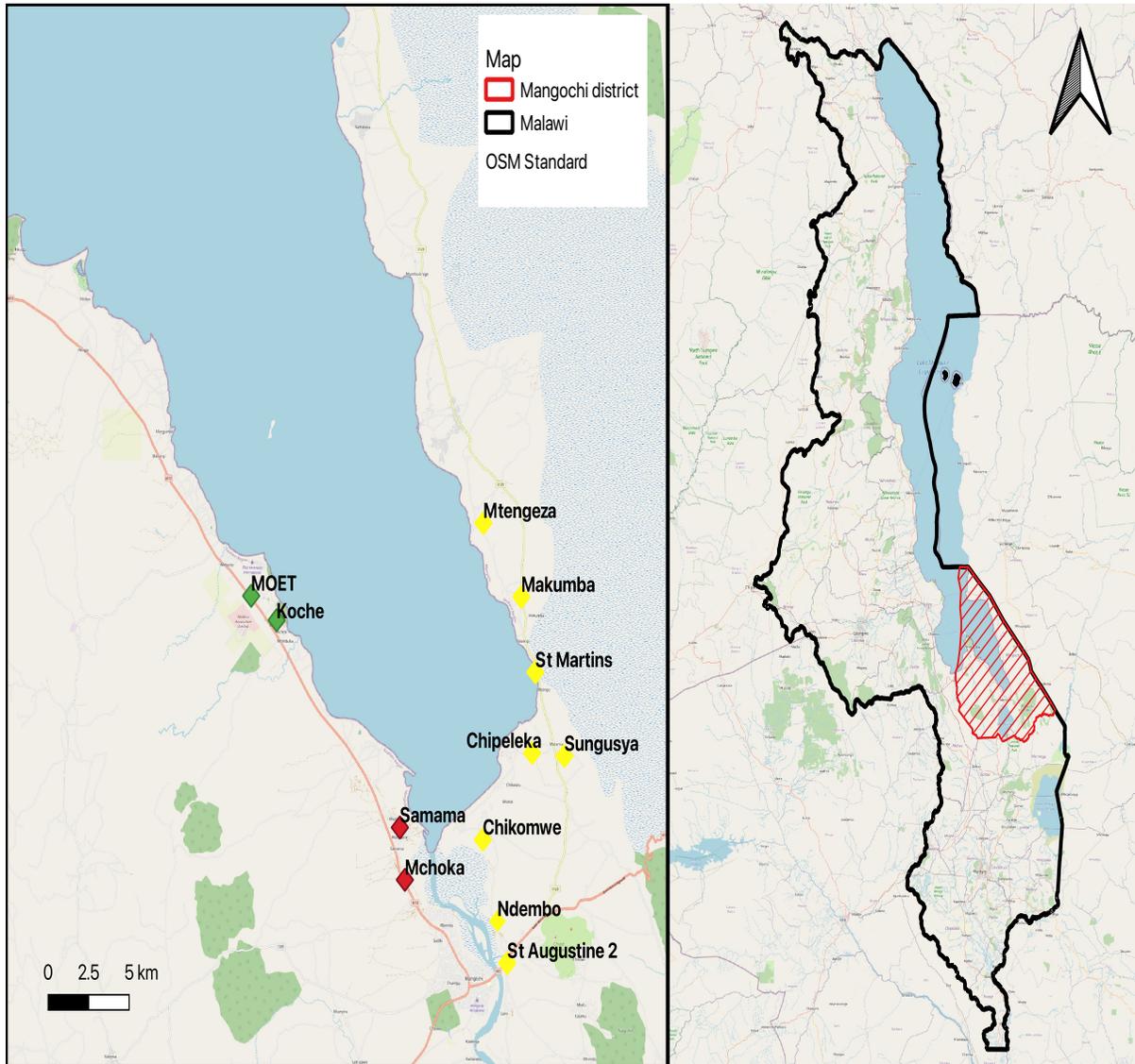


Figure 2.1: Locations of the schools sampled in the primary study, a) red markers represent a repeat of the previous collection (80 SAC sampled), green markers represent collections newly known to *Biomphalaria* intermediate host locations (60 SAC sampled) and yellow markers represent rapid mapping of the shoreline (30 SAC sampled), b) map indicating the location of Mangochi District.

In this secondary analysis of primary data reported by Kayuni *et al.* 2020 [127], our two aims were: i) to determine if general relationships between age of SAC and prevalence of IS, UGS and co-infection could be determined, and ii) to assess heterogeneities in infection-age profiles across sampled schools.

2.2 Methods

2.2.1 Dataset

The primary dataset reported by Kayuni *et al.* [127] which this secondary analysis is based on, was originally collected in late May/June 2019 from cross-sectional school-based surveys in Mangochi District along the shoreline of southern Malawi (Figure 2.1) [127]. In brief, the study carried out a mixture of rapid diagnostic tests, parasitological examinations and questionnaire surveys on 520 primary school children, aged 6 –15 years old in twelve schools, after parental consent was given. The study was split into three phases during May/June 2019: 80 SAC each from Samama and Mchoka schools – annual follow up [11]; 60 SAC each from Moet and Koche schools – an assessment of the two schools near known locations for *Biomphalaria*; and 30 SAC each from 8 further schools along the lake shoreline – a rapid surveillance exercise. The SAC were randomly sampled after being stratified by gender and age, with sample sizes at each school calculated by standard sample size methodology [127]. As reported by Kayuni *et al.* [127], all participants provided a urine sample. Sampling was accompanied by a questionnaire on demographics, water contact behaviour, praziquantel treatment history and travel. A visual inspection of the urine samples was carried out before samples underwent on-site testing using the circulating cathodic antigen (CCA) test for IS, and 10 ml well-mixed urine was filtrated for UGS [127]. The former was used to estimate prevalence of IS and the latter for UGS [56, 128].

Ethical approval for this study was obtained from the National Health Sciences Research Committee, Mangochi District Health Office Research Committee and LSTM's Research Ethics Committee.

2.2.2 Statistical analysis

The primary data were cross-checked with any ambiguities resolved against paper records, then secondary analyses were carried out in R version 3.6.1 with RStudio. The CCA antigen test and urine filtration results were used as binary response variables to measure the prevalence of infec-

tion. The responses were categorised into two subgroups “1 = Positive” and “0 = Negative” in our study. For CCA antigen tests in the original study an additional “trace” result was recorded. In our study, we carry out two analyses: one as “T+” (Trace positive) and one as “T-” (Trace negative). “T+” is where all trace responses are considered “Positive” and “T-” is where all trace responses are considered “Negative”. The main analysis reports T+ results only (allows for low intensity infections) whereas T- results are in the appendix [63, 129].

As a visual exploration tool, heatmaps were used to inspect the empirical age-prevalence profile of *S. mansoni* and *S. haematobium* in each school. The order of the schools on the heatmaps reflected a highest to lowest prevalence ranking for *S. mansoni*.

For both *Schistosoma* species assessed, our response data were binary: an individual was denoted positive (1) or negative (0) for infection and for co-infection an individual was positive (1 and 1) for both infections. Our response data were from independent tests detecting different infections. We assumed therefore, given the characteristics of a child, their age and school, that test results are independent between children. The CCA and urine filtration tests behave the same with respect to school and age and such we assumed that the sensitivity and specificity do not change with respect these characteristics.

We assumed the diagnostic data followed a Bernoulli distribution and therefore used a logistic regression framework. Since our exploratory data analysis (Appendix B Figs. B.1 and B.2) suggested a non-linear relationship between log odds of infection and age, we fitted age using a thin plate spline. School was additionally fitted as a categorical explanatory variable to adjust for systematic school-level variation in baseline prevalence. The resulting logistic generalised additive model (GAM) enables estimation of a smooth, though non-linear, relationship between age and prevalence as a trend summary of our otherwise noisy observational data [97]. GAMs were fitted using the “mgcv” package in R version 3.6.1 which fits the model by penalized likelihood maximisation (penalised regression splines) with an associated smoothing parameter [130] (Appendix B Fig. 2.3).

After fitting these models, smooth age-prevalence curves were reconstructed for each outcome. Confidence intervals were created using Parametric bootstrapping. To calculate the 95% confidence limits of the age prevalence curves the parametric bootstrap method was used. The parametric bootstrap method was used as we have a known Binomial distribution with unknown parameters [64]. We apply the method to our data, by first taking the sampled number of positives cases x in an age group in the data and dividing by the overall number of individuals in that age group n ,

$$\hat{p} = \frac{x}{n}, \tag{2.1}$$

where \hat{p} is the mean. Our bootstrap samples can be generated by simulating random samples of size n , where we have,

$$x^* \sim \text{Binomial}(n, \hat{p}). \quad (2.2)$$

Then using our generated bootstrap sample x^* , we divide this by n again, as follows

$$\hat{p}^* = \frac{x^*}{n}, \quad (2.3)$$

where \hat{p}^* is the estimated bootstrap sample mean and this process is repeated for 10000 samples, to find a good estimate of the sampling distribution of \hat{p} . We find the 95% confidence limits for population mean, \hat{p} by sorting the estimated bootstrap sample means \hat{p} from lowest to highest, and removing the estimates 2.5% and 97.5%. Hence, the 95% confidence limits are the smallest and largest values of each end of the confidence interval.

Model fit was assessed by plotting the average of binned residuals against the fitted values as shown in Appendix B Figs. B.6 and B.7 [131].

We did consider sex, gender and water contact covariates (Appendix B.4 Figures B.8 and B.9 shows initial work) to be added to the model however due to lack of data/robustness we did not consider these in our models and only considered school and age as our main focus.

2.3 Data methods

In brief, the GAMs in our study took the form of a logistic regression using a Bernoulli distribution with mean probability p_{ij} . Let Y_{ij} be the diagnostic binary response for individual SAC i at a named school j . Two cases were considered: For dual-infection focus, Y_{ij} is either $Y_{ij} = 1$ if the SAC had a positive result for *S. haematobium* or *S. mansoni* or $Y_{ij} = 0$ if the SAC had a negative result for *S. haematobium* or *S. mansoni* at named school j . For co-infection focus: This follows that Y_{ij} is either $Y_{ij} = 1$ if the SAC had a positive result for both *S. haematobium* and *S. mansoni* or $Y_{ij} = 0$ if the SAC had at least one negative result. The GAM model takes the following form:

$$\text{logit}(p_{ij}) = \log(p_{ij}/(1 - p_{ij})) = \alpha + \beta_i + s(\text{age}_{ij}; k), \quad (2.4)$$

with intercept α , β_i is vector of each school location with i th subject $i = 1, 2, \dots, n$. s is a thin-

plate spline function for age_{ij} where age_{ij} denotes the age of the child i at school j , and k denotes the number of knots (estimated from the data). In all our analysis, the level of significance was set as “indication of significance”. $p < 0.1$ or “significant”; $*p < 0.05$, $** p < 0.01$ or “highly significant” $***p < 0.001$ and 95% confidence intervals were calculated for each model.

2.4 Results

As reported in the primary study, 520 children were tested using urine CCA-dipsticks for *S. mansoni* and urine filtration for *S. haematobium* [127]. Our provisional secondary analysis found that the prevalence of *S. mansoni* at each school ranged from 67.5% to 96.7%, with overall pooled prevalence of 82.5% [T+]. *Schistosoma haematobium* prevalence ranged from 3.3% to 60.0% with an overall pooled prevalence of 24.0%. Co-infection prevalence by school, ranged from 1.67% to 56.7% with overall pooled prevalence of 21.0% (Table 2.1). Ages of the SAC were between 6 and 15 years, with mean age 10.4. Ndembo school had the lowest mean age sampled with 9.77, whereas Mtengza had the highest with 10.7. Trace negative [T-] prevalence summary can be found in Appendix B Table B.1.

2.4.1 Prevalence heatmaps

Fig 2.2 shows that there was considerable heterogeneity between the schools. Further, *S. haematobium* shows a similar pattern of prevalence among SAC to co-infection. Trace negative result [T-] can be found in Appendix B Fig. B.3.

2.4.2 Generalised additive models

The thin-plate spline for age, adjusted for school, used in the GAM enables us to construct a smooth function of the log odds ratio of infection with respect to age. For the average binned residuals, no evidence of outliers or systematic model bias was found, suggesting model fitted well (Appendix B Figs. B.6 and B.7).

Fig. 2.3a shows very strong evidence for a non-linear relationship between *S. mansoni* infection and age [T+: $p = 8.45 \times 10^{-4}$], Fig. 2.3c shows strong evidence for a non-linear relationship between co-infection and age [T+: $p = 7.81 \times 10^{-3}$], whereas there is no evidence of a non-linear relationship between *S. haematobium* infection and age [$p = 0.114$]. This is visualised in Fig. 2.3a, *S. mansoni* [T+], where the smoothing coefficient for age goes from a negative to positive from

ages 6 to 11 before decreasing back to negative, and similarly for co-infection in Fig. 2.3b. For *S. haematobium* there was no clear pattern between prevalence and age for all the schools (Fig. 2.3c). *Schistosoma mansoni* [T-] and co-infection [T-] GAM adjusted for age and school result can be found in Appendix B Fig. B.4.

In comparison to Mchoka School (baseline), the log odds of being positive for *S. mansoni* infection increased by 0.767 [T+] per year [$p = 4.40 \times 10^{-2}$, 95% CI: 2.06×10^{-2} , 1.51] at Samama. Similarly, the log odds of being positive for *S. mansoni* infection [T+] increased by 1.52 [$p = 2.37 \times 10^{-3}$, 95% CI: 0.540, 2.50] at Koche, 2.63 [$p = 1.21 \times 10^{-2}$, 95% CI: 0.576, 4.69] at St Augustine 2, 1.48 [$p = 2.50 \times 10^{-2}$, 95% CI: 0.186, 2.78] at Sungusya, 1.47 [$p = 2.67 \times 10^{-2}$, 95% CI: 0.169, 2.76] at St Martins, and 1.47 [$p = 2.68 \times 10^{-2}$, 95% CI: 0.168, 2.76] at Chipelekera school per year (Table 2.2). For *S. haematobium*, the following significant coefficient estimates suggest that as a SAC ages, the log odds of being positive for *S. haematobium* infection decreases by -2.62 [$p = 1.24 \times 10^{-2}$, 95% CI: -4.68, -0.657] at Koche, increases by 1.75 [$p = 1.99 \times 10^{-6}$, 95% CI: 1.03, 2.47] at Samama, 1.19 [$p = 1.12 \times 10^{-2}$, 95% CI: 0.271, 2.12] at St Augustine 2, and 1.74 [$p = 2.52 \times 10^{-4}$, 95% CI: 0.807, 2.67] at Ndembo per year compared to Mchoka (Table 2.2). Trace negative [T-] result of the GAM can be found in Appendix B Table B.2.

For co-infection, the coefficient estimate for Samama school suggests a significant relationship with age such that the log odds of being positive for co-infection increases by 1.81 [$p = 7.17 \times 10^{-6}$, 95% CI: 1.01, 2.59] [T+] per year compared to Mchoka. Similarly, the log odds of being positive for co-infection decreases by -2.26 [$p = 3.32 \times 10^{-2}$, 95% CI: -4.34, -0.180] at Koche and increases by 1.43 [$p = 4.58 \times 10^{-3}$, 95% CI: 0.440, 2.41] at St Augustine 2, 1.89 [$p = 1.37 \times 10^{-4}$, 95% CI: 0.920, 2.86] at Ndembo and 1.05 [$p = 4.51 \times 10^{-2}$, 95% CI: 28×10^{-2} , 2.07] at Mtengeza per year.

Given our fitted GAMs, Fig. 2.4 provides a prediction of the age-prevalence profile for each outcome in each school. The predictions indicate that prevalence was highly heterogeneous between schools, and highly heterogeneous in terms of the modelled infection outcome. The trace negative result [T-] can be found in Appendix B Fig. B.5.

2.5 Discussion

To our knowledge, the secondary analysis reported here is the first to analyse the IS infection-age relationship within SAC in a newly established and novel co-infection focus. The newly emerging focus of IS was first noted by Al-Harbi *et al.* [11] then described in greater detail by Kayuni *et al.* [127]. Even though MDA has been ongoing, this focus of IS and co-infections thereof, is being further documented as it seemingly spreads along the southern part of the shoreline of Lake

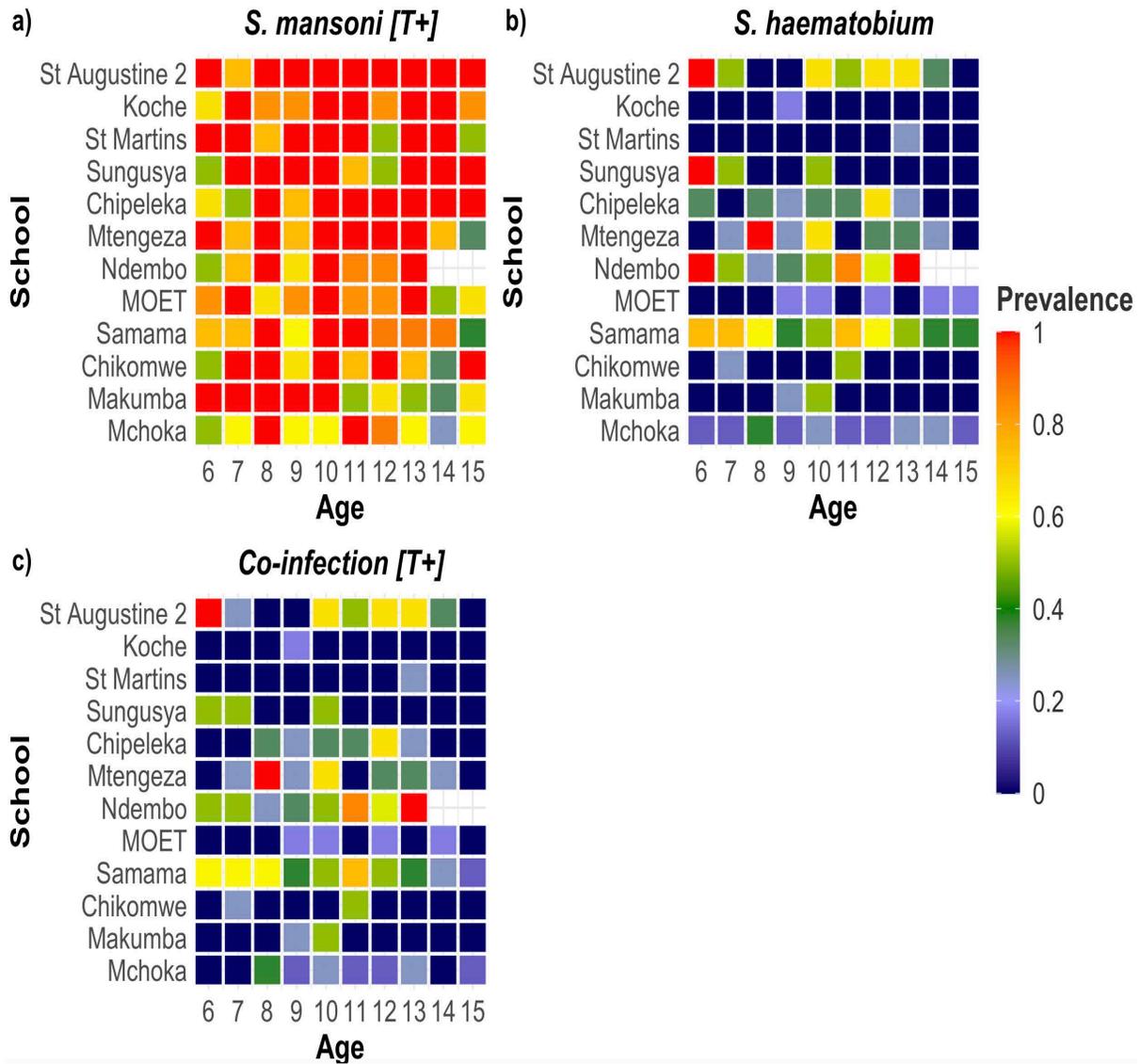


Figure 2.2: Heatmap showing the age of the children vs school prevalence for a) *S. mansoni* [T+] b) *S. haematobium* c) co-infection [T+]. Order of schools on heatmap was by highest to lowest prevalence ranking for *S. mansoni* and showed that there was considerable heterogeneity between the schools. Further, *S. haematobium* shows a similar pattern of prevalence among SAC to co-infection.

Table 2.1: Summary of prevalence of *S. mansoni* [T+], *S. haematobium* and co-infection [T+]

Name	<i>S. mansoni</i> (CCA)[T+]		<i>S. haematobium</i> (Urine filtration)		Co-infection [T+]	
	No. Positive (%)	95% CI	No. Positive (%)	95% CI	No. Positive (%)	95% CI
Mchoka (N=80)	54 (67.5)	57.5-77.5	15 (18.9)	10.0-27.5	11 (13.8)	6.25-21.3
Samama (N=80)	65 (81.3)	72.5-88.8	45 (56.2)	45.0-67.5	38 (47.5)	36.3-58.8
MOET (N=60)	49 (81.7)	71.7-90.0	5 (8.33)	1.70-15.0	4 (6.67)	1.67-13.3
Koche (N=60)	54 (90.0)	81.7-90.0	1 (1.67)	0.00-5.00	1 (1.67)	0.00-5.00
St Augustine 2 (N=30)	29 (96.7)	90.0-100	13 (43.3)	26.7-60.0	12 (40.0)	23.3-56.7
Ndembo (N=30)	25 (83.3)	70.0-96.7	18 (60.0)	43.3-76.7	17 (56.7)	40.0-73.3
Sungusya (N=30)	27 (90.0)	76.7-100	5 (16.7)	3.33-30.0	4 (13.3)	3.33-26.6
St Martins (N=30)	27 (90.0)	80.0-100	1 (3.33)	0.00-10.0	1 (3.33)	0.00-10.0
Chikomwe (N=30)	24 (80.0)	63.3-93.3	3 (10.0)	0.0-23.3	3 (10.0)	0.00-23.3
Chipeleka (N=30)	27 (90.0)	76.7-100	8 (26.7)	13.3-43.3	7 (23.3)	10.0-40.0
Makumba (N=30)	23 (76.7)	60.0-90.0	2 (6.67)	0.00-16.7	2 (6.67)	0.00-16.7
Mtengeza (N=30)	25 (83.3)	70.0-96.7	9 (30.0)	13.3-46.7	9 (30.0)	13.3-46.7
Total (N=520)	429 (82.5)	79.2-85.8	125 (24.0)	20.4-27.7	109 (21.0)	17.5-24.4

Malawi in the Mangochi District. For *S. mansoni* infection detected by CCA dipsticks from the primary data [127], our secondary analysis finds that a positive association between IS prevalence and age was observed up to the age of 11, after which there was a decreasing trend [T+: $p = 8.45 \times 10^{-4}$]. As might be expected, co-infection showed a similar pattern [T+: $p = 7.81 \times 10^{-3}$], largely mirroring the IS pattern. By contrast, no clear age-infection pattern for UGS was identified [$p = 0.114$].

Other studies on *Schistosoma* infection carried out in sub-Saharan Africa have found varied peak age-infection profiles and all to be expected to arise around early to mid-adolescence [10–15 years] [17, 65, 76, 132, 133]. The earlier observed peak in *S. mansoni* and co-infection prevalence at the age of 11 in this study compared to up to 15 years may be a result of the newly established transmission potential of this species locally alongside growing acquired immunity in exposed children. This is due to their cumulative exposure to parasite antigens as adult worms die within the body after treatment or natural senescence, or as egg antigens present. Literature reports varied age-infection profiles considered to be dependent on the transmission rates and focality [123].

In classic infection epidemiology of schistosomiasis, changes in the “peak shift” are known [109, 123]. These can be explained by site-specific factors, for instance, water exposure, environmental, socio-economic, genetic, MDA compliance, as well as age and sex profiles within a community [121]. In our instance, the expansion of the underlying distribution of *Bi. pfeifferi* both in time and space is an influential transmission potential driver of IS. A key observation is the contrasting age-prevalence by schistosome species, yet each share a common infection pathway, *viz.* exposure to unsafe water. The occurrence of the snail species present in unsafe water is an underlying heterogeneity of the fine-scale distributions of intermediate snail hosts, *viz.* *Biomphalaria* for *S. mansoni* and *Bulinus* for *S. haematobium*. The latter genus of snail is also undergoing a reappraisal

as cryptic species, with as of yet unknown transmission potentials, as described in [134]. Whilst Kayuni *et al.* 2020 [127] and Al-Harbi *et al.* 2019 [11] presented information on the presence and absence of *Biomphalaria*, a similarly precise map for *Bulinus* is starting to emerge [134]. A recent study of *Bi. pfeifferi* has confirmed a year-on-year expanding distribution of this species along the shoreline of the lake, with clear evidence of schistosome DNA in examined snails from several independent locations [135]. It is reasonable to speculate that further transmission foci for intestinal schistosomiasis will continue to appear in the lake and along its periphery.

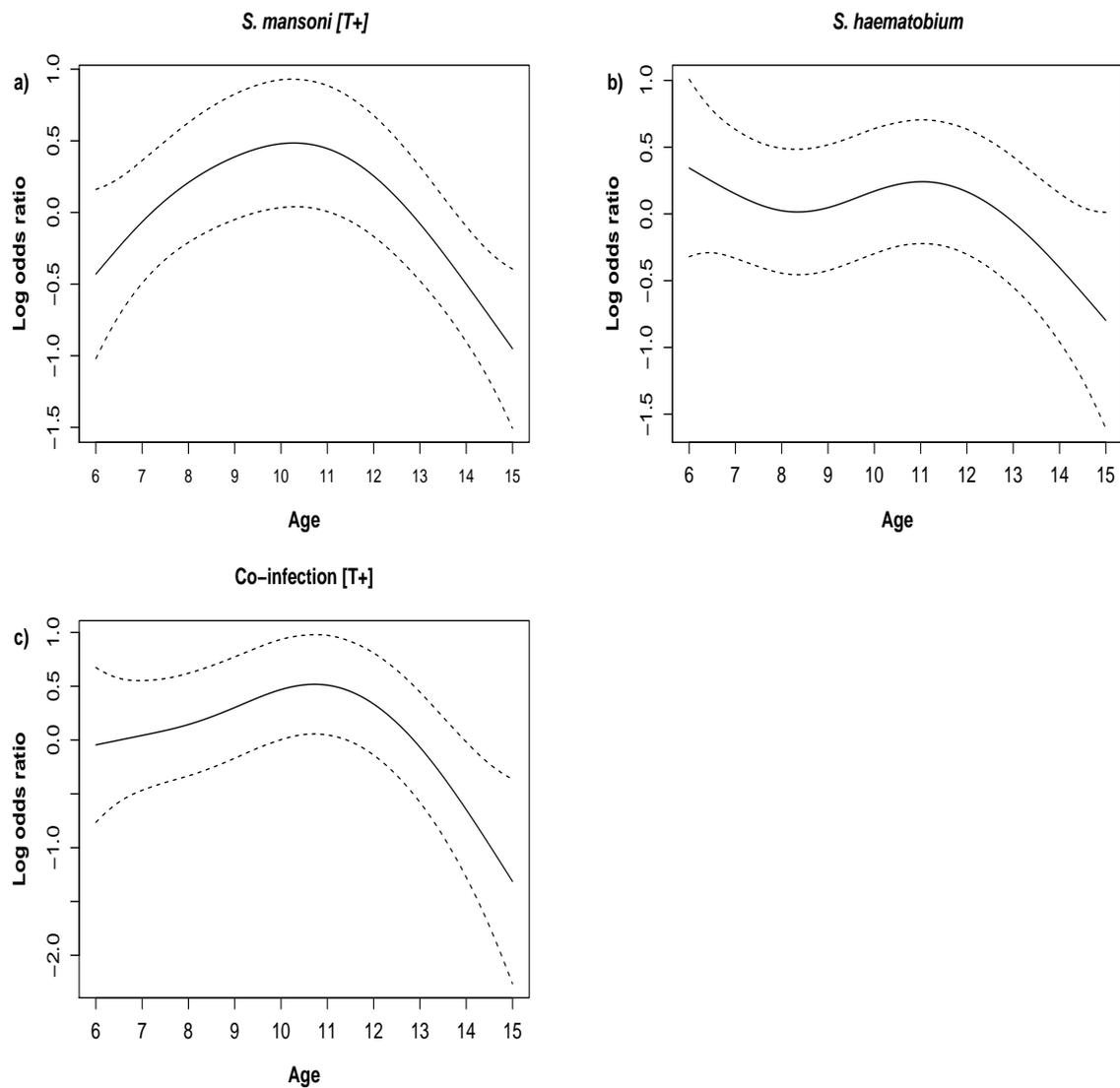


Figure 2.3: Thin plate spline functions of the log odds ratio of *Schistosoma* infection by age in SAC for (a) *S. mansoni* [T+] (b) *S. haematobium* (c) coinfection [T+]. A general trend towards SAC between 9 and 12 years old having the highest odds of infection was seen in all cases, though that for UGS is not statistically significant.

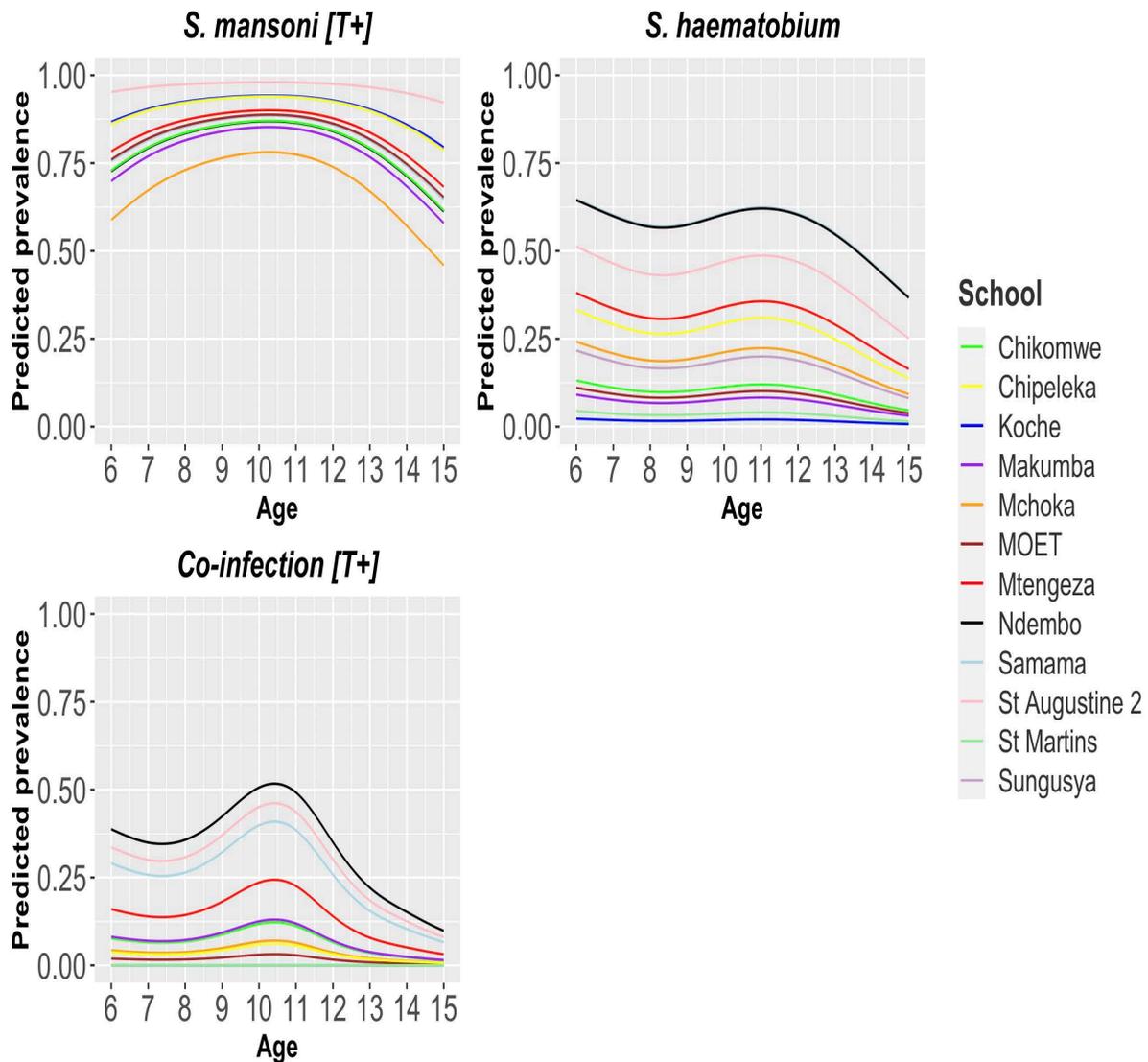


Figure 2.4: Smoothed age-specific prevalence of *Schistosoma* association with age of SAC for each school. a) *S. mansoni* [T+] b) *S. haematobium* C) co- infection [T+]. Light Green: Chikomwe, Yellow: Chipeleka, Dark Blue: Koche, Purple: Makumba, Orange: Mchoka, Brown: MOET, Red: Mtengeza, Black: Ndembo, Light Blue: Samama, Pink: St Augustine 2, Dark Green: St Martins, Mauve: Sungusya.

Table 2.2: Coefficients for the GAM with age smoothing term adjusted for school

	<i>S. mansoni</i> [T+]		<i>S. haematobium</i>		Co-infection [T+]	
	95% CI		95% CI		95% CI	
Smooth term (p-value)						
Age	$8.45 \times 10^{-4}***$		0.114		$7.81 \times 10^{-3}**$	
Factor (estimated coefficient)						
School						
Samama	0.767*	(0.0206,1.51)	1.75***	(1.03, 2.47)	1.81***	(1.01, 2.59)
MOET	0.796.	(-0.0251,1.62)	-0.940.	(-0.202,0.138)	-0.815	(-2.02, 0.391)
Koche	1.52**	(0.540, 2.50)	-2.62*	(-4.68, -0.567)	-2.26*	(-4.34, -0.180)
St Augustine 2	2.63*	(0.576, 4.69)	1.19*	(0.271, 2.12)	1.43**	(0.440, 2.41)
Ndembo	0.621	(-0.465, 1.71)	1.74***	(0.807, 2.67)	1.89***	(0.920, 2.86)
Sungusya	1.48*	(0.186, 2.78)	-0.143	(-1.26,0.975)	-4.33×10^{-2}	(-1.28, 1.20)
St Martins	1.47*	(0.169, 2.76)	-1.92.	(-3.99, 0.155)	-1.57	(-3.67, 0.532)
Chikomwe	0.634	(-0.393, 1.66)	-0.745	(-2.07, 0.578)	-0.387	(-1.75, 0.97)
Chipeleka	1.47*	(0.168,2.76)	0.445	(-0.546, 1.43)	0.628	(-0.444, 1.70)
Makumba	0.484	(-0.503, 1.47)	-1.16	(-2.70, 0.386)	-0.788	(-2.37, 0.791)
Mtengza	0.928.	(-0.160, 2.02)	0.656	(-0.314,1.63)	1.05*	(0.0228, 2.07)
Mchoka	0	-	0	-	0	-

Heterogeneities in prevalence among the schools were also found in our secondary analysis, with Mchoka School having the lowest prevalence and St Augustine 2 the highest prevalence for *S. mansoni* infection. Clearly, this heterogeneity indicates that there are many further un-considered factors that affect the transmission of *Schistosoma* infection within SAC, such as location, local environmental and socio-economic factors. For instance, SAC living and attending school in areas near the lake shoreline have been found to have increased and different age-infection profiles compared to inland villages [76]. As longitudinal data were not collected in the primary study, we were not able to assess seasonal or long-term variation in prevalence; however it is possible that the force of infection could vary spatially and temporally. For instance, environmental changes such as increases or decreases in water levels of the lake, flooding events or fluctuations in aquatic vegetation could impact SAC water contact [11, 127]. Another factor not studied was reinfection after preventive chemotherapy, as PZQ only affects the adult worms, and any immediate snail exposure could therefore lead to reinfection. Further studies into the relationship between age, water exposure rates and treatment in the future could enhance our perspective of age infection profiles. Identifying the peak of infection prevalence within SAC at school level using a GAM increases the interpretability of our findings by turning noisy data into useful and assessable information which in turn will help better our understanding of epidemiology of infection and other control methods along the southern part of the shoreline [136].

A further limitation of this secondary analysis was that, owing to available resourcing, sample size taken from each school in the primary data was constrained [127]. Further, the schools may not had been a representative samples as for example, Moet and Koche were chosen as near newly identified *Biomphalaria* snail habitat whereas 8 of the schools were randomly sampled. Further

trace positive result were mainly considered instead of trace negative results, which allows for low intensity infections to be considered however, can overestimated the prevalence of infections [63, 129]. More generally, GAMs are sometimes known to smooth out underlying relationships excessively. Also, they have higher computational load compared to linear models and unstable behaviours at the boundaries of smooth splines [137]. Nevertheless, for the purpose of our study, the general age-prevalence relationships were detected adequately, and provide a useful insight for future research into the causal mechanisms driving this infection biology.

2.6 Conclusion

Our study which is a secondary analysis of recently collected epidemiological data concerning a newly emergent focus of IS against an existing background of UGS, provides evidence for the peak of prevalence for *Schistosoma* infection being around 11 years for both *S. mansoni* mono-infection and co-infection with *S. haematobium* along the southern part of Lake Malawi in Mangochi district. However, considerable heterogeneity still remains in terms of baseline prevalence between schools, and investigating this in terms of demographics and *Schistosoma* transmission dynamics requires further research. In particular, understanding how SAC exposure is related to water access will require both further prevalence and malacological niche mapping. Coupling these conclusions into statistically-grounded infection modelling techniques will advance the understanding of the dynamics of *Schistosoma* infection, and hence inform future intervention programmes.

2.7 Declaration of Competing Interest

None.

2.8 Data availability

Anonymised epidemiological data are available from the corresponding author upon request.

2.9 Acknowledgments

With thanks to the field teams and children in Malawi who participated in the surveys. Thank you to Dr. Claudio Fronterre at Lancaster University for helpful discussions during the preparation of this manuscript. We gratefully acknowledge the MRC DTP LSTM & Lancaster University (MR/N013514/1) for funding for AR.

Chapter 3

A geospatial analysis of local intermediate snail host distributions provides insight into schistosomiasis risk within under-sampled areas of southern Lake Malawi

Authors list: Amber L Reed¹, Mohammad H Al-Harbi^{2,3}, Peter Makaulai⁴, Charlotte Condemine², Josie Hesketh², John Acher², Sam Jones², Sekeleghe A Kayuni^{2,4}, Janelisa Musaya⁴, Michelle C Stanton⁵, J Russell Stothard², Claudio Fronterre¹, Christopher Jewell⁶

Author Affiliations

1. Lancaster Medical School, Lancaster University, Bailrigg House, Bailrigg, Lancaster, LA1 4YE
2. Tropical Disease Biology, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA
3. Ministry of Health, Buraydah 52367, Saudi Arabia
4. Malawi Liverpool Wellcome Trust Programme of Clinical Tropical Research, Queen Elizabeth Central Hospital, College of Medicine, P.O.Box 30096, Blantyre, Malawi
5. Vector Biology, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA
6. Mathematics and Statistics, Lancaster University, Bailrigg House, Bailrigg, Lancaster, LA1 4YE

Keywords: *Bulinus*, *Biomphalaria*, Snail abundance, Bayesian multilevel models, Geospatial analysis, Gaussian latent process, Remote sensing

Abstract

Background: Along the southern shoreline of Lake Malawi, the incidence of schistosomiasis is increasing with snails of the genera *Bulinus* and *Biomphalaria* transmitting urogenital and intestinal schistosomiasis, respectively. Since the underlying distribution of snails is partially known, often being focal, developing pragmatic spatial models that interpolate snail information across under-sampled regions is required to understand and assess current and future risk of schistosomiasis.

Methods: A secondary geospatial analysis of recently collected malacological and environmental survey data was undertaken. Using a Bayesian Poisson latent Gaussian process model, abundance data were fitted for *Bulinus* and *Biomphalaria*. Interpolating the abundance of snails along the shoreline (given their relative distance along the shoreline) was achieved by smoothing, using extracted environmental rainfall, land surface temperature (LST), evapotranspiration, normalised difference vegetation index (NDVI) and soil type covariate data for all predicted locations. Our adopted model used a combination of two-dimensional (2D) and one dimensional (1D) mapping.

Results: A significant association between normalised difference vegetation index (NDVI) and abundance of *Bulinus* spp. was detected [log risk ratio -0.83, 95% CrI: -1.57, -0.09]. A qualitatively similar association was found between NDVI and *Biomphalaria* sp. but was not statistically significant [log risk ratio -1.42, 95% CrI:-3.09, 0.10]. Analysis of all other environmental data were considered non-significant.

Conclusions: The spatial range in which interpolation of snail distributions is possible appears less than 10km owing to fine-scale biotic and abiotic heterogeneities. The forthcoming challenge is to refine geospatial sampling frameworks with future opportunities to map schistosomiasis within actual or predicted snail distributions. In so doing, this would better reveal local environmental transmission possibilities.

3.1 Introduction

Schistosomiasis is a freshwater snail-borne neglected tropical disease (NTD) common across much of sub-Saharan Africa. Two forms of schistosomiasis occur, urogenital and intestinal schistosomiasis. Their respective transmission can only occur if permissive intermediate snail hosts of the genus *Bulinus* and *Biomphalaria* respectively occur. While various species of *Bulinus* are present in Lake Malawi, with *Bulinus globosus* and *B. nyassanus* responsible for most *S. haematobium*

transmission, only from 2017 was *Biomphalaria* first formally noted along its southern shoreline. The expanding distribution of *Biomphalaria pfeifferi* in this area has facilitated transmission of *Schistosoma mansoni*, which causes intestinal schistosomiasis, which has now transitioned from emergence to outbreak [11, 127, 134].

Owing to the singular importance of this newly invasive *Bi. pfeifferi*, subsequent malacological surveys were undertaken to track its presence alongside concurrent parasitological surveys in local children attempting to define the extent of schistosomiasis, particularly IS. These surveys demonstrated the need for further surveillance of snails, alongside emphasis upon updated and tailored interventions and policies for control of schistosomiasis in this lacustrine setting [11, 127, 134]. However, as snail distributions can be patchy or focal, owing to their dependency on local habitats, many gaps in current cartography and predictive mapping are exposed [138]. Indeed, variation in such local characteristics creates difficulties in outlining either permissive or refractory areas where snails may or may not be found, thereby confounding control strategies.

A combination of climate change and human behaviour is thought to be the primary reason for *Biomphalaria* invasion and colonisation into new areas [85]. Characteristics such as vegetation, temperature, rainfall (precipitation), evapotranspiration and soil type have been reported as possible effects on determining snails' presence and abundance, increasing potential heterogeneity in snail populations over a wide area [85–87]. Changes in climate and seasonal patterns are therefore likely to alter transmission of schistosomiasis over both space and time, increasing the need for identifying snail habitats to target appropriate control interventions [85]. However, although snail distribution within a geographical area can be measured through malacological surveillances, physically collecting freshwater snails is expensive and time consuming, and it is therefore unfeasible to sample every possible location. Thus, effective sampling remains incomplete.

Lake Malawi dominates the eastern side of Malawi, being 600km long and 75km wide. It is known as the second deepest lake in Africa [76] and is vital for those using it for irrigation, agriculture, water supply, fishing industries and tourism [139]. Due to the lack of sanitation in Malawi, human urine and faecal materials continuously contaminate the shoreline facilitating the transmission of schistosomiasis, amongst other water-borne pathogens [140]. In Mangochi District, representing the southern part of Lake Malawi, the eastern side of the lake is mountainous with high elevation (1000–1500m), whereas the western side is flat and with lower elevation (<500m) [141, 142]. Lower temperatures and higher winds are reported on the eastern side [143], with low-lying areas such as the upper Shire River margins vulnerable to flooding [144]. More broadly, the climate of this southern part of the shoreline is affected by the migration of the Inter-Tropical Convergence Zone (ITCZ). This leads to the dry season with cooler temperatures occurring between May and August, hotter temperatures between September to November and wet season between December

and April [145, 146]. Rainfall is dependent on altitude and time of the year [147]. Lake water levels vary over time and are at their highest during wet season, which also affects evapotranspiration and outflows to the Shire River [127, 144]. Most important perhaps, is an increasing human and livestock population which is leading to more frequent water contact, enhancing opportunities for transmission of schistosomiasis [127, 138].

The World Health Organisation (WHO) has supplied new guidelines to target elimination of schistosomiasis by interrupting transmission [63]. The ability to be able to identify locations where freshwater snails are most abundant aids targeted control methods, preventing initial infection and re-infection and hence helping eliminate or reduce transmission [67, 148, 149].

Here in the Lake Malawi setting, we undertook a secondary analysis of primary malacological data first reported by Al-Harbi *et al.* 2019 [11] and Kayuni *et al.* 2020 [127]. Our study models the snail distributions as a function of environmental and climate data measured along the shoreline aiming to i) interpolate and predict the distribution of the snails along the shoreline of Lake Malawi where the snails had not been sampled and ii) assess the association between environment data and snail distributions. In turn, we hoped to clarify the extent of environmental heterogeneities for schistosomiasis transmission along the shoreline of Lake Malawi and inform the targeting of control programmes to the most appropriate snail breeding sites.

3.2 Methods

The data used in this study consist of observations of snail abundance at a small number of discrete locations on the Lake Malawi shoreline together with remote-sensing data used to describe snail habitat. These are described separately below.

3.2.1 Snail abundance

The primary dataset reported in Al-Harbi *et al.* 2019 [11] and Kayuni *et al.* 2020 [127] which this secondary analysis is based on, originally collected malacological surveys between 2017–2019 as shown in Figure 3.1. Pilot surveillance data from November 2017 identified *Biomphalaria* sp. and *Bulinus* spp. along the shoreline. May/June 2018 and 2019 malacological surveys resampled some of the original locations and added new sites based on satellite imagery or randomly based on their surrounding environment suitable for breeding sites to confirm the emergence and outbreak of IS. The Danish Bilharziasis Laboratory key was used to identify *Bulinus* and *Biomphalaria* according to shell morphology. Figure 3.1b shows a map of sampling sites, together with their relationship

to primary schools in the region, demonstrating human proximity to the lake shore and hence potential for exposure to infected snails.

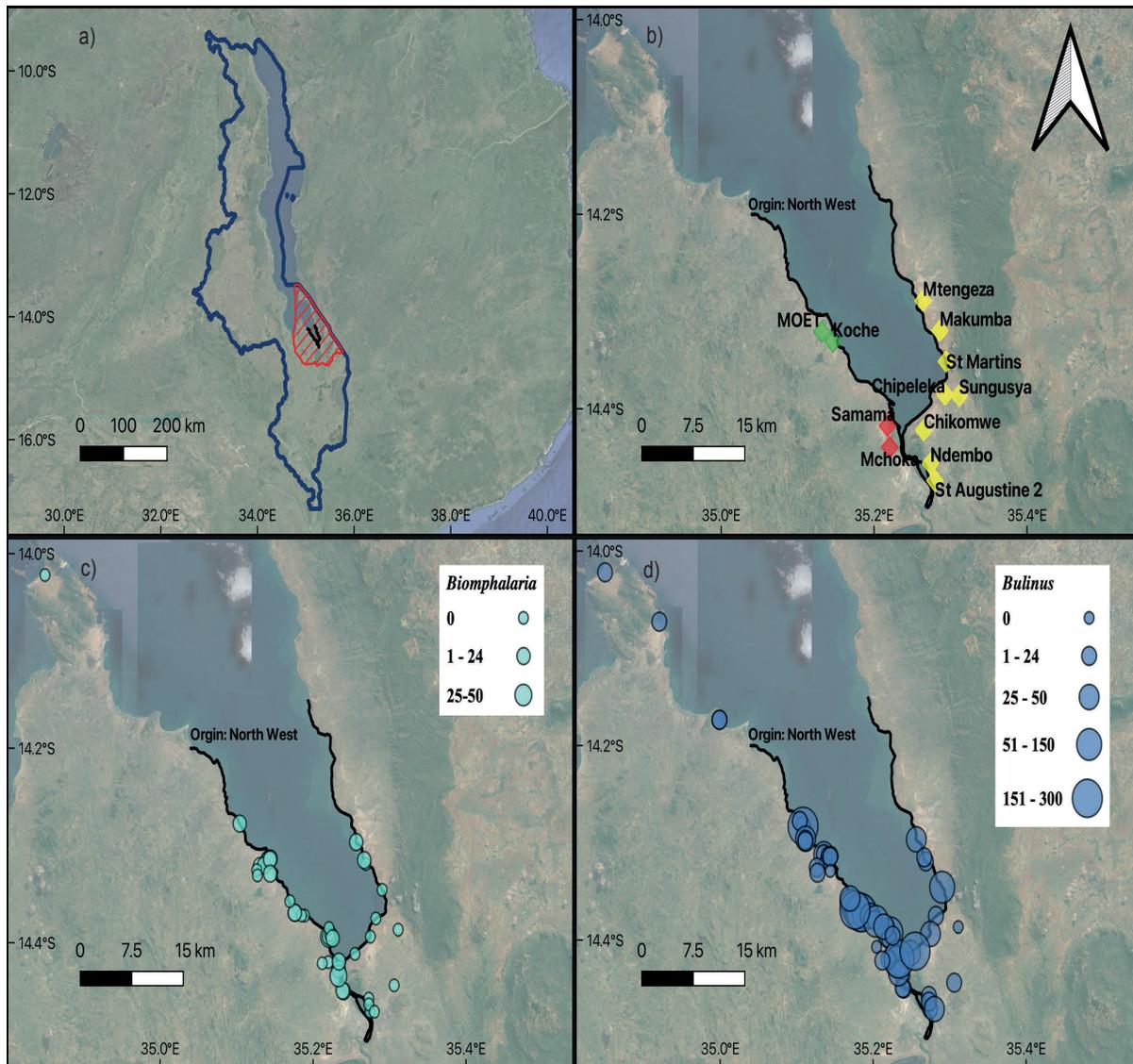


Figure 3.1: Primary dataset collected data a) Map of Malawi in dark blue. Red crossed: study area; black line: prediction points. Parasitological surveys: b) Primary school locations along the shoreline. Malacological surveys: c) observed *Biomphalaria* snails; d) observed *Bulinus* snails.

The snail abundance counts taken from the primary dataset snail counts were numerical counts or in some cases reported as approximate values, e.g 300+. In our study we took these approximate value and assumed these values to be the closest lowest value, e.g. 300. The recorded sites considered in our study are shown in Figures 3.1c and 3.1d.

3.2.2 Remote sensing data

Publicly available continuously collected satellite sensory systems were used to extract environmental and climatic data measured adjacent to the shoreline as shown in Figure 3.2. Rainfall estimates were extracted between 1st January 2017 to 30th June 2019 from Tropical Applications of Meteorology using satellite data and ground-based observations (TAMSAT) with a monthly frequency at 4km resolution [150–152]. Land Surface Temperature (LST), evapotranspiration, Normalised Difference Vegetation Index (NDVI) raster data were obtained from Land Processes Distribution Active Archive Center (LPDAAC) [153–156]. LST data between 1st November 2017 to 30th June 2019 were extracted from Moderate Resolution Imaging Spectroradiometer (MODIS)/Terra LST/Emissivity 8-Day L3 Global 1km SIN Grid raster (MOD11A2v061) [153]. Evapotranspiration data were extracted between 1 January 2014 to 1 January 2019 (5-year time frame) from Modis/Terra Evapotranspiration Gap-Filled Yearly L4 Global 500m SIN Grid raster (MOD16A3GFv061) [154]. NDVI data between 1th November 2017 to 30th June 2019 were extracted from Modis/Terra vegetation indices 16-Day L3 Global 1km SIN Grid raster (MOD13A2v0061) [155]. Soil type polygon data were taken from the International Soil Reference and Information Centre (ISRIC) World Soil Information and were derived from the Soil Terrain Database for Malawi (SOTER) at a scale of 1:1M [157]. After extracting the values, the temporal covariates were aggregated by taking the mean of the values over the time frame.

3.2.3 Construction of 200 prediction points along the shoreline

Snail abundance was predicted in 1D representation to allow us to interpolate the values along the whole linestring. We made this assumption on the basis that snails live along the shoreline, in habitats that are associated with human water contact and entry, so correlations between snail locations are affected by distance along the shoreline and not, for example, by stretches of deep, open water, e.g. mouth of a bay.

The 1D shoreline was represented by computing the distances between a sequence of 200 vertices obtained from the 2D linestring representation. To achieve this, we used the following method : i) A 2D linestring was drawn by hand following the shoreline as shown by Google Satellite imagery (Figure 3.1). ii) the linestring was re-sampled to 4000 equally spaced vertices, iii) each observed sampling site location was snapped to its nearest vertex; iv) the distance along the line from the origin (northwest-most vertex) to each of the snapper observed sampling site locations was computed (Appendix C Figures C.1 and C.2). Additionally, we sub-sampled the 4000 vertices at equal intervals to a set of 200 prediction points.

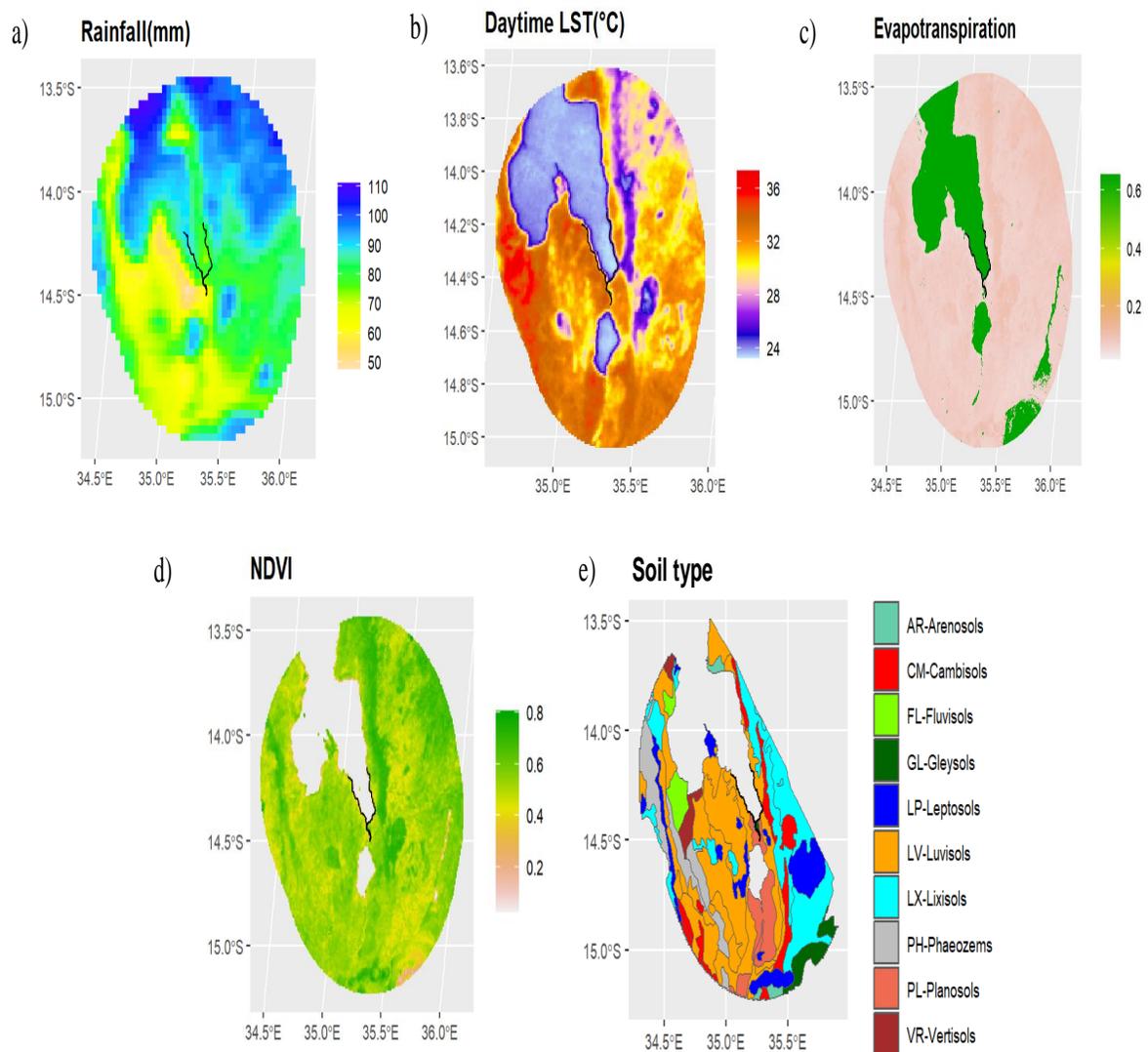


Figure 3.2: Raster plot of extracted covariate data. a) Rainfall (mm), b) Daytime LST ($^{\circ}C$), c) Evapotranspiration d) NDVI. e) Soil types for southern part of Lake Malawi shoreline, adapted from Dijkshoorn *et al.* 2016 [157]. Black line shows the shoreline template where covariate values were extracted from.

3.2.4 Extraction of remote sensing data to linestring vertices

The covariate data were created by extracting the values of each remotely sensed covariate layer data variable surface at each of 200 linestring vertices. To do this, the mean of raster pixels within a 1km buffer around each vertex was computed. Where missing values were found for a vertex, the buffer took the calculated mean value for the previous corresponding vertex working away from the origin. In cases where missing values were present as the first sampling point, the next collected value was taken. From the 200 prediction points created, 196 points were actually predicted as removed CM soil type from the analysis due to lack of data points.

3.2.5 One dimensional Poisson latent Gaussian process regression

3.2.5.1 Bayesian Multilevel Model

A Bayesian Poisson multilevel model (BMLM) with a Gaussian latent process (GP) was developed using STAN programming language [158], which uses a Markov Chain Monte Carlo (MCMC) algorithm to regress snail abundance data onto the remotely sensed covariate data, accounting for (1D) spatial correlation along the shoreline. We assume that the number of snails observed at a sampling location was Poisson distributed, with log-mean given by a coefficient-weighted sum of the covariates plus a spatially correlated error term. Covariance between the error terms was represented as the sum of spatially correlated variance (using quadratic, exponential, or Matérn) κ uncorrelated (or nugget) variance [159]. Suitably weakly informative priors were applied to the model coefficients and variance terms, with MCMC run for 10,000 iterations. Posterior summaries (mean and 95% credibility intervals) were computed for the fitted model, as well as predictive distributions for each of the linestring vertices conditional on the data. All data processing and analysis were performed in R version 4.1.1. See Model Formulation (next section 3.2.6) for a mathematical explanation of the model.

3.2.6 Model Formulation

3.2.6.1 Model

We used a Bayesian log-linear Gaussian Process model to smooth observations of snail abundance along the shoreline. We assume that the number of snails y_i at observed (sampling) locations x_i expressed as distance along the shoreline relative to an arbitrary reference point follow a Poisson distribution with mean μ_i :

$$y_i \sim \text{Poisson}(\mu_i), \quad (3.1)$$

The mean number of snails is then modelled on the log-scale as sum of two components: a set of covariates expressed through the design matrix Z and a Gaussian process :

$$\log(\mu_i) = \alpha + Z^T(x_i)\beta + S(x_i). \quad (3.2)$$

S is modelled using a 1-Dimensional (1D) Gaussian process with exponential correlation function with standard deviation σ and fixed length scale $\phi=3.33\text{km}$ (5% range of measured shoreline). We also included an additional uncorrelated variance term τ (nugget effect). We place priors on our unknown model parameters,

$$\alpha \sim N(0, 100), \quad (3.3)$$

$$\beta_i \sim N(0, 5), \quad (3.4)$$

$$\sigma \sim \Gamma(1, 1), \quad (3.5)$$

$$\tau \sim \Gamma(0.1, 0.1). \quad (3.6)$$

Where Γ distribution is defined using shape and rate parameters. The STAN probabilistic programming language was used to fit the model, allowing us to project the Gaussian process onto a fine grid of points along the shoreline to create predictions of snail abundance at arbitrary points in our study region.

3.2.6.2 GP prediction

To get the GP prediction values we find the mean of the Poisson distribution variables.

$$\log(\hat{\mu}_i) = \hat{\alpha} + Z^T(x_i^*)\hat{\beta} + \hat{s}(x_i^*), \quad (3.7)$$

where $\hat{\mu}_i$ is our GP prediction of the number of snails present, x_i^* is the covariate values for each prediction point, where $i = 1, \dots, N$.

3.2.6.3 Latent Gaussian process

$$s(\vec{x}) \sim SGP(0, \Sigma^2), \quad (3.8)$$

where Σ is the covariance matrix. The exponential covariance function is represented as the following:

$$\Sigma_{ij}^2 = \begin{cases} \sigma^2 \exp\left(\frac{-\|x_i - x_j\|}{\sigma}\right) & \text{if } i \neq j, \\ \sigma^2 + \tau^2 & \text{if } i = j, \end{cases} \quad (3.9)$$

where $\|\cdot\|$ denotes the distance between x_i and x_j .

3.3 Results

3.3.1 Observed data

After cross-checking the observation data, as shown in Figure 3.3, we obtained 33 locations where *Biomphalaria* sp. and 63 locations where *Bulinus* spp. were present. The mean number of snails for *Biomphalaria* sp. was 6.03 ranging from 0 to 50 snails, with the most snails found at 46.17km along the shoreline from the origin. The mean number of snails for *Bulinus* spp. was 28.20, ranging from 0 to 300 snails, with the most snails found at 14.66km along the shoreline from the origin. For observed *Biomphalaria* sp. data the extracted environmental data ranges were rainfall with mean 78.8 [63.01-89.51]mm and LST with mean 29.68 [24.97-32.44 ($^{\circ}C$)] For *Bulinus* spp. data the extracted environmental data ranges were rainfall with mean 80.62 [63.01-89.5]mm and LST with mean 30.28 [24.97-32.44 ($^{\circ}C$)]. Appendix C shows the observed data for 1D (Figure C.3) and 2D (Figures C.4 and C.5). A histogram of the centred and scaled covariates is shown in Appendix C.3 Figure C.6.

3.3.2 Environmental data prediction points

The extracted environmental data prediction point ranges were: rainfall [59.82-90.37mm], LST [24.68-32.46 ($^{\circ}C$)], NDVI [0.29-0.61] and evapotranspiration [0.10-0.66] along the prediction points of the shoreline. Evapotranspiration was lowest and NDVI highest along the River Shire, with eastern shoreline having the most rainfall and lowest LST ($^{\circ}C$) compared to the western shoreline. Luvisolic (LV) soil type was absent around the River Shire compared to Gleysolic

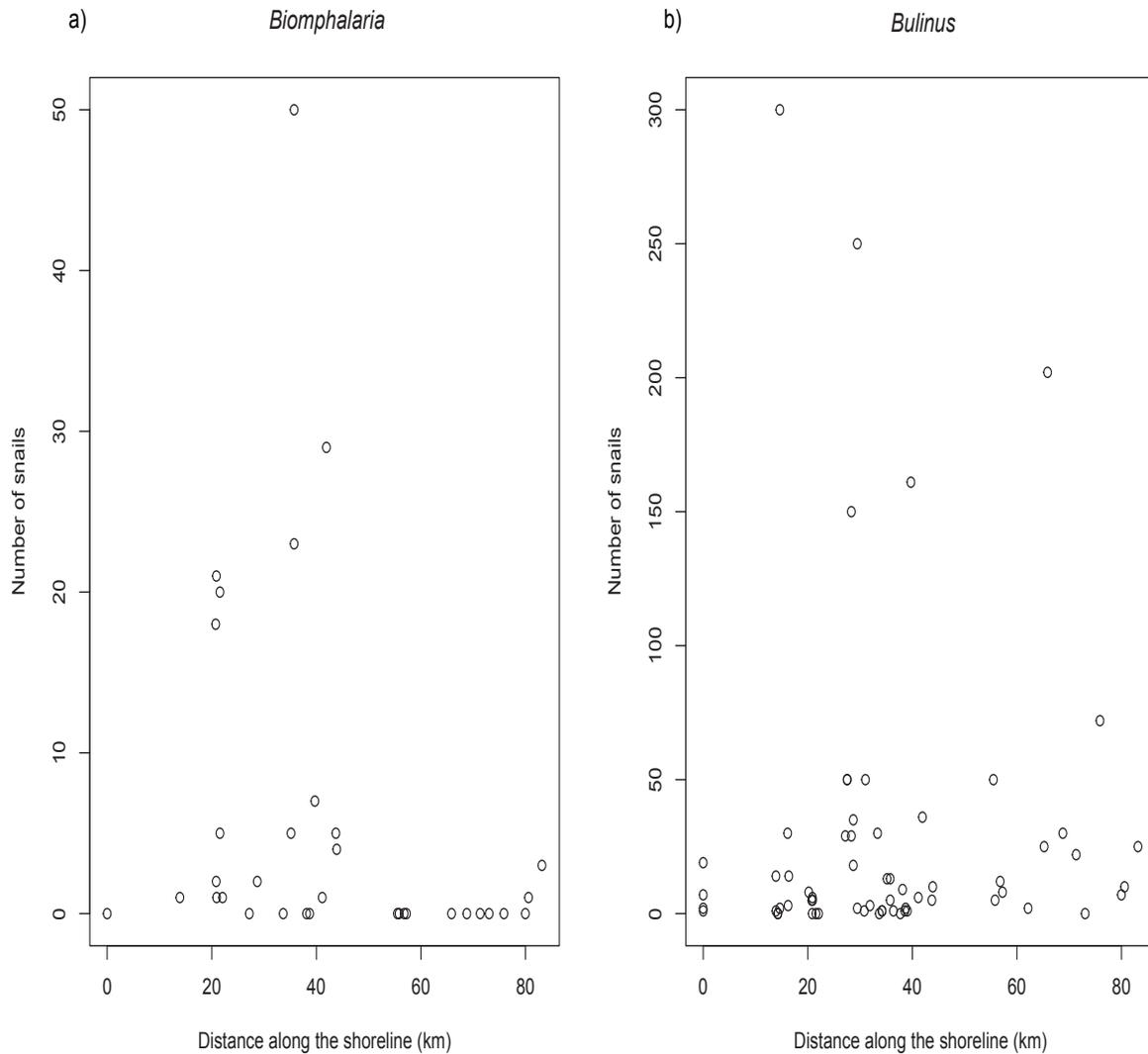


Figure 3.3: Scatter plot of absolute snails numbers observed at sampling points versus distance along the shoreline in km. a) *Biomphalaria* sp; b) *Bulinus* spp..

(GL) soil type; Planosolic (PL) soil type was presence at the entrance to River Shire and south of the River Shire compared to GL soil type. The distributions of the values of the environmental covariates are shown in Figure 3.4 and a ID version can also be seen in Appendix C.3 (Figures C.7a and C.7b).

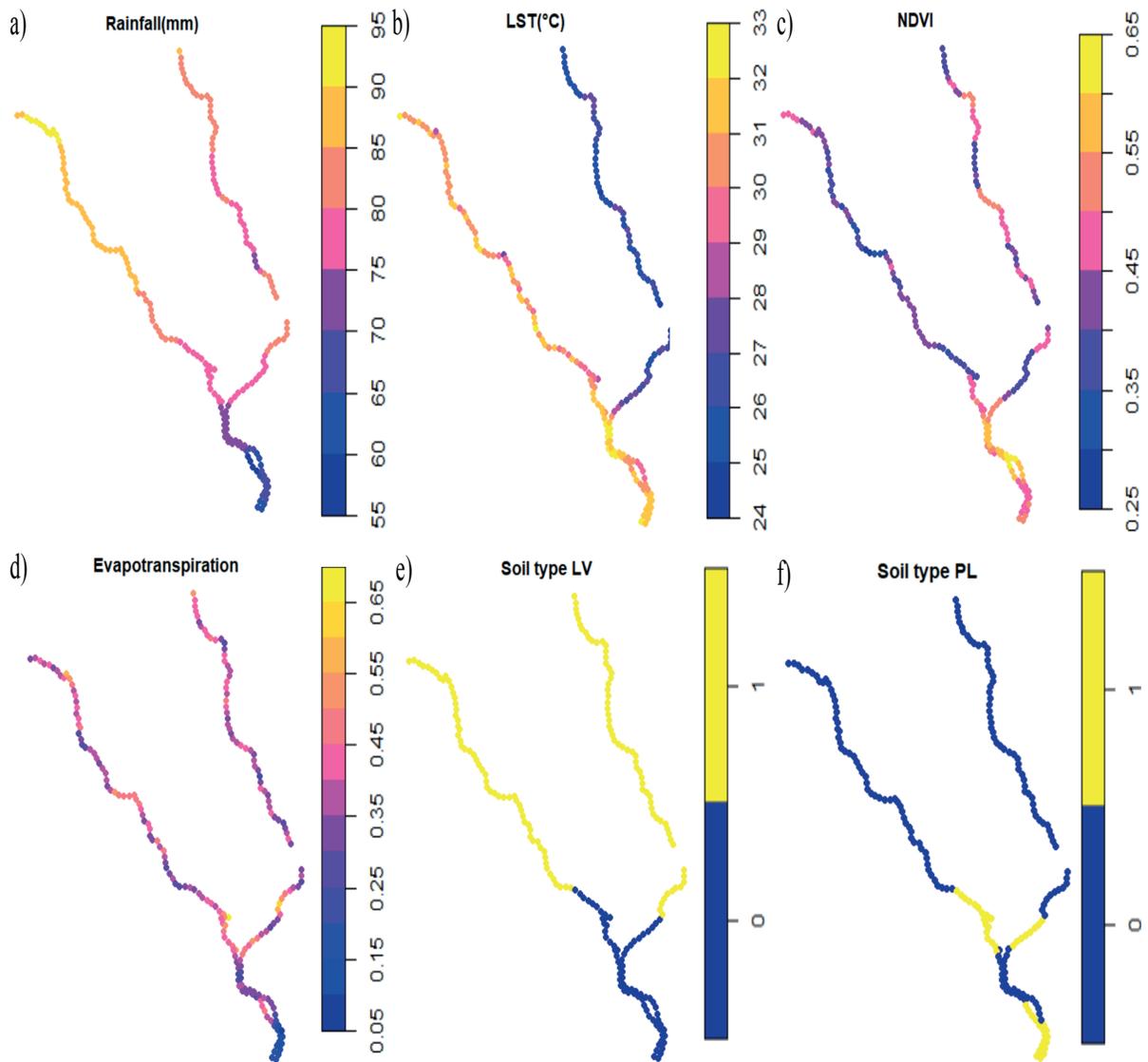


Figure 3.4: Environmental data values extracted for each prediction point. a) LST($^{\circ}C$) b) rainfall (mm) c) evapotranspiration d) Normalised Difference Vegetation Index (NDVI) e) Luvisolic (LV) f) Planosolic (PL). e and f are compared with Gleysolic(GL) soil type. Gap in shoreline is due to the removal of CM soil type.

3.3.3 Covariance function comparison

As shown in Appendix C: Figure C.8, the exponential quadratic covariance function was found to over-fit the model (smooth out the snail abundance excessively), and the Matérn ($\kappa=1.5$) smoothed the results, whereas exponential covariance function was the roughest fit of the model. Furthermore, there seemed to be no difference in predicted $\log(\hat{\mu}_i)$ against distance along shoreline for either *Biomphalaria* sp. or *Bulinus* spp. for the different covariance functions as shown in Appendix C.4 Figure C.9. This suggests that the effect of the covariates (environmental data) is more prominent than in the Gaussian process.

3.3.4 Model fit

The Bayesian log-linear Gaussian process model converged well according to the trace plots of the estimated parameters, and the priors were appropriately selected as shown Figure 3.5 and in Appendix C.5 Figure C.10.

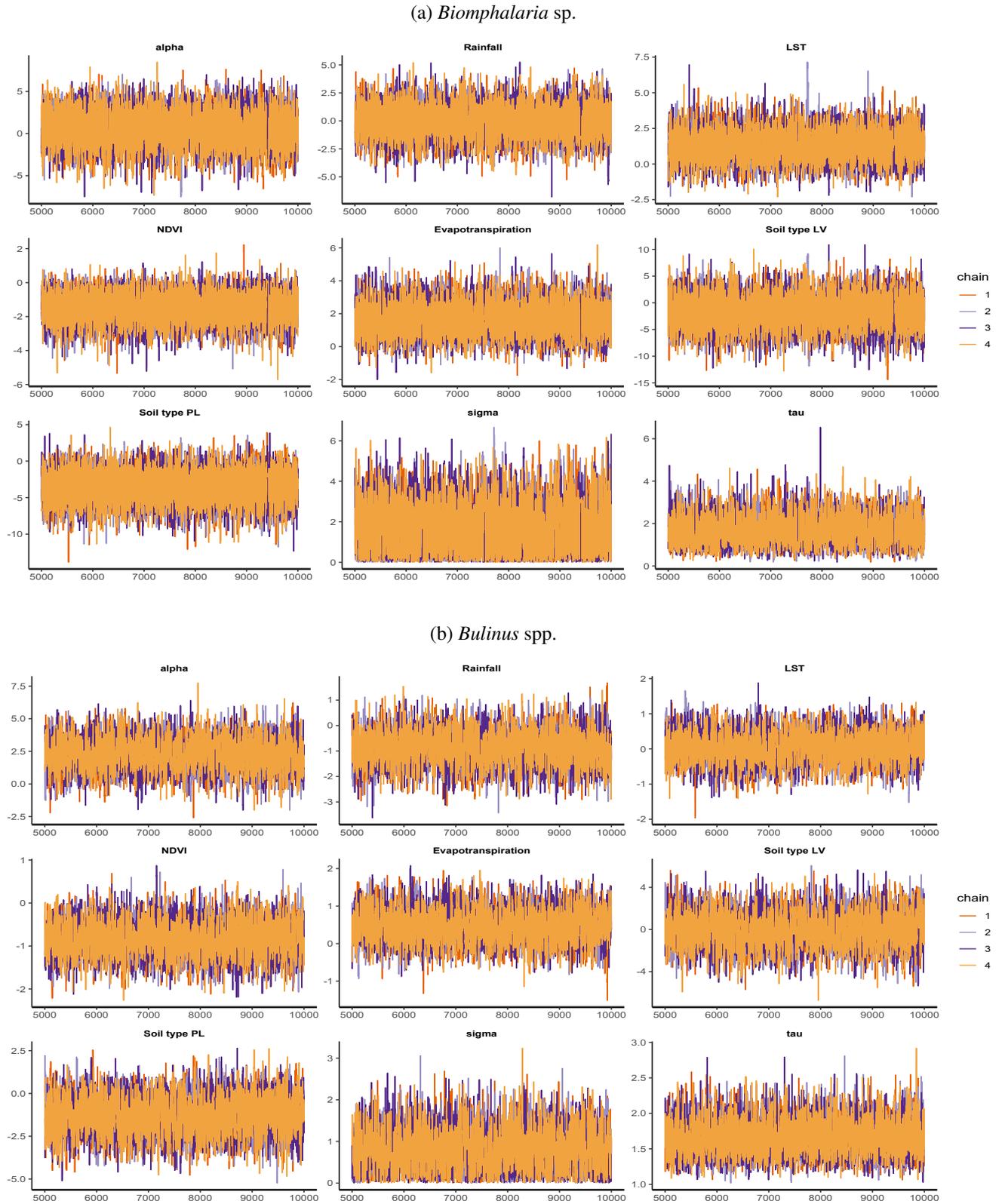


Figure 3.5: Trace plot of Bayesian log-linear Gaussian Process fitted model in Stan a) *Biomphalaria* sp. b) *Bulinus* spp.

3.3.5 Covariate effects

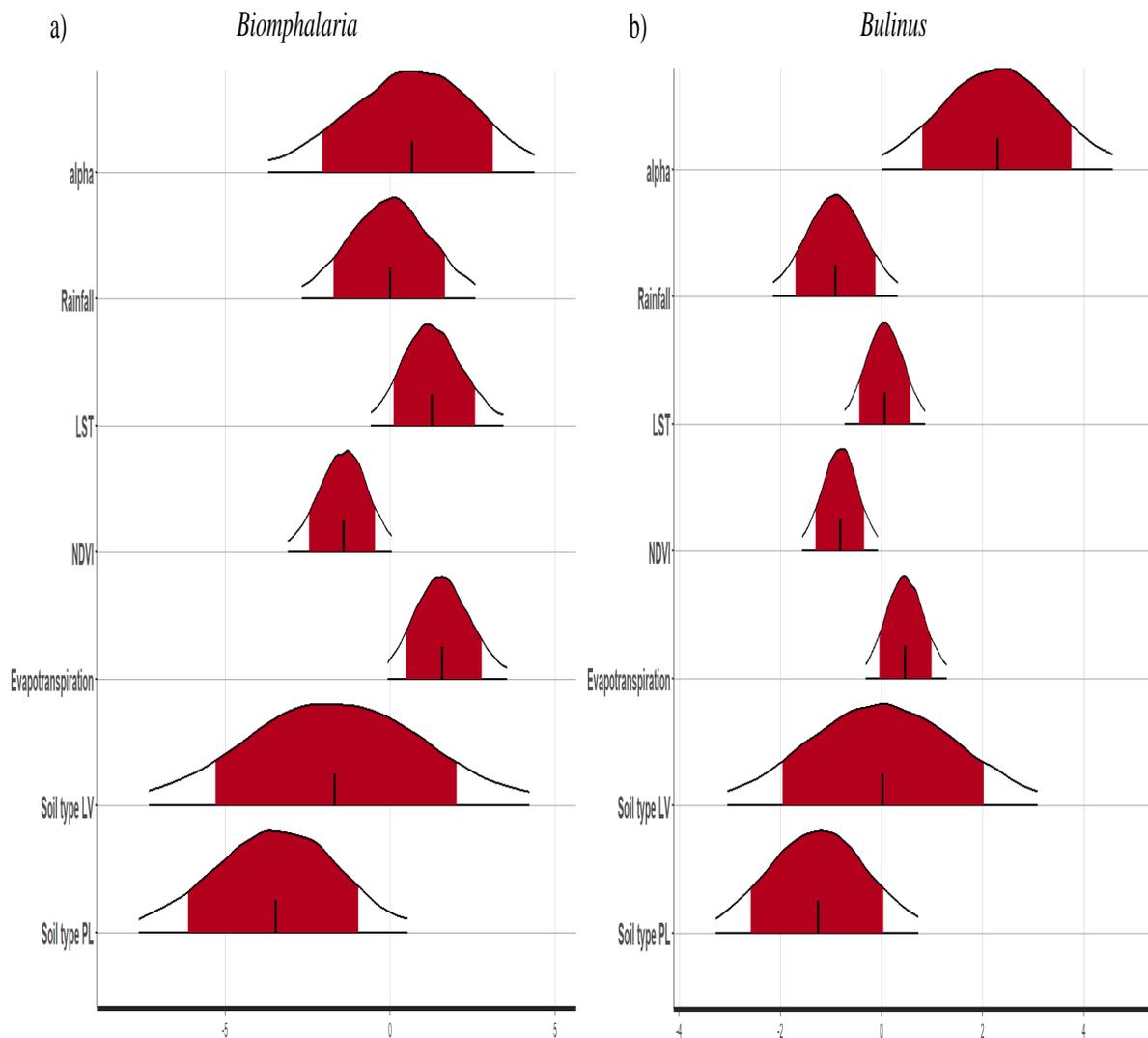


Figure 3.6: Posterior plot for each species a) *Biomphalaria* sp.; b) *Bulinus* spp. Red shaded area represents the 80% credible intervals (CrI) and the extent of the curve is the 95% CrI.

Figure 3.6 shows the posterior distributions for the environmental covariate effects (on the log scale) for each species of snail with 95% credible intervals (CrI). As shown in Figure 3.6 and table 3.1, a significant result was reported for NDVI, where 1-SD increase in NDVI had a -0.83 [CI: -1.57, -0.09] reduction in the $\log \mu_i$, mean *Bulinus* spp. snail abundance at location i .

All other covariates were not significant however, the following were still found of interest: For a 1-SD increase in the NDVI, the $\log \mu_i$ mean *Biomphalaria* sp. abundance changes by -1.42 [CI: -3.09, 0.10] (reduction). For a 1-SD increase in rainfall, the $\log \mu_i$ mean *Bulinus* spp. abun-

Table 3.1: Estimated parameter values, mean and CrI

Parameter	<i>Biomphalaria</i>		<i>Bulinus</i>	
	Mean	95% CrI	Mean	95% CrI
α	0.46	[-3.98,4.40]	2.32	[3.69,4.56]
Rainfall	-0.05	[-2.74,2.75]	-0.88	[-2.15,0.33]
LST	1.30	[-0.56,3.40]	0.04	[-0.72,0.81]
NDVI	-1.42	[-3.09,0.10]	-0.83	[-1.57,-0.09]
Evapotranspiration	1.61	[-0.04,3.56]	0.46	[-0.34,1.28]
Soil type LV	-1.51	[-7.13,4.08]	-0.04	[-3.07,3.11]
Soil type PL	-3.48	[-7.56,0.52]	-1.29	[-3.29,0.71]
Soil type GL				

a. Credible interval (CrI), Land surface temperature (LST), Normalized difference vegetation index (NDVI), Luvisols (LV), Planosols (PL), Gleysols (GL).

dance changes by -0.88 [CI: -2.15, 0.33] (reduction). For a 1-SD increase in LST, the $\log \mu_i$ mean *Biomphalaria* sp. abundance changes by 1.30 [CI: -0.56, 3.4] (increased). For a 1sd increase in evapotranspiration, the $\log \mu_i$ mean *Biomphalaria* sp. abundance changes by 1.61 [CI: -0.04, 3.56]; similarly, the $\log \mu_i$ mean *Bulinus* spp. abundance changes by 0.46 [CI: -0.34, 1.28] (increased). For an increase of PL soil type compared to the baseline GL soil type, the $\log \mu_i$ mean *Biomphalaria* sp. abundance changes by -3.48 [CI:-7.13,0.52] (reduction); similarly, the $\log \mu_i$ mean *Bulinus* spp. abundance changes by -1.29 [CI: -3.29,0.71] (reduction). No association could be found for LV soil compared to GL soil type.

3.3.6 Model predictions

For *Biomphalaria* sp., we predicted the greatest number of snails present to be close to Moet and Koche schools. For *Bulinus* spp., a higher number of snails was predicted over a wider area, close to Moet, Koche, Mtengeza, Chipeleka and Sungusya schools. However, for both *Biomphalaria* sp. and *Bulinus* spp., there was great uncertainty around all locations (2D version Figure 3.7; 1D version, Appendix C.7 Figure C.11).

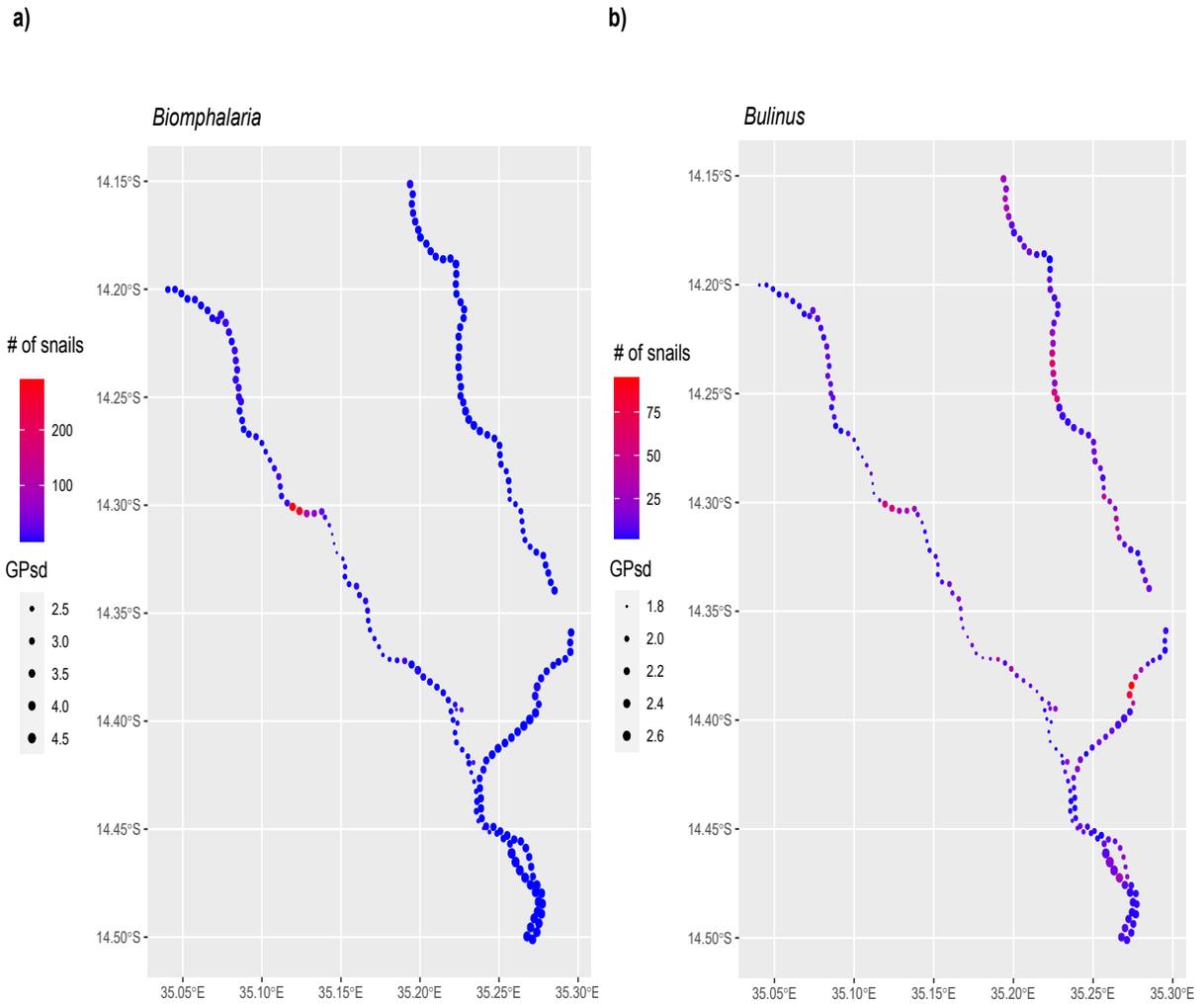


Figure 3.7: 2D mean GP prediction of number of snails $\log(\hat{\mu}_i)$ along the shoreline (km) a) *Biomphalaria* sp. b) *Bulinus* spp. Legend: Blue to red stands for exponential of mean GP number of snails. Dot size represents the standard deviation of the posterior predictive distribution at each vertex.

3.4 Discussion

Our secondary spatial analysis has made a seminal attempt to analyse, interpolate and then predict *Biomphalaria* and *Bulinus* snail distribution in unsampled locations in the southern part of Lake Malawi, Mangochi District. Our study found a significant negative association between NDVI and snail abundance for *Bulinus* spp. Analysis of our results are also indicative of a similar association between NDVI and *Biomphalaria* sp. abundance, although this was not significant given our currently available data. Other covariates considered in the model were all non-significant, as reported in Table 3.1; despite their uncertainty, we reported an increase in rainfall along the shoreline, which causes a reduction in the mean snail abundance found along the shoreline for *Bulinus* spp.. However, an increase in evapotranspiration and in LST along the shoreline, may each cause an increase in the mean snail abundance found along the shoreline for both *Bulinus* spp. and *Biomphalaria* sp.. For soil type, we found that an increase in PL or LV caused a reduction in the mean abundance found along the shoreline compared with GM. The characteristics of the shoreline of the southern part of Lake Malawi are known to vary considerably over focal areas (Figure 3.8) and in turn can increase or decrease snail abundance. We discuss our findings below upon consideration of other studies and establish how this could help to identify risk of schistosomiasis transmission risk locally.

In most previous studies, increasing vegetation (higher NDVI) was shown to have a positive association with snails found due to vegetation providing more suitable breeding sites, whereas our study suggests a negative association [160–162]. This difference in result is likely due to our focus on Lake Malawi, where molluscivorous fish may be present, as opposed to a more general area including smaller bodies of stagnant water, which typically lack such predatory fish.

The presence of land vegetation around the shoreline may well be descriptive of the land topology and hence the depth of the water in the immediate vicinity-deeper water is likely less conducive to snail habitats because of the absence of aquatic flora. Furthermore, the type of vegetation and whether it is submerged or nonemergent floating vegetation are known to be important as the freshwater snails need protection from wave action and food resources, aiding egg-laying, and this was not considered in our model [130, 163].

There is an indication that an increase in rainfall decreases snail abundance in our model despite its uncertainty. First this result could be due to the water flow increasing and spreading to new locations not yet colonized, causing the disruption of the freshwater snail habitats [161]. Second, an increase in rainfall has been reported to increase turbidity of water and in turn decrease the presence of snails [161, 162]. Lastly, increases in rainfall and water flow have also been reported to cause rapid changes in temperature causing thermal shock and reduced egg-laying of the freshwater



Figure 3.8: A collection of location photographs representative of the variation of the Southern part of Lake Malawi. Pictures taken during field work studies carried out during August 2022 showing the east side of southern part of the Lake Malawi shoreline (unpublished). Pictures taken by Alexandra Juhasz.

snails, causing an overall reduction in snail abundance [163].

In contrast to our result, some cases studies have found increase in snail abundance during increase in rainfall. For instance, when excess rainfall (flooding) occurs, new areas of snail habitat can occur where previously snails were not present or eliminated. Runoff water can create new pools adjacent to the shoreline or inland, allowing more breeding sites to be colonized by the intermediate snail host and thus increasing freshwater snail abundance [85]. Consequently, flooding can change the human-snail contact interplay, through an indirect effect on human behaviour, and thus the associated risk of schistosomiasis transmission [135, 164]. However, other studies have suggested that during flooding, these newly established pools of water can lead to humans visiting these new sites instead of the Lake Malawi with a possible decreased likelihood of snails being present already in these new sites, which could lead to a reduction in schistosomiasis transmission [160, 164]. Adding to the complexity, rainfall and water levels are known to oscillate over time, with a general decrease in lake levels reported more recently, with ongoing localised peaks of lake levels occurring through time [127]. This could impact the snail abundance and its presence spatially and temporally and indirectly affect human behaviour as mentioned before [127, 135, 164]. For example, if the lake levels are regulated by needs for hydroelectricity or because many individuals prefer to make contact with shallower and safe areas of the lake [165, 166].

Analysis of our results suggest that an increase in LST increases *Biomphalaria* sp. and *Bulinus* spp. snail abundance. Many laboratory studies have been carried out to try find the optimal temperature for snail survive. For *Biomphalaria* sp. snails the optimum temperature has been found to between 15 and 30 °C, where there is a decrease in snail abundance above 30 and 35°C, and no snails survive above 35°C [135, 165]. For our prediction points along the shoreline, the LST ranged between 25 and 32°C, which suggests *Biomphalaria* sp. snail abundance still increases above 30°C. This difference could be due to it being in natural environment where snails are able to adapt to climate change [167]. It has also been reported that freshwater snails move further into the lake when temperatures increase, which we did not consider in our model because it was constrained to the shoreline and buffer area [168].

Similarly, there is an indication that increased evapotranspiration increased the *Biomphalaria* sp. and *Bulinus* spp. snail abundance in our study. The increase in evapotranspiration, also known as the increase in evaporation of water, is known to have an impact on pH, salinity (salt concentration), conductivity and temperature of water through unpublished field studies, these finer physical characteristics need to be further investigated. This suggests an increase in evapotranspiration causes these unexplored covariates to become more habitable for intermediate snail hosts, causing an increase in local snail abundance. How these unexplored covariates interact and their effect on snail abundance are not considered in our study but have has been investigated in other studies

[85].

Our study found PL soil type decreases snail abundance compared to GL soil type. PL soil types are clay-based, plinthic soils with high concentrations of iron. GL soils are mineral soils, which are a mixture of sand, silt and clay. Both muddy when rainfall occurs (become water-logged) [169]. A previous study by Koch *et al.* 2004, found the opposite result with muddy soil being reported to improve the survivability of *Biomphalaria* sp. by preventing them from losing moisture in the hot and dry seasons compared with sandy ones and stony and decomposing material [86]. The difference between PL and GL soils, is that GL is known for its iron reduction [170]. Kulina *et al.* 2018 [171], reported an increase in *Schistosoma* transmission risk in groundwater with higher iron concentration [171]. We found a different result which suggested another chemical within the soil type could be interacting with the snail abundance and affecting transmission. Furthermore, there was uncertainty in our results. The soil types from SOTER database are for wide scales; lower level data are needed to improve the information on more localised soil types [157]. Other resources have been created, for example SoilGrids for Africa, which, if time permitted and it provided lower level data for southern Malawi, could be applied to our study in the future [172].

Our secondary analysis study shows substantive heterogeneities in snail distributions along the lake's shoreline, with certain schools being close to areas of increased abundance of snails. Hence SAC attending these schools may be more likely to be exposed to schistosomiasis. Moet and Koche schools were predicted to be nearest to the highest number of *Biomphalaria* sp. present along the shoreline, suggesting that more *S. mansoni* infections probably occur at these schools compared to the 12 other schools. Whereas, Moet, Koche, Mtengeza, Chipeleka and Sungusya were all predicted to be nearest to the highest number of *Bulinus* spp. However, for both *Biomphalaria* sp. and *Bulinus* spp., predicted presence along all the shoreline had large uncertainty. Furthermore we cannot be certain about the exposure risk for the SAC as this secondary analysis does not consider their water contact patterns, including where they visit (how far they travel to) the shoreline, frequency, type of contact and how long they remain at the shoreline. This needs to be further investigated as previous studies have reported increased snail abundance in localised areas where more water contact is occurring [168, 173]. In addition, the ability to measure exposure risk for SAC from our secondary analysis is dependent on presence of snails in an area being indicative that freshwater snails present are shredding cercariae, but it is difficult to be certain of this [163].

There are many more physical, chemical and environmental factors (abiotic and biotic) which could impact *Schistosoma* intermediate snail habitats and their relative snail abundance; these were not considered in our model, because of time constraints or non-accessible data e.g. pH, salinity, conductivity, flow velocity, turbidity, calcium and bicarbonate concentration, dissolved oxygen, soil density and water capacity [163, 164, 174]. Furthermore, other factors such as food source,

pollution (e.g. discarded plastics), parasitism and even the competition for snail habitat with other organisms within an area were not considered in our model [174]. Variation in human movement patterns can make it difficult to locate the location of acquired infected. A land use and human influence index could have been included in our model if time had permitted [83].

One limitation of our study is the restricted study period (November 2017 to June 2019, except for 5-year evapotranspiration time frame) as well as taking the mean values for each prediction location. Rabone *et al.* 2019 [168] reported seasonality affecting snail abundance, with higher snail abundance during the dry season compared to the wet season. For instance, seasonality can affect growth of vegetation and therefore the freshwater snail's life due to the variation in sunlight, therefore leading to changes in snail abundance [174]. In the future, we would like to investigate how seasonality affects the snail distribution using our model. We reported on the seasonal changes of the covariate data in Appendix C.2 (Figures C.3, C.4 and C.5); this allows observation on how covariate data change over time, although this was not considered in our model.

As mentioned before, another known limitation is that snails are not only found on the shoreline of Lake Malawi but also in pools adjacent to the lake or rivers, ponds and streams. This has been reported to affect snail abundance by affecting the microhabitat, for instance by changes temperatures [168, 175]. Unpublished field work studies in 2021, showed that on southern lake slopes in areas, the western side of the shoreline had longer shallower areas. The area near the Upper Shire River is known to have more vegetation and swampy areas than the rest of the shoreline. Bathymetric data for water depth were originally considered in our secondary analysis, taken from the GLOBal Bathymetric (GLOBathy) dataset, which relies on HydroLakes dataset [175]. However, it was excluded from the study because of missing River Shire values as shown in Appendix C.8 (Figures C.12a, C.12b, C.13 and C.14). Therefore, water depth needs to be further investigated. As mentioned before, water levels are known to vary over time, leading to changing water depth. *Bulinus* spp. and *Biomphalaria* sp. have different preferences regarding water depth and vegetation [11, 127, 134].

An important main limitation of our analysis is the resolution of the raster data we used as covariates. Many remotely sensed metrics are known to be inaccurate over water. Therefore we positioned our shoreline linestring just inland of the water's edge. Thus, any associations between land-based measurements and habitat conditions in the water are likely to only indirectly affect snail abundance. A repeat study, using direct observations of shoreline habitat composition, perhaps using towed arrays of sensors behind a boat or done directed close to the water's edge, may be able to provide a more accurate map of predicted snail abundance.

3.5 Conclusion

Our study provides a preliminary method of predicting the abundance of *Biomphalaria* sp. and *Bulinus* spp. snails along the shoreline of Lake Malawi, given malacological data collected at sparse locations and remotely sensed environmental data. Furthermore, our study shows substantive heterogeneities in snail distributions along the lake and abundance information which may be used to develop further statistically grounded study designs to improve the identification of likely snail habitats posing a high risk for schistosomiasis transmission.

Chapter 4

Development of a dynamical model to enhance understanding of epidemiology of schistosomiasis in school aged children

Authors list: Amber L Reed¹, Sekeleghe A Kayuni^{2,3}, Janelisa Musaya³, J Russell Stothard², Claudio Fronterre¹, Chris Jewell⁴

Author Affiliations

1. Lancaster Medical School, Lancaster University, Bailrigg House, Bailrigg, Lancaster, LA1 4YE
2. Tropical Disease Biology, Liverpool School of Tropical Medicine, Pembroke PI, Liverpool L3 5QA
3. Malawi Liverpool Wellcome Trust Programme of Clinical Tropical Research, Queen Elizabeth Central Hospital, College of Medicine, P.O.Box 30096, Blantyre, Malawi
4. Mathematics and Statistics, Lancaster University, Bailrigg House, Bailrigg, Lancaster, LA1 4YE

Keywords: Schistosomiasis, *Biomphalaria*, *Bulinus*, Transmission dynamics, Optimisation

Abstract

Background: School-aged-children are most vulnerable to schistosomiasis as exemplified by age-prevalence profiles although current understanding of these patterns needs improvement. Using epidemiological data from the southern shoreline of Lake Malawi, we investigated the dynamics of *Schistosoma* transmission and the main determinants of *Schistosoma* infection risk using transmission dynamic model and considered urogenital and intestinal schistosomiasis, respectively. Specifically, we assessed whether the proximity of primary schools to the immediate shoreline was a major geospatial and epidemiological determinant.

Methods: Cross-sectional parasitology and malacological data previously collected and analysed was used, including age-infection profiles and interpolated predicted snail distributions for the southern part of Lake Malawi shoreline. A disease SEIRS ordinary differential equation model was created, and an observation prevalence model was formed using a binomial sampling distribution using the already published dataset. An optimisation using L-BFGS-B algorithm with upper/lower bounded box constraints was carried out to calibrate the model to find the best parameter values for each infection state transition given the disease model and dataset. The aim was to recapture the age-structure dynamics shown in the observation model representing the already published age-infection profiles.

Results: Concerning intestinal schistosomiasis, the best model for *Biomphalaria* sp. was the use of a single transmission rate for all the school's and no spatial effect. By contrast, for urogenital schistosomiasis, the best model for *Bulinus* sp. was found when using an independent transmission rate for each school and no spatial effect.

Conclusion: There was some evidence that we were able to capture the age-structured dynamics of infection in SAC despite the expected outcome differing to statistical output due to sparse data. Within our study area, we found there was no significant effect on SAC exposure to *Schistosoma* infection risk based on school distance from the shoreline. Further, there was heterogeneity between schools in transmission rates estimated, although these did not have significantly different confidence intervals. However, schools considered in our study were all relatively close to cercaria infested shoreline. Further studies using a longitudinal cohort study could improve understanding of *Schistosoma* infection dynamics and allow for improved control method application.

4.1 Introduction

Schistosomiasis is a water-borne neglected tropical disease caused by the trematode worm *Schistosoma*. Two forms of the disease exist, with intestinal schistosomiasis (IS) caused by *S. mansoni* and urogenital schistosomiasis (UGS) caused by *S. haematobium* infections. To complete their lifecycle, these species of *Schistosoma* require the presence of aquatic snail intermediate hosts, *Biomphalaria* sp. and *Bulinus* spp., respectively. Thus, exposure to the snail habitats is a known major risk factor for human *Schistosoma* infection [17]. School-aged-children (SAC) are known to be particularly vulnerable to schistosomiasis, with signs and symptoms of infection including malnutrition, anaemia, and neurological and developmental delays caused by the accumulation of trapped eggs causing tissue inflammation [121]. Children are thought to be first infected soon after birth upon freshwater contact(s) with prevalence increasing with cumulative parasite exposure(s) up to adolescence, although the age-profiles are known to oscillate over time due to many undefined factors [17].

There are many known risk factors for *Schistosoma* infection transmission including repeated water contact, type of water use, animal contact, age and treatment [17, 125, 149, 176]. Recent studies have reported an increase in transmission of *Schistosoma* infection through a changing ecological environment in the lake, an increasing human population and reduced molluscivore fish in the lake. This has created an increase in snail populations and new at-risk locations for water contact [138]. Changes in human behaviour affect age-dependent exposure, especially among SAC who are known for frequent water contact [127, 138, 177]. Reitzug *et al.* 2023 [177], reported its importance in human exposure behaviour in driving *Schistosoma* transmission [177]. Other factors such as schistosomiasis infection immunity, formally known as resistance to reinfection, is still equivocal and being researched. Human immunity developing over long repeated exposure to *Schistosoma* infection and how this interacts with age-dependent exposure is much debated [178]. Currently studies have reported partial immunity building up in individuals over time [5, 25, 53, 178]. Most SAC are considered to have very low to no acquired immunity allowing for reinfection to occur. This is possibly shown by the Reed *et al.* 2023 [176] (Chapter 2) study where the age profiles of SAC increase up to 11 years before decreasing there afterwards, although this is conjecture and requires more investigation [176, 179].

People are more likely to make use of water facilities close to where they live and if that water facility happens to be a *Schistosoma* infection transmission site then the risk is likely to be higher for these people than people living further away leading to a higher need for control application in communities closer to these high-risk areas [180]. Further, Madsen *et al.* 2011 [76] reported higher *Schistosoma* transmission in shoreline villages compared to in-land villages [76]. Further-

more, other studies have also found the proximity to the lake to increase *Schistosoma* transmission risk [78, 181]. Ecological niche mapping and fine-scale malacological mapping of areas have previously been used as a technique to try to measure exposure risk dependent on location; these studies find that there is localised variation in areas with different biological effects impacting the transmission for each focal location [182, 183]. Other studies have researched water-contact related activities and association with exposure risk [177]. However, there is a lack of knowledge of how to measure SAC age-dependent exposure and the associated *Schistosoma* transmission risk, which is limiting appropriate application of control methods [176]. The difficulty of measuring this exposure could be due to multiple factors, for example, immunity, age of child, water exposure patterns and treatment [110]. Currently prevalence among SAC is measured at survey school locations using parasitological methods to measure prevalence and intensity of infection [184]. Further as mentioned before, age prevalence profiles among SAC have been studied along the southern part of the Lake Malawi shoreline and reported to oscillate over time [176].

Mathematical modelling can allow us to model the spread of disease and transmission dynamics to inform policy decision makers for intervention programmes. Many mathematical modelling studies have been carried out over the years to aid control programmes [111, 185–187]. Both Nelson & Macdonald carried out pioneering work on transmission dynamics of *Schistosoma* infection [106–108]. Anderson *et al.* 2016 [110] recently reviewed transmission models and control of *Schistosoma* infection by Mass Drug Administration (MDA). The World Health Organisation (WHO) has supplied new guidelines to target elimination of schistosomiasis by interrupting transmission [63]. Intrinsic factors such as age and their related exposure have a possible importance in interrupting transmission when moving towards elimination [116].

In this study, we developed a susceptible - exposed - infection - recovered - susceptible (SEIRS) dynamical model with age-related-immunity to model the transmission dynamics of *Schistosoma* infection for SAC aged 6 to 15 years, using published *Schistosoma* age-stratified prevalence [176], and considering proximity to snail-infested water as a proxy for exposure to the *Schistosoma*. We fit the SEIRS model over prevalence data via maximum likelihood to improve our understanding of *Schistosoma* transmission dynamics and what drives SAC age-profiles. The main aims of the study were the following: i) To identify the main determinants of *Schistosoma* infection risk ii) To determine whether school distance from the lake shoreline determines the exposure of SAC *Schistosoma* risk. iii) To improve our understanding of the non-linear relationship between age and prevalence found by Reed *et al.* 2023 [176], and whether this was partially due to the exposure to the snails or immunity develop by the SAC. Hence, we produce models with immunity to try to reproduce the age-profiles of these SAC found in the previous study [176]. Consequently, we hope this analysis will help improve understanding of *Schistosoma* transmission dynamics and in turn,

improve the application of schistosomiasis control within SAC.

4.2 Methods

In this section we describe our modelling approach. We provide a brief description of our schistosomiasis prevalence training dataset, the state transition model used to model the disease process in the school, and how this relates to the observation process. All data processing and analysis was performed in R version 4.1.1.

4.2.1 Dataset

This secondary analysis study uses the outcomes of our previously published, Reed *et al.* 2023 [176] on *Schistosoma* age prevalence profiles (see Chapter 2), and Reed *et al.* 2024 [188] on geospatial modelling of snail distributions based on the parasitological and malacological data from Kayuni *et al.* 2020 [127] and Al-Harbi *et al.* 2019 [11] (see Chapter 3). These data were collected in 2019 for the age-profiles and the snail distribution predictions were interpolated from aggregated data from 2017 to 2019 along the southern part of the Lake Malawi shoreline. Therefore, in the study we use cross-sectional single point in time data.

We removed St Augustine 2 school from *Biomphalaria* sp. dataset as it had no variation in prevalence (effectively 1). For *Bulinus* spp. all schools in the dataset were used including St Augustine 2.

4.2.2 Disease model formulation

We consider the incidence of schistosomiasis in SAC within a school as a function of childrens' exposure to snail-infested waters. To investigate how their age and proximity to snail habitats affects the age-prevalence relationship in SAC, we develop an age-stratified SEIRS model where children are assumed to start susceptible (S) to infection, progress to being exposed (E) (i.e. infected but not yet infectious), infectious (and detectable, I), and recovered (R) with immunity as shown in figure 4.1. We then allow for immunity to wane and the children to return to being susceptible.

Within each school we assume that children are divided into 10 age-grades, with a class of size of 30, represented as m . Children are assumed to enter school at age 6 at a rate of $\frac{1}{30}$ per year (i.e. 30 children entering the school per year) having had no prior infection by *Schistosoma* spp. ($\alpha = 0\%$)

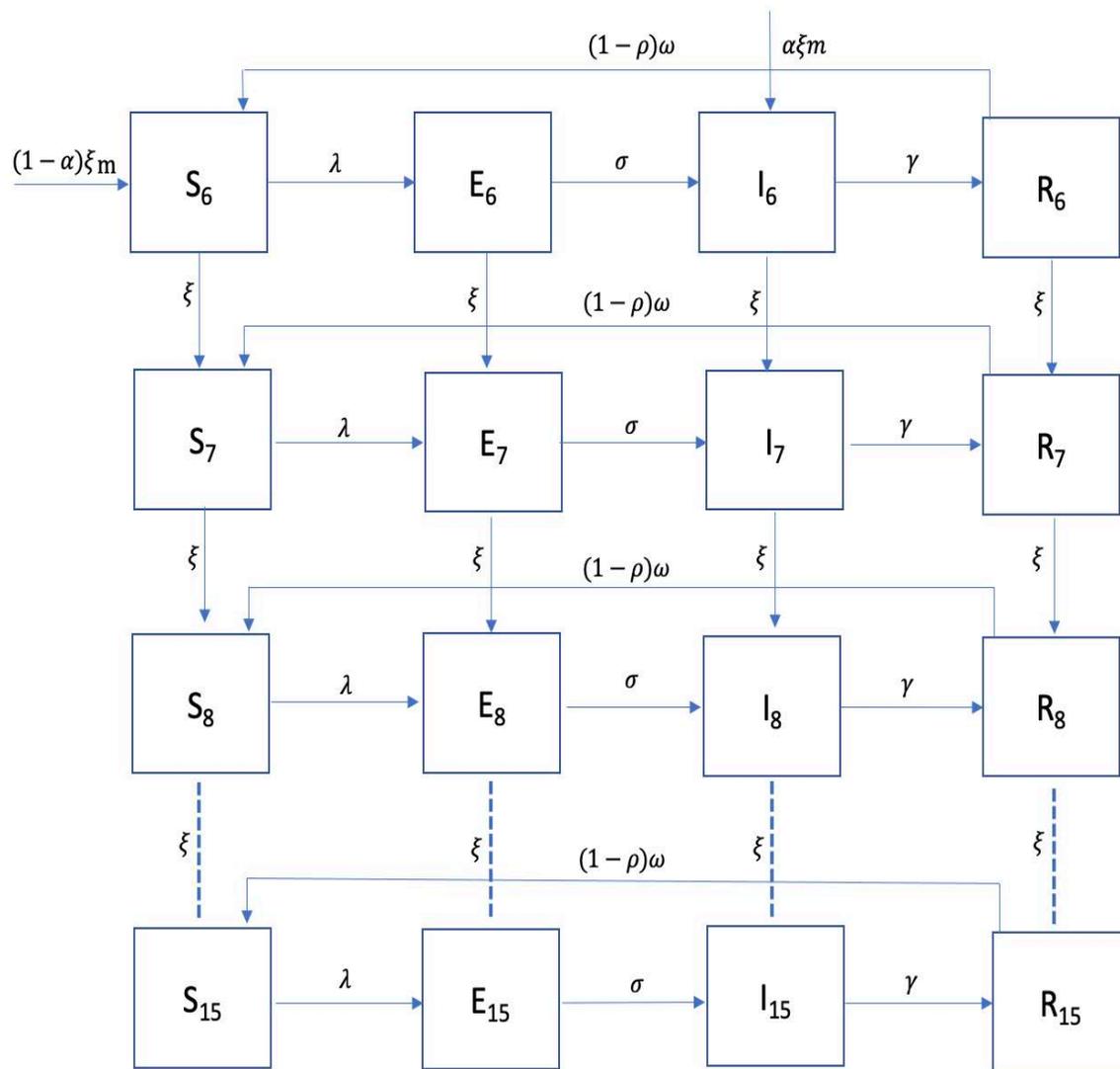


Figure 4.1: SEIRS transmission compartment model with SAC age ranging from 6 to 15. Transmission parameters are discussed in the text.

infection. We also test the assumption where prior *Schistosoma* infection for SAC entering into age group 6, α by testing different values (see sensitivity section). They then progress between the age classes at a rate of $\frac{1}{365}$ per day giving ξ as an aging rate of the SAC. We assumed SAC only enter the school via age 6 and do not leave the school till aged 15.

Given the states, we model the rate of transition between S and E according to a function of exposure to snail habitats, mediated by distance to the lake shore. Letting λ_i be the infection rate experienced by a child in location i , we have:

$$\lambda_i = \int \beta K_{ij} a(x_j) dx_j \simeq \frac{\beta}{n} \sum_{j \in x} K_{ij} a(x_j), \quad (4.1)$$

where β is the transmission rate coefficient for *Schistosoma* infection, $K_{ij} = \exp\left(-\frac{\|x_i^2 - x_j^2\|^2}{\phi^2}\right)$ is a function that decays with Euclidean distance between i and j , $a(x_j)$ is the snail abundance, x_j is the location of snails (see appendix C Figure C.1) and x_i is the location of the school for each SAC, i for $i = 1, \dots, n$. Contact with free living cercariae was assumed to be the only means through which children become infected (i.e. transition from S to E). Furthermore, all snails were assumed infected and snails' infection status is assumed independent of contact with SAC. In addition, we assume that SAC do not infect each other, and no mother-to-child transmission of disease occurs among humans. These assumptions are supported by the fact that we only model school children, who are themselves a small fraction of the wider population and can therefore be assumed to be approximately independent within the overall human-snail-schistosome lifecycle.

Then we model the rate of transition from E to I as σ , which is the rate at which immature worms become adult worms residing in intestinal or urogenital region and producing eggs. Then we model the rate of transition from I to R, with γ as the rate of infective SAC moving to the R. Next, we assume that the rate of age-specific loss of immunity, $(1 - \rho(\text{age}))\omega$, i.e. age-related, such that the rate of transition from R to S (Figure 4.1) follows a logistic growth rate with an inverse relation to age represented as:

$$\rho(\text{age}) = \frac{1}{1 + e^{-\kappa(\text{age}-C)}}, \quad (4.2)$$

where κ governs the rate of change with respect to age, and C is a constant which controls the midpoint i.e. the age at which $\rho(\text{age}) = \frac{1}{2}$. For parameter identifiability reasons, we assume C to be 11 years which is the midpoint between age groups 6–15 years and was the age of peak infection prevalence found in the published secondary analysis paper, Reed *et al.* 2023 [176](see Chapter 2) on *Schistosoma* prevalence age-profiles. We assumed the older the child, the slower they transition

from R to S. In any transition of the model, no deaths from schistosomiasis or MDA treatment were considered.

The age-structured dynamics of schistosome transmission in this system are then represented by the following differential equations:

$$\frac{d\vec{S}_t}{dt} = -(\vec{\lambda} + \xi) \odot \vec{S}_t + \omega(1 - \vec{\rho}) \odot \vec{R}_t + \xi \vec{S}_t^+, \quad (4.3)$$

$$\frac{d\vec{E}_t}{dt} = \vec{\lambda} \odot \vec{S}_t - (\sigma + \xi) \vec{E}_t + \xi \vec{E}_t^+, \quad (4.4)$$

$$\frac{d\vec{I}_t}{dt} = \sigma \vec{E}_t - (\gamma + \xi) \vec{I}_t + \xi \vec{I}_t^+, \quad (4.5)$$

$$\frac{d\vec{R}_t}{dt} = \gamma \vec{I}_t - (\xi + \omega(1 - \vec{\rho})) \odot \vec{R}_t + \xi \vec{R}_t^+, \quad (4.6)$$

where, t stands for time in days.

The age structure of the dynamical model are represented as the following equations:

$$\vec{S}_t^+ = ((1 - \alpha)m, S_t^6, S_t^7, S_t^8, S_t^9, S_t^{10}, S_t^{11}, S_t^{12}, S_t^{13}, S_t^{14})^T, \quad (4.7)$$

$$\vec{E}_t^+ = (0, E_t^6, E_t^7, E_t^8, E_t^9, E_t^{10}, E_t^{11}, E_t^{12}, E_t^{13}, E_t^{14})^T, \quad (4.8)$$

$$\vec{I}_t^+ = (\alpha m, I_t^6, I_t^7, I_t^8, I_t^9, I_t^{10}, I_t^{11}, I_t^{12}, I_t^{13}, I_t^{14})^T, \quad (4.9)$$

$$\vec{R}_t^+ = (0, R_t^6, R_t^7, R_t^8, R_t^9, R_t^{10}, R_t^{11}, R_t^{12}, R_t^{13}, R_t^{14})^T. \quad (4.10)$$

For a given set of parameters, $\theta = (\beta, \sigma, \gamma)$, we solve this set of ODEs using Euler's method as implemented in the R package "deSolve", R version 4.1.1 (reduce computation time required). The solver is run until the SEIRS system achieves equilibrium, which by experimentation we find to be by timestep $t = 1000$ days for a large range of parameter values.

For simplicity, and exposition of the inference methods in the next section, we abstract the ODE model into a mathematical function, of the parameters θ and initial conditions $X_0 = \{\vec{S}_0, \vec{E}_0, \vec{I}_0, \vec{R}_0\}$.

$$S^*, E^*, I^*, R^* = g(\theta, x_0, z), \quad (4.11)$$

where the vector S^*, E^*, I^*, R^* denotes the number of children in each age-group in each epidemiological state at equilibrium, z represents our input data (snail abundances, distance to the shoreline), and $\theta = (\beta, \sigma, \gamma)$ our model parameters. Where necessary, we subscript these quantities to denote that they relate to a specific school, for example $\theta_s = (\beta_s, \sigma, \gamma)$ to denote the condition where all schools share σ and γ but have individual β_s for the s th school.

4.2.3 Observation of prevalence

From our study we have observation of childhood infection prevalence in school. For school s , given that $S_s^*, E_s^*, I_s^*, R_s^* = g(\theta, x_0, z)$, we assume observed number of positive children y_{as} in age-group a given a sample of n_a children is Binomially distributed given that

$$y_{as} \sim \text{Binomial}(n_{as}, \pi_{as}), \quad (4.12)$$

where $\pi_{as} = I_{as}^*/N_{as}$ where I_{as}^* is the modelled number of infected children in age-group a in school s , and $N_{as} = 30$ is the size of the class.

4.2.4 Inference

In the model describe above, we have parameters β, σ, γ which remain unknown. In order to estimate these we fitted the model using maximum likelihood estimation. Since the ODE model is a deterministic function of the input parameters and covariate data, the log-likelihood is a product over the 10 age groups and 12 schools

$$l(\theta; \pi, n) = \sum_{a=1}^{10} \sum_{s=1}^{12} y_{as} \log(\pi_{as}) + (n_{as} - y_{as}) \log(1 - \pi_{as}) + k, \quad (4.13)$$

where k is a constant. We calculated estimates for θ by optimising the log-likelihood numerically using the L-BFGS-B method provided by R's "optim" function. Parameters bounded at zero were log-transformed for ease of optimisation, with marginal log-likelihoods generated for each param-

eter to visually inspect the quality of optimisation and suggest starting points for the optimisation algorithm. Parallel computing was used to quicken the computer processing time, exploiting conditional independence of the schools in the model, enabling the ODE solvers for each school to run simultaneously.

Due to data in this model being cross-sectional data for a single point in time, we cannot easily identify a timescale for the disease process. If we double all transition rates, we find the same steady state conditions. To accommodate for this, we chose to fix κ and ω for identifiability purposes. Since β , γ , and σ are rate parameters, and therefore have positive support, they were optimised on the log scale, e.g. optimising $\beta^* = \log(\beta)$. This improved the efficiency of the optimisers by constraining the parameter space to be positive. κ was fixed to be $\kappa = 0.5 \text{ days}^{-1}$ as this allowed for a moderate age-related loss of immunity slope in the logistic regression function with $C = 11.0$ years as shown in Figure 4.2. ω was fixed to be $\omega = 0.5 \text{ days}^{-1}$ such that the rate of loss of immunity was approximately 0.5 days^{-1} for 6-year-olds. We also fixed the spatial decay parameter ϕ which enabled the identification of the β parameters given the length scale. ϕ was fixed at $\phi = 4.48\text{km}$ as represents spatial decay to 0 by 100km (the approx. distance along the shoreline of Lake Malawi for our data used) and the distance was divided by 10km to allow for numerical stability.

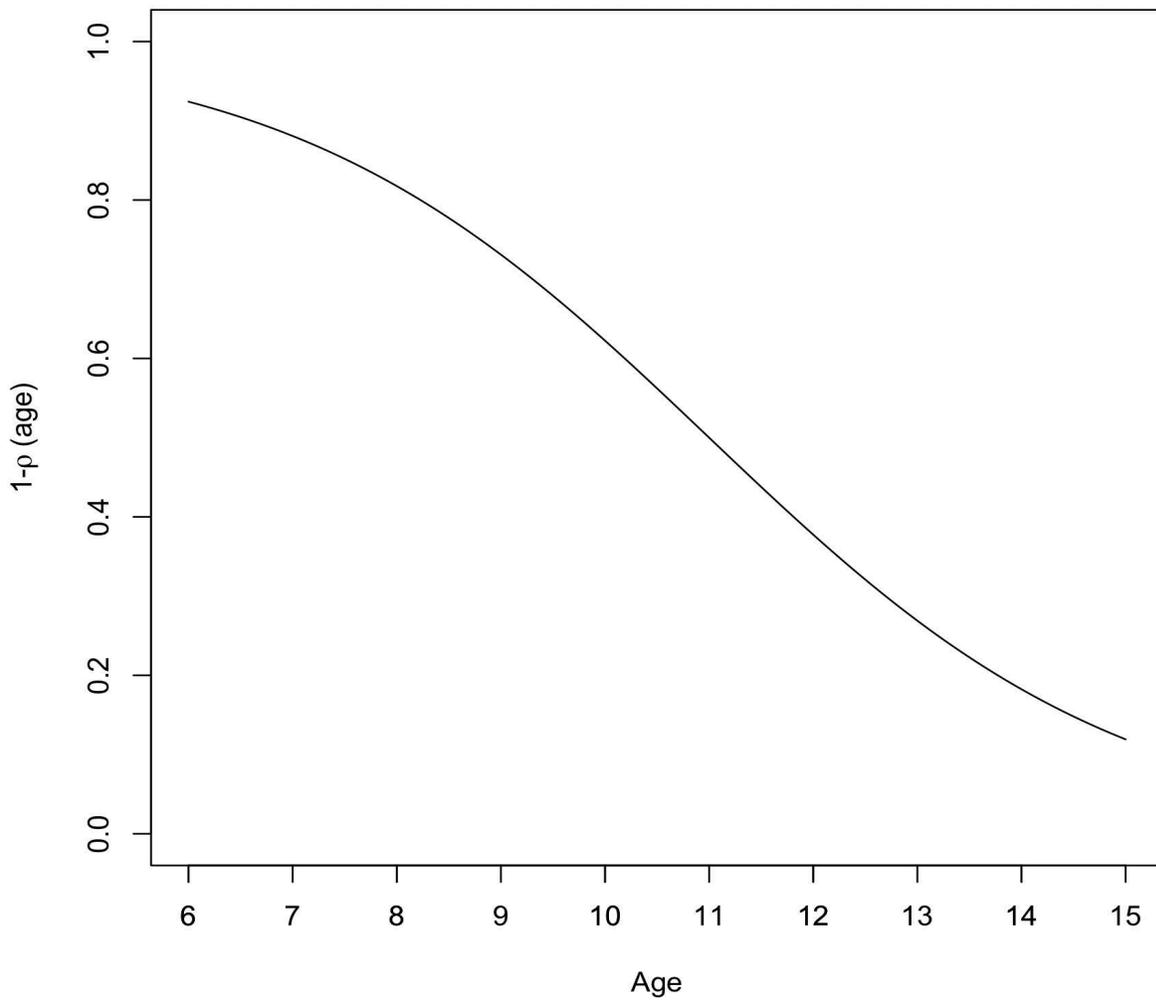


Figure 4.2: Logistic age-specific loss of immunity function, $1 - \rho(\text{age})$ versus age. For $\kappa = 0.5$ days⁻¹, $C = 11$ years, age ranging between 6 and 15 years.

To allow the optimiser to find the best-estimate values of the parameters in our models with reduced computation time, we set initial values for the non-fixed parameters by plotting the marginal log-likelihoods and starting the optimiser near to the maximum (peak) of the log-likelihood. We constrained the optimiser search space by ± 1 either side of the graphical maximum to stop the optimiser entering unstable regions of the parameter space [189].

The fitted optimiser estimates were plotted against the joint log-likelihood to check if the maximum log-likelihood (MLE) was found. Optimisation of certain parameters was carried out as shown in Table 4.1.

The following four models were fitted for each species using the optimiser as shown in Table 4.1.

- 1) different (independent) β_s value (transmission rate) for each school s with spatial effect
- 2) single- β value for all schools with spatial effect
- 3) no spatial effect by making ϕ large with multiple (independent) β_s 's values for each school
- 4) no spatial effect by making ϕ large with single β 's values for each school.

In the third and fourth model we removed the spatial effect component, modifying the snail force of infection to test whether we still need a model with multi- β_s 's or single- β for each school.

We then calculated the Akaike Information criterion (AIC) to compare the fit of different models. Then the fit of each model was plotted for each school against age, as well as the observed prevalence against age. In addition, to assess the parameter estimates further, mean squared error (MSE) was computed for the parameter estimates for each school to compare between the observed and predicted prevalence using the following:

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2, \quad (4.14)$$

where a vector of n predictions were sampled from our n data points ($i = 1, \dots, n$) for all variables, Y_i is the vector of the observed values of prevalence (at school location x_i) that are being used for the prediction and \hat{Y}_i are the predicted parameter values of model. Then we took the mean of (4.14) (e.g. square of the errors) to find MSE value for each model and compared them.

We also computed the hessian matrix to allow us to find the approximate confidence intervals (CI) using the Wald method for multi- β_s models (spatial and no spatial effect). Then we plotted the estimated coefficients values (parameters) and the CI together.

Table 4.1: Parameters used for the 4 models fitted to *Bulinus* spp. and *Biomphalaria* sp. data: single vs school-specific β , with and without the spatial term

Parameters	Definition	Multi- β space	Single- β space	Multi- β no space	Single- β no space
β	Transmission rate coefficient	Multiple	Single	Multiple	Single
σ	Rate of exposed SAC becoming infective (shredding eggs)	Unknown	Unknown	Unknown	Unknown
γ	Rate of infective SAC Recovering	Unknown	Unknown	Unknown	Unknown
ξ	School recruitment rate and rate at which SAC age	$1/365 \text{ days}^{-1}$			
κ	Smoothness of age-dependent immunity loss curve (logistic curve)	0.5 days^{-1}			
C	Constant (Age range of SAC)	11.0 years			
ω	Rate of loss of immunity of SAC aged 6 years	0.5 days^{-1}			
ϕ	Spatial decay constant	4.48km	4.48km	10^6	10^6

4.2.5 Sensitivity analysis

A sensitivity analysis was carried out by allowing for prior *Schistosoma* infection prevalence of SAC entering age group 6, α as 5% (0.05), 10% (0.10) and 20% (0.20).

4.3 Results

4.3.1 Check best value of ϕ

If we fixed all parameters to the estimated values from multi- β_s space model and unfixed ϕ the model converges quickly and finds the best value of $\phi = 4.48km$.

4.3.1.1 Snail abundance

Figure 4.3 shows that the SAC snail abundance exposure decreases till $\sim 100km$ distance from school to shoreline when $\phi = 4.48km$. Whereas, when $\phi = 10^6$, the snail exposure rate is the same all along the shoreline, which shows there is no spatial effect in this case.

4.3.2 Model optimisation

All models converged to the maximum likelihood for each school as shown in appendix D.2 Figures D.31, D.32, D.33, D.34, D.35 and D.36, although there were some identifiability issues be-

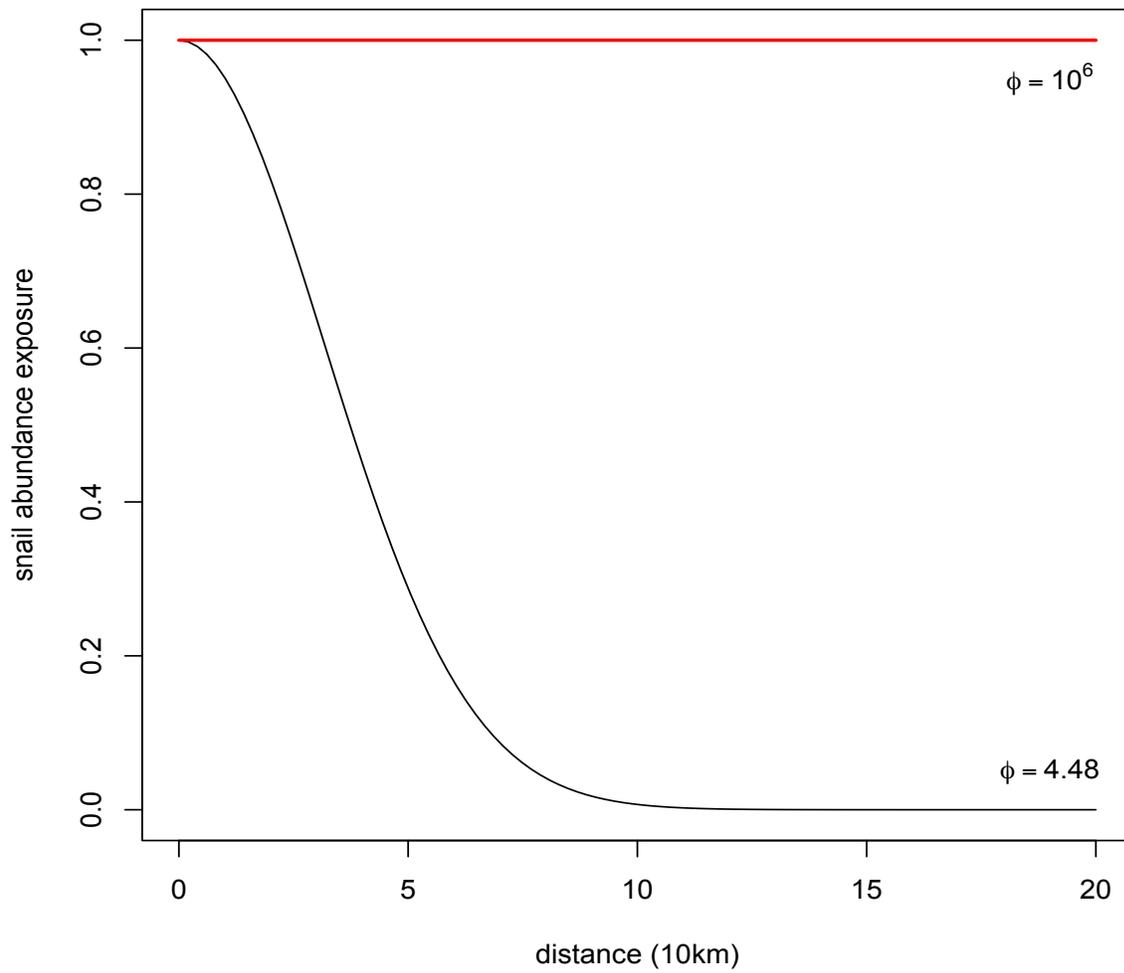


Figure 4.3: Distance (10km) of school from lake shoreline versus snail abundance distance. Black line: $\phi = 4.48km$, red line: $\phi = 10^6$.

tween β and γ . As shown in tables 4.2, 4.3, 4.4 and 4.5 the following was found: the best model (AIC=250) for *Biomphalaria* sp. was the single- β with no spatial effect. When testing multi- β_s or single- β with spatial effect (AIC=258 & AIC=269) and multi- β_s with no spatial effect (AIC=250) we found a worse model fit compared to single- β with no spatial effect (AIC=250). Whereas for *Bulinus* sp. the best model (lowest AIC) was the multi- β_s with no spatial effect (AIC=248). When testing the multi- β_s with spatial (AIC=250), single- β with spatial effect (AIC=341) and single- β with no spatial effect (AIC=350) we found a worse model fit compared to the multi- β_s with spatial effect (AIC=248).

MSE score was the same for *Biomphalaria* sp. multi- β_s with spatial effect and no spatial effect. For both *Biomphalaria* sp. and *Bulinus* spp., there were cases where the MSE was slightly lower (better value) for particular parameters estimates for a school despite the model having a higher AIC score compared to the other models. For instance, *Biomphalaria* sp. multi- β_s with spatial effect (AIC=258), Chikomwe school MSE score was 0.487 whereas for single- β with spatial effect (AIC=269) Chikomwe school MSE score was 0.475.

From the best fit model, for *Biomphalaria* sp. the following parameter values were found for all schools, $\log \beta$ [-1.42, CI: -4.95, 2.00], for $\log \sigma$ [0.147, CI: -3.34, 3.63] giving $\frac{1}{\sigma} = 0.863$ days [ranging from 0.0265 to 28.2 days] exposed period and for $\log \gamma$ [-2.61, CI: -2.85, -2.36] giving $\frac{1}{\gamma} = 13.6$ days [ranging from 0.0573 to 17.3 days]. For the best fit model, for *Bulinus* spp. the following parameter values were found all schools, $\log \sigma$ was estimated to be -2.50 [CI:-8.49,3.38] giving $\frac{1}{\sigma} = 12.2$ days [ranging from 0.03 to 4870 days] and $\log \gamma$ [-3.39, CI: -8.49, 1.72] giving $\frac{1}{\gamma} = 29.7$ days [ranging from 0.179 to 4866 days] infectious period for all the schools. For *Bulinus* spp. at each school, the following $\log \beta$ parameters were found: Mchoka [-7.26, CI:12.4, -2.14], Samama [-4.80, CI: -11.0,1.35], MOET [-8.26, CI:-13.4,-3.12], Koche [-9.98 ,CI:-15.4,-4.55], St Augustine 2 [-5.73 CI: -11.2,0.289], Ndembo [-4.20, CI:-12.1,3.70], Sungusya [-7.41, CI:-12.6,-2.19], St Martins [-9.26, CI:-14.7,-3.81], Chikomwe [-8.05, CI:-13.3,-2.82], Chipeleka [-6.72, CI:-12.0, -1.50], Makumba [-8.51, CI:-13.8,-3.22], Mtengeza [-6.53, CI:-11.80,-1.24].

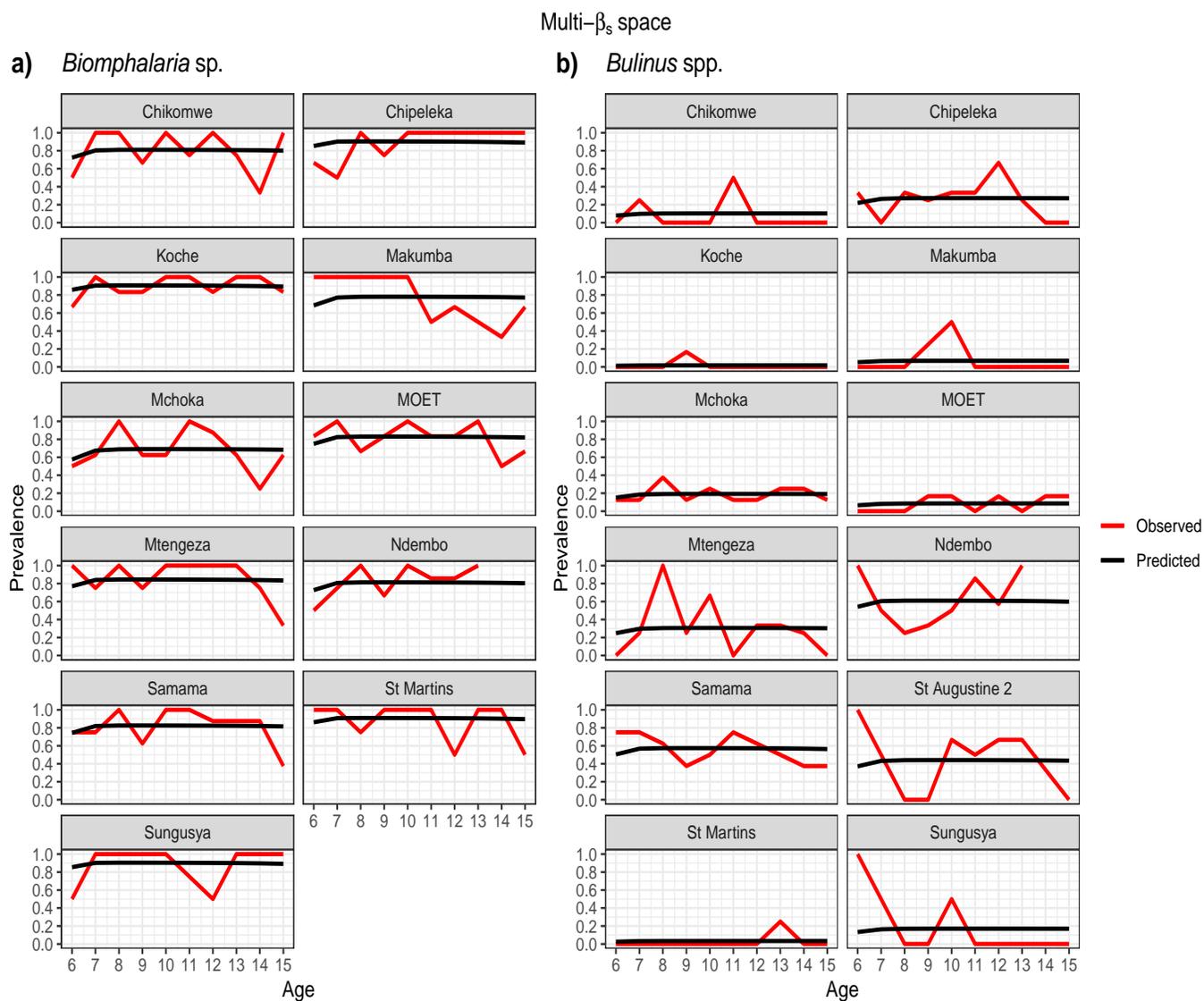


Figure 4.4: Multi- β_s with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

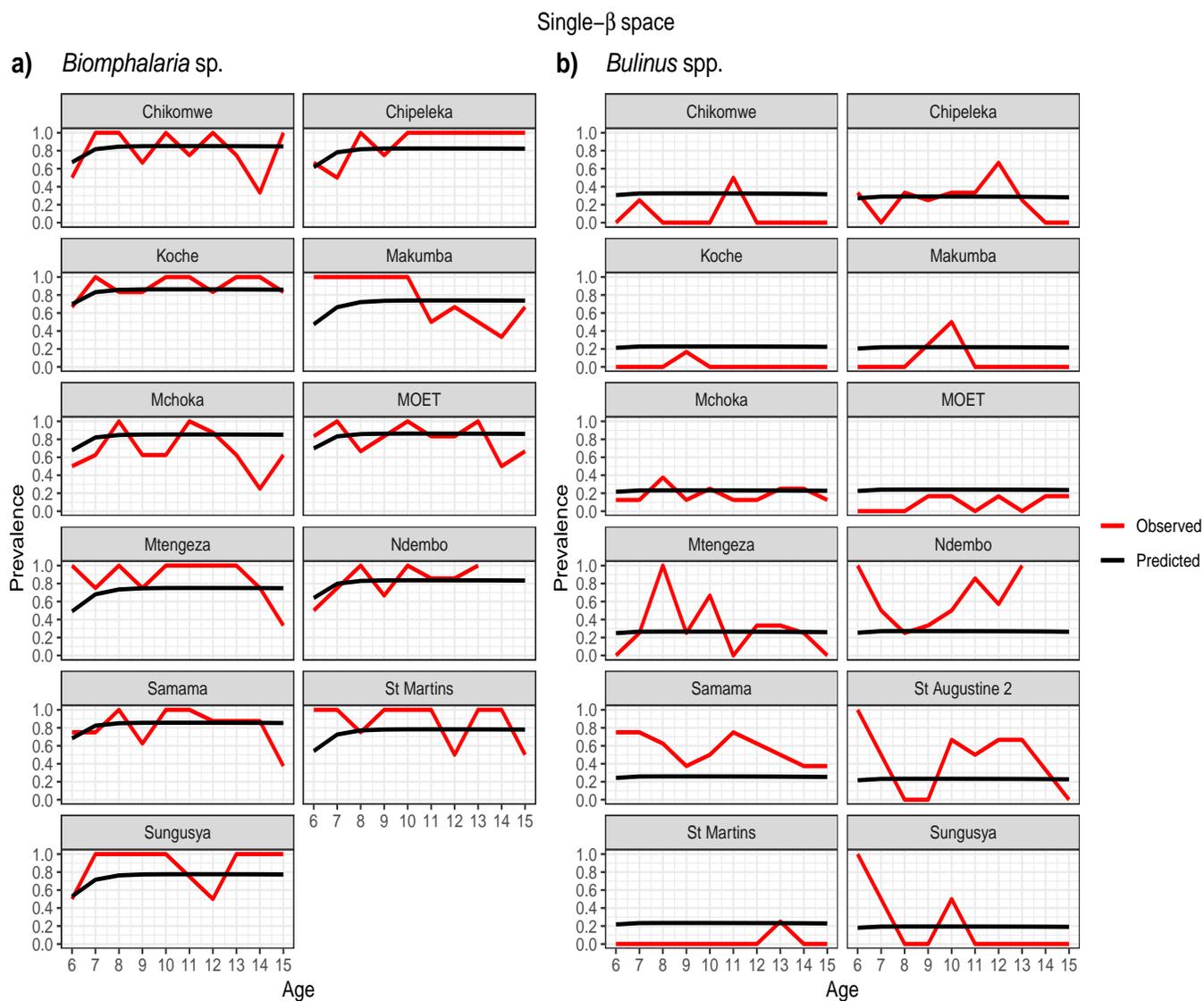


Figure 4.5: Single- β with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

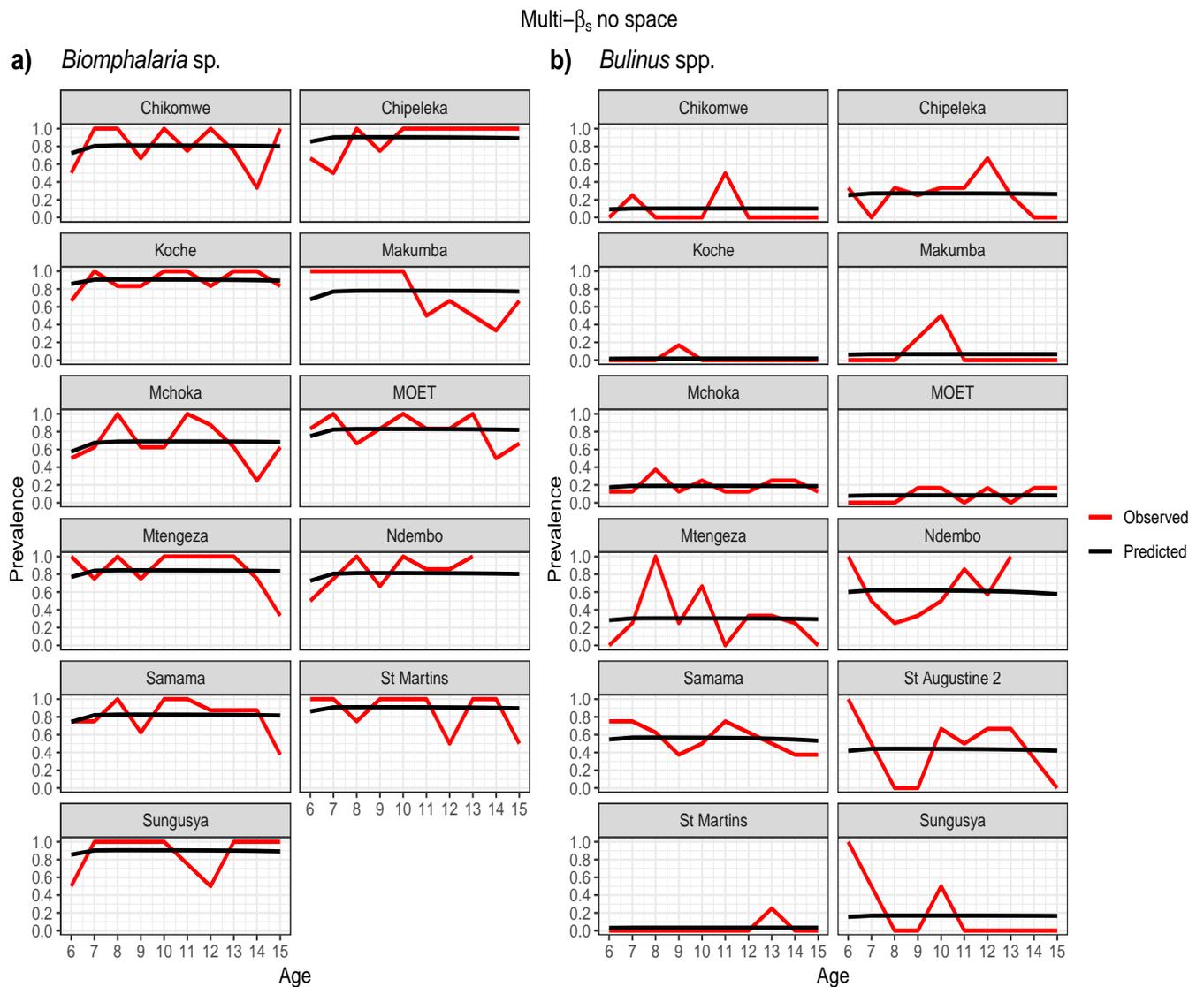


Figure 4.6: Multi- β_s with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black) against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

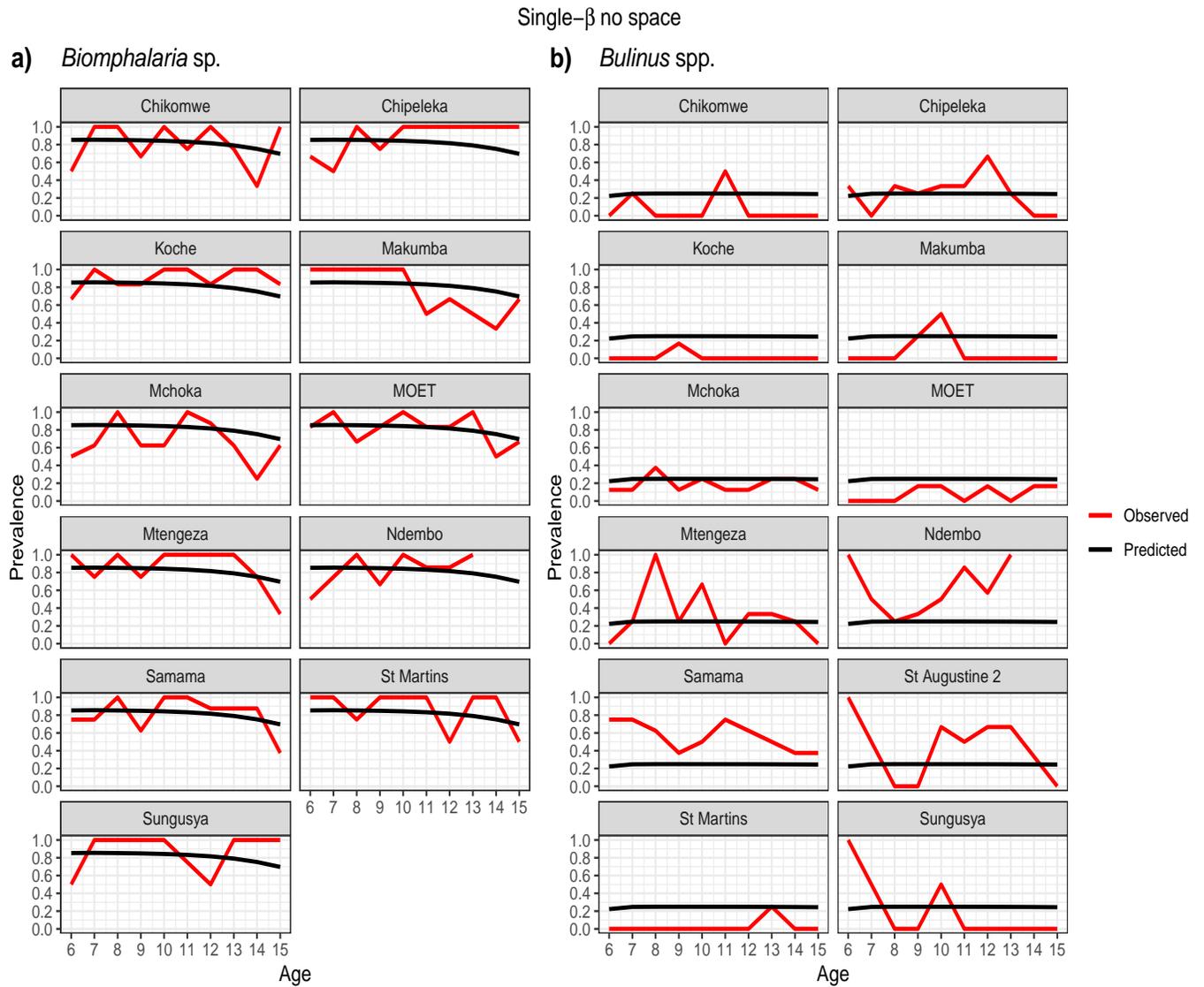


Figure 4.7: Single- β with no space effect model optimisation prevalence prediction, $\hat{\pi}_{as}$ (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

Table 4.2: Parameter estimates for multi- β_s spatial model for each species

Multi- β_s space								
	<i>Biomphalaria</i> AIC=258			MSE	<i>Bulinus</i> AIC=250			MSE
Parameters/Schools [CI]	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-2.78 [-5.31, -0.253]	-2.59 [-9.07, 3.89]	-5.40 [-7.68, -3.13]	0.442	-6.31 [-8.30, -4.31]	-2.80 [-7.93, 2.33]	-4.65 [-6.50, 2.80]	0.0661
Samama	-2.15 [-5.83, 1.54]			0.339	-4.15 [-7.42, -1.62]			0.254
MOET	-4.65 [-8.45, -0.85]			0.252	-7.32 [-9.38, -5.25]			0.0656
Koche	-2.97 [-13.1, 7.15]			0.103	-8.95 [-11.65, -6.26]			0.0248
St Augustine 2				-5.02 [-7.45, -2.59]	1.15			
Ndembo	-0.922 [-4.49, 2.64]			0.186	-4.38 [-7.58, -1.18]			0.657
Sungusya	1.66 [-7.96, 11.3]			0.380	-6.24 [-8.38, -4.09]			1.18
St Martins	1.72 [-9.70, 13.1]			0.420	-8.26 [-11.0, -5.56]			0.0564
Chikomwe	-1.77 [-5.28, 1.74]			0.487	-7.54 [-9.76, -5.33]			0.261
Chipeleka	0.769 [-8.82, 10.4]			0.291	-6.14 [-8.30, -3.96]			0.397
Makumba	0.438 [-2.74, 3.61]	0.671	-7.45 [-9.8, -5.10]	0.254				
Mtengeza	0.919 [-3.52, 5.35]	0.451	5.84 [-8.04, -3.64]	0.870				

Table 4.3: Parameter estimates for single- β spatial model for each species

Single- β , space

Parameters/Schools [CI]	<i>Biomphalaria</i> AIC=269			MSE	<i>Bulinus</i> AIC=341			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-0.514 [-0.526, -0.502]	-4.60 [-5.18, -4.02]	-6.46 [-7.00, -5.92]	0.684	-5.00 [-9.07, -0.927]	0.500 [-12.2, 13.2]	-3.55 [-7.58, 0.489]	0.0856
Samama				0.357				1.17
MOET				0.291				0.307
Koche				0.107				0.459
St Augustine 2								1.49
Ndembo				0.134				1.64
Sungusya				0.469				1.12
St Martins				0.681				0.482
Chikomwe				0.475				0.859
Chipeleka				0.307				0.400
Makumba				0.894				0.455
Mtengeza				0.756				0.909

Table 4.4: Parameter estimates for multi- β_s no spatial model for each species

Multi- β_s no space								
Parameters/Schools [CI]	Biomphalaria AIC=258			MSE	Bulinus AIC=248			MSE
	log β	log σ	log γ		log β	log σ	log γ	
Mchoka	-6.07 [-8.59, 3.56]	-2.59 [-9.08, 3.90]	-5.40 [-7.66, -3.15]	0.442	-7.26 [-12.4, -2.14]	-2.50 [-8.49, 3.38]	-3.39 [-8.49, 1.72]	0.0683
Samama	-5.12 [-8.79, -1.45]			0.339	-4.80 [-11.0, 1.35]			0.215
MOET	-5.07 [-8.86, -1.29]			0.252	-8.26 [-13.4, -3.12]			0.0682
Koche	-3.80 [-13.9, 6.29]			0.103	-9.98 [-15.4, -4.55]			0.0249
St Augustine 2					-5.73 [-11.2, -0.289]			1.08
Ndembo	-5.23 [-8.78, -1.68]			0.186	-4.20 [-12.1, 3.70]			0.623
Sungusya	-3.89 [-13.5, 5.68]			0.380	-7.41 [-12.6, -2.19]			1.13
St Martins	-3.73 [-15.1, 7.66]			0.420	-9.26 [-14.7, -3.81]			0.0568
Chikomwe	-5.26 [-8.75, -1.76]			0.486	-8.05 [-13.3, -2.82]			0.261
Chipeleka	-3.90 [-13.4, 5.63]			0.291	-6.72 [-12.0, -1.50]			0.390
Makumba	-5.5 [-8.66, -2.34]			0.672	-8.51 [-13.8, -3.22]			0.256
Mtengeza	-4.91 [-9.33, -0.489]			0.451	-6.53 [-11.8, -1.24]			0.884

Table 4.5: Parameter estimates for single- β no space model outcome for each species

Single- β_s no space								
Parameters/Schools [CI]	<i>Biomphalaria</i> AIC=250			MSE	<i>Bulinus</i> AIC=351			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-1.48 [-4.95, 2.00]	0.147 [-3.34, 3.63]	-2.61 [-2.85, -2.36]	0.613	-5.50 [-14.8, 3.75]	-4.62 [-11.8, 2.64]	-3.7 [8.67, 1.27]	0.100
Samama				0.275				1.24
MOET				0.190				0.329
Koche				0.233				0.547
St Augustine 2								1.45
Ndembo				0.261				1.78
Sungusya				0.519				1.16
St Martins				0.372				0.541
Chikomwe				0.535				0.542
Chipeleka				0.476				0.390
Makumba				0.506				0.541
Mtengeza				0.326				0.924

Figures 4.4, 4.5, 4.6 and 4.7 show visually that the prevalence prediction model, $\hat{\pi}_{as}$ at steady state was able to capture the age-prevalence structure of the Reed *et al.* 2023 paper [176] (see Chapter 2). Visually there is no evidence to suggest these models do not fit the data for *Biomphalaria* sp. with either single- β or multi- β_s cases and for *Bulinus* spp. only the multi- β_s case. For *Bulinus* spp. the single- β does not visually capture the observed prevalence model for certain schools including Samama, Moet, Koche, Ndembo and St Martins.

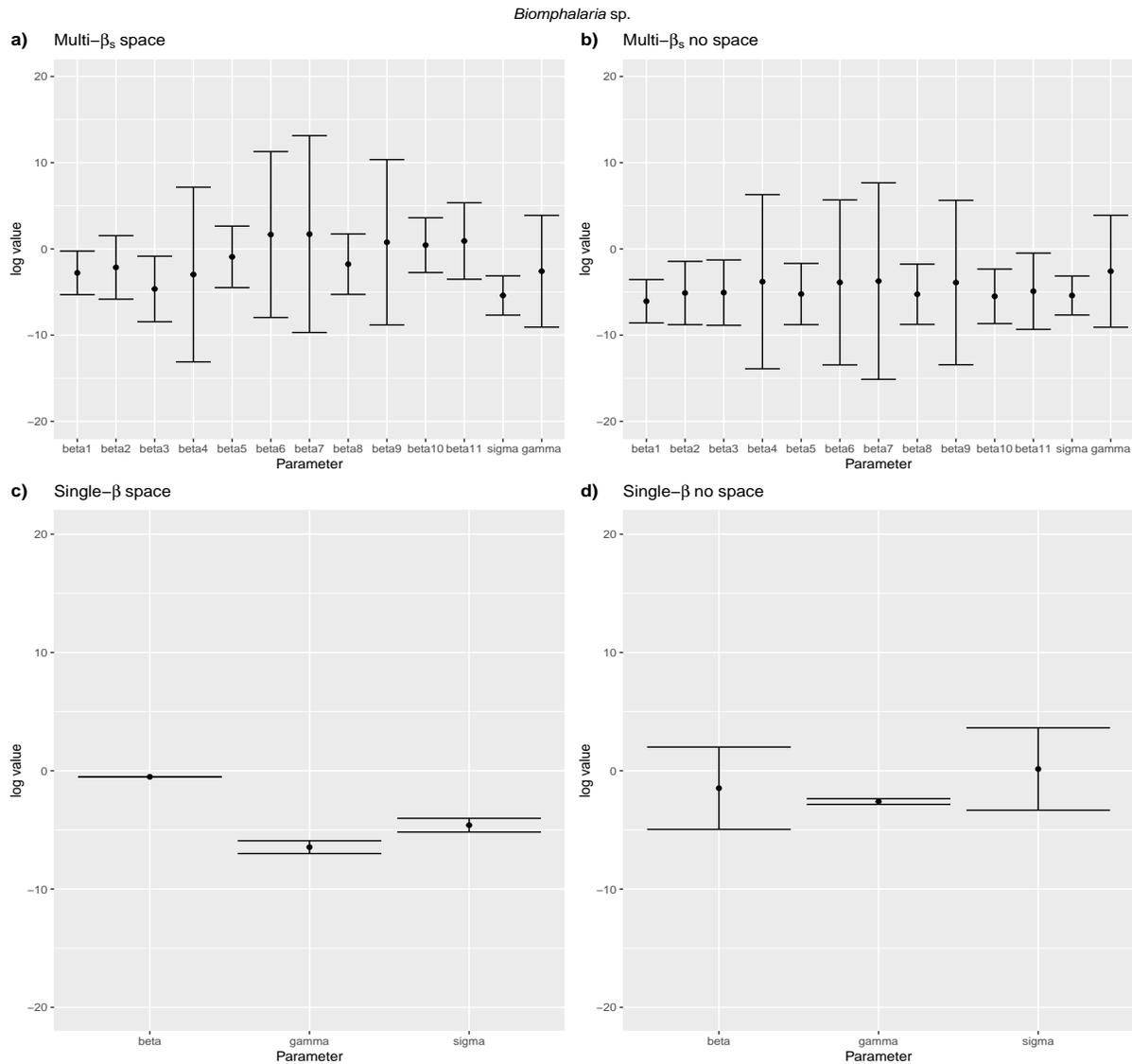


Figure 4.8: Confidence intervals for parameter estimates for *Biomphalaria* sp. models with SAC prevalence at age 6 set as α set to zero a) Multi- β_s space, *Biomphalaria* sp. b) Multi- β_s no space, *Biomphalaria* sp. c) Single- β space d) Single- β no space

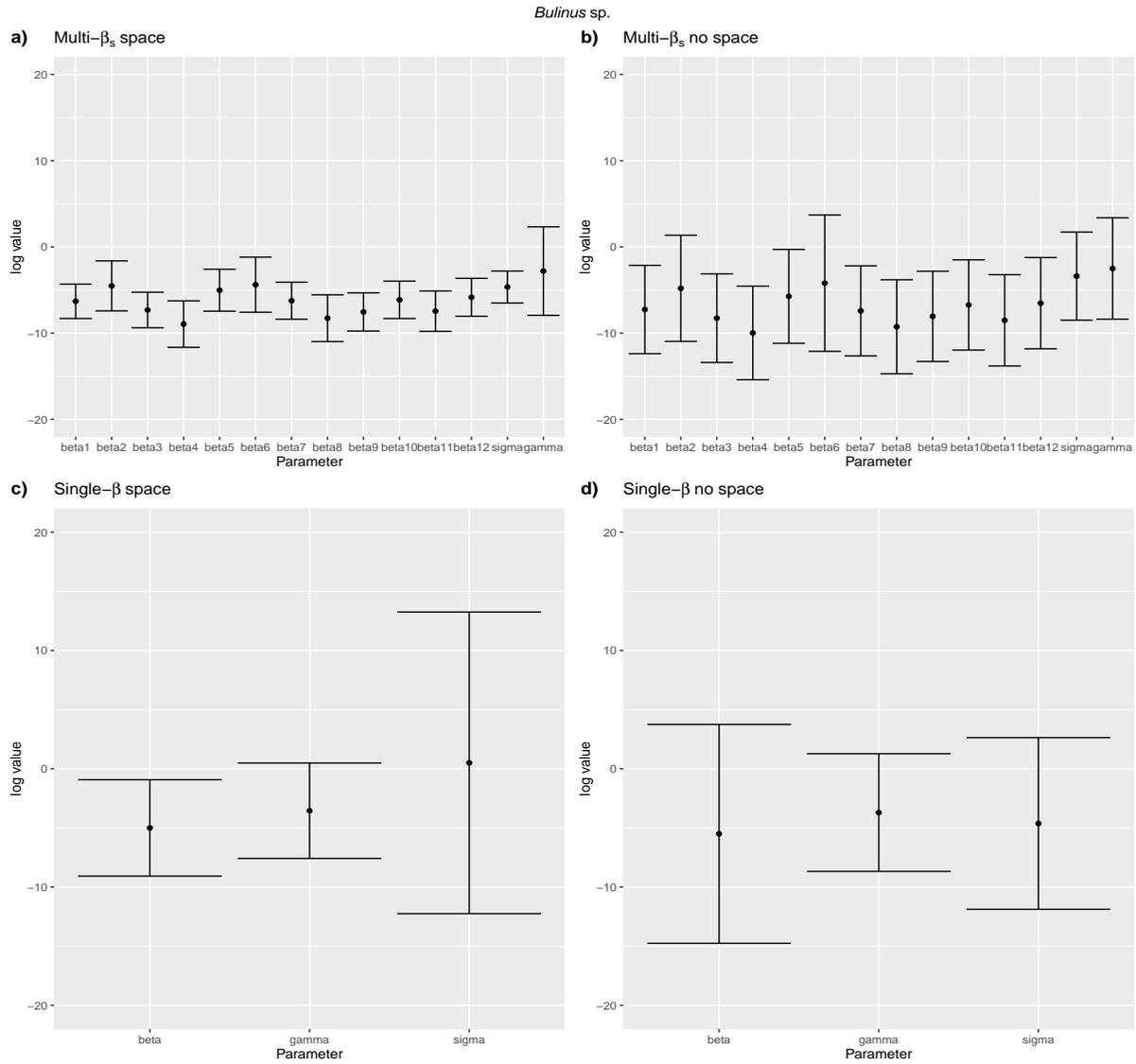


Figure 4.9: Confidence intervals for parameter estimates for *Bulinus spp.* models with SAC prevalence at age 6 set as α set to zero a) Multi- β_s space b) Multi- β_s no space c) Single- β space d) Single- β no space

4.3.3 CI Intervals

We assessed the likely impact of school-level differences in baseline transmission rate by comparing $\log \beta$ values. Considering $\log \hat{\beta}_1$ as a reference, the confidence intervals for all other $\log \hat{\beta}_{2,\dots,12}$ include $\log \hat{\beta}_1$ such that we have no strong evidence that schools 2, \dots , 12 differ from school 1. In other words, a single β_1 for all schools (single- β model) in the model suffices (Figures 4.8 and 4.9). For *Biomphalaria sp.* multi- β_s space and no space models, Figure 4.8a and 4.8b shows visually the approximate baseline β_1 value to be between $\log \beta = -3$ and $\log \beta = -7$ from average of the CIs for all schools. For *Biomphalaria sp.* single- β space and no space models, Figure 4.8c and 4.8d shows visually the approximate baseline β_1 value to be between $\log \beta = -0.5$ and $\log \beta = 1.5$ from average of the CIs for all schools. For *Bulinus spp.* multi- β_s space and no space models, Figure 4.9a and 4.9b shows visually the approximate baseline β_1 value to be between $\log \beta = -6$ and $\log \beta = -7$ from average of the CIs for all schools. For single- β space and no space models, Figure 4.9c and 4.9d shows visually the approximate baseline β_1 value to be between $\log \beta = -5$ and $\log \beta = -5.5$ from average of the CIs for all schools.

4.3.4 Sensitivity analysis

When testing different values of SAC age 6 prevalence, α we found that the parameter estimates did not change a lot and the AIC and MSE scores were similar, suggesting a similar fit. In addition, for $\alpha = 0.10$ and $\alpha = 0.20$ for *Bulinus spp.* we found that there were non-identifiability issues between γ and β as shown in Figure 4.10, where essentially the optimiser is just contouring around the distribution, following the contours, instead of finding the estimated parameter value(s).

In addition the plots of the fits to the age-prevalence profiles for the different α values are shown in Appendix D.1 D.1.1, Figures D.1, D.2, D.3 and D.4, D.5, D.6, D.7 and D.8, D.9, D.10, D.11 and D.12.

The tables of the parameter estimates for the different α values are shown in Appendix D.1 D.1.2, Tables D.13, D.14, D.15 and D.16, D.17, D.18, D.19 and D.20, D.21, D.22, D.23 and D.24. The confidence intervals are shown in appendix D.1, D.1.3, Figures D.25, D.26, D.27, D.28, D.29, D.30.

The confidence intervals for the parameter estimates for the different α values are shown in Appendix D.1 D.1.3, Figures D.25, D.26, D.27, D.28, D.29 and D.30.

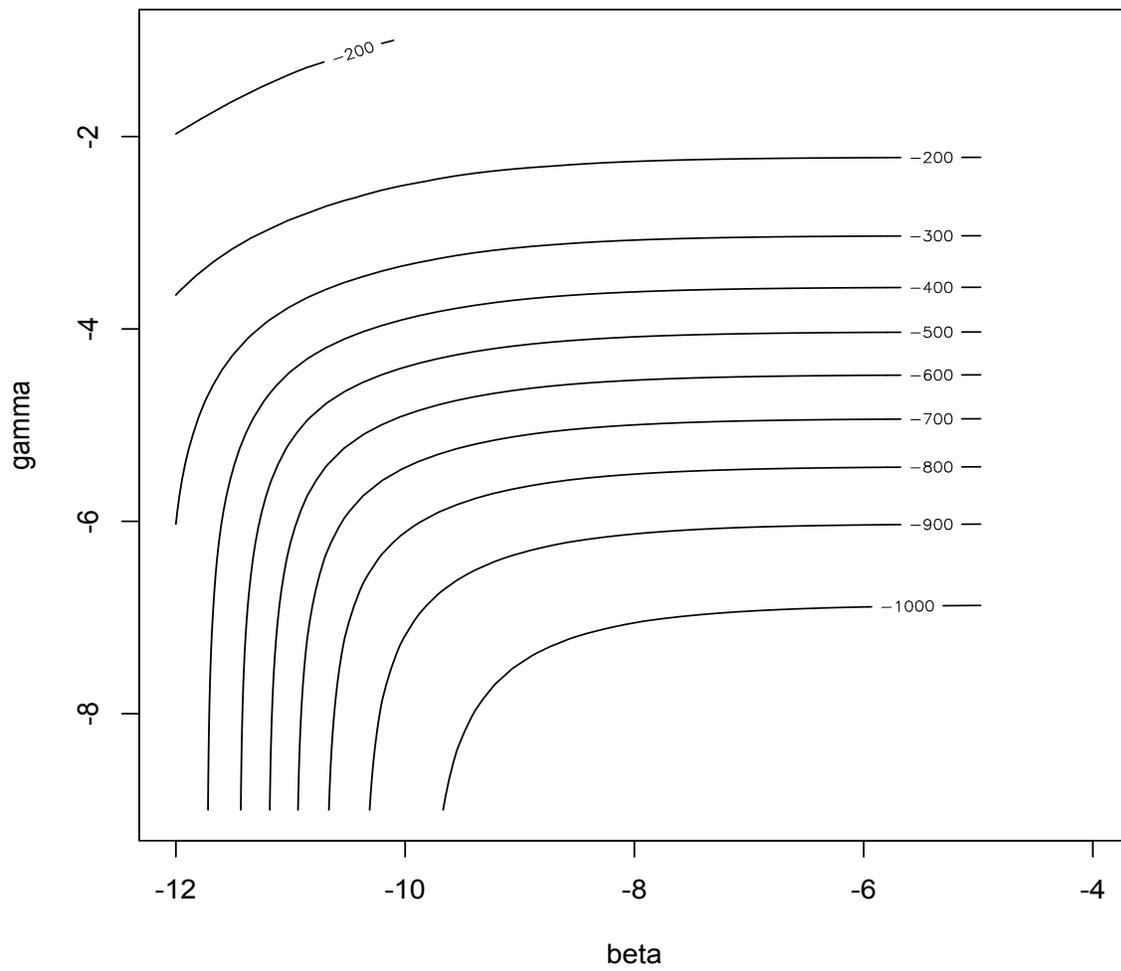


Figure 4.10: Contour plot showing non-identifiability issue between γ and β parameters

4.4 Discussion

Our analysis has made attempts to model the dynamics of *Schistosoma* transmission at a single point in time, and estimate the best possible values of our parameters given our data. This allowed us to determine their associated *Schistosoma* infection risk and to determine whether school distance from the lake shoreline determines the exposure of SAC *Schistosoma* risk, and further, whether we can reproduce the SAC prevalence age-profiles shown in Reed *et al.* 2023 [176] (see Chapter 2). This was conducted using a novel SEIRS ODE model and carried out using the L-BFGS-B optimisation algorithm. Our model is novel in the sense that we focus on capturing the disease process within children as a function of exposure to a quasi-static snail population, as opposed to modelling the entire *Schistosoma* lifecycle as is common in other studies [114, 115]. Our study found the best fitting model for *Biomphalaria* sp. was the single- β no spatial effect model (AIC=250), which supports the use of single- β model for all schools and with no spatial decay of the force of infection from snails with distance from the lake shoreline. Hence, we cannot conclude that transmission differs appreciably between schools or that school distance from the shoreline has an effect on SAC exposure. For *Bulinus* spp., however, the best model was the multi- β_s with no spatial effect model, which supports a hypothesis that school-level factors determine the apparent prevalence of urinary schistosomiasis though we have no evidence that the spatial relationship to snail exposure affects prevalence. Therefore, this could mean all the schools are close enough to the lake for the variation in the distance to not have an effect on SAC exposure.

There was possible heterogeneity in transmission between schools. Based on our AIC scores the model preferred incorporating different transmission rates for each school and this was also shown in Figures 4.4 and 4.6 which suggests this is likely the case. In particular, one noted result was Samama, Moet, Koche, Ndembo and St Martins predictive estimate for single transmission rate for all schools with *Bulinus* spp. did not capture the data that we observed and so this suggested that we needed different transmission rates in the schools. However, we found no statistical evidence that transmission rates are different from each other based on CIs given our data. This may well be due to the noisy nature of the observed prevalences, and associated wide confidence intervals, as well as the possibility that our approximate confidence intervals were not capturing the true correlation structure in the joint likelihood surface. Since, Figures 4.4, 4.5, 4.6 and 4.7 indicate that the multi- β models indicate a better predictive fit, it is perhaps not surprising that these are preferred by AIC. Our conclusion, therefore, is that the multi- β model should be preferred on the grounds of improved predictive performance, rather than model parsimony.

From the best fit models, the exposed period (S to E group) was estimated to be 0.863 days for *Biomphalaria* sp. and 12.2 days *Bulinus* spp. compared to 14–84 days [usually between 35–

42 days] usually reported in other studies [111, 190]. In addition, the infectious period (E to I group) was estimated to be 13.6 days and 0.357 days, whereas in other studies the infectious period (γ) (I to R group) has been found to be estimated around 35 days (14–70 days); we found lower rate in the infectious period than previous studies [53]. These differences in exposed and infectious period could have been due to some identifiability issues (i.e. Figure 4.10) that were noted for our unknown parameters β , σ , γ making it difficult for the model to capture the observed prevalence in our dataset. Further, the shorter exposed and infectious predictions could suggest our model fit is artefact of data collection or even the ODE fitted to a small population (in which there is a continuous approximation to a discrete state space of numbers of individuals). A SIR model could have been the more parsimonious model to fit, albeit with the obvious violation of a biological principle (the fact that your infection is not immediately detectable when you first get infected).

4.4.1 Limitations

A limitation of this study is that we only accounted for distance from the lake shoreline (data aligned to the lake shoreline) and did not account for inland bodies of water, for instance, ponds and pools of water inland in the dataset. Further we do not know in detail the local environment from the inland schools, only the local environment and their relative exposure along the lake shoreline. If there is any effect of differential exposure to snails by SAC at the different schools, this may not be the best representation using the simple distance from the lake. This could be due to individual human spatial movement having predictable patterns influenced by their environment and socio-economic factors. For instance, SAC schools that have direct and easy access to the lakeshore could be more likely to be exposed to snails as they visit the lake shoreline more often than rural villages [191]. This could have also been due to all our schools in our study being close enough to the lake that they guarantee SAC visiting the shoreline often. Although, water contact can also be dictated by gender or age [177]. Our study was place specific so there may be a different result at other locations with different snail species or seasonality. Moreover, the study could have been expanded to other schools in the area.

We also consider SAC to only travel using Euclidean distance (straight) from their school to the lake shoreline. This is a misrepresentation as often SAC will go to different parts of the shoreline, with more easily accessible areas than others so repeated exposure is more likely at these focal locations. Further, we did not know where the residences of the SAC were, only their school location was considered in our models. Future studies, using a quantitative social study are required to improve our understanding of how exposure to snails along the shoreline is affected with dis-

tance from SAC school or residential living area [180]. For instance, other socio-economic factors including education, wealth, trade, or health could affect how often, where, and when SAC are exposed to snail habitats, and hence effect their *Schistosoma* transmission risk [192].

We only had cross-sectional data for a single point in time (aggregate 2017 to 2019 data) due to having too small samples for each year. Ideally, in future studies, a longitudinal cohort study should be carried out to attain more data, where we recruit a panel of children and then follow them every year. Firstly, this would allow us to better understand the sensitivity and specificity of the tests used as they are repeated. Secondly, this would allow us to study how the dynamic of infection changes over time; for instance, allow us to understand what is driving the age profile over time instead of single-point in time.

Another limitation of the study is that we fixed some parameters for identifiability purposes (i.e. Figure 4.10). For instance, we assumed immunity to be under one day for age 6 SAC due to the optimiser difficulty in identifying omega against the other parameters, however estimating omega would be useful if we had more data on the children's partial immunity to be able to find their actual immunity accumulated over time. Kura *et al.* 2021 [178], reported that in the presence of acquired immunity MDA programmes were less effective in decreasing the prevalence of infection compared to no acquired immunity [178]. Hence, the ability to estimate immunity within SAC could improve application of appropriate control methods. Further, we could have also accounted for treatment from MDA, which can affect the children's infection status and load of infection (light/heavy) which was not considered in this study.

To improve how our model captures the dynamics of *Schistosoma* transmission and try find significant statistical evidence, future work could be carried out using a bayesian approach to look at the structure of the posterior very carefully and work out whether its the non-identifiability of parameters in the model that mean we can't trust the Gaussian approximation that we need for calculating the CI in the way we designed our models.

Furthermore, a limitation of the model was that we assumed the all-snail infection drops off by 100km, although this is a sensible assumption as it is unlikely SAC would travel more than 100km from the shoreline: this is a weakness of the model given our data. Another limitation of our study is we assumed that all snails are infected, so presence of snail indicates exposure to *Schistosoma* infection. This may not be the case in reality. However, in spite of these several limitations, our model provides insight into how proximity to snail abundance might affect the patterns of age prevalence in schools nearby to the shoreline.

4.5 Conclusion

Our study reproduces the non-linear age-prevalence profiles by modelling the *Schistosoma* infection transmission dynamics with age-related immunity using cross-sectional parasitology and malacological data already collected and analysed. We find the best estimates values of the main determinants of *Schistosoma* risk with SAC at schools along the southern part of Lake Malawi. One of the main outcomes of our study is that school measured distance from the shoreline has no tangible effect on the SAC exposure. However, our dataset was limited by our assumptions of the model, size and single-time-point. We were able to a certain extent to reconstruct the age-infection profile using a simple state transmission model. However, further studies could be carried out using a longitudinal cohort study, which would be much more powerful than a single cross-sectional study to understand *Schistosoma* infection dynamics and how this affects SAC age-infection profiles and the implications on how control methods are applied.

Chapter 5

General Discussion

5.1 Chapter overviews

The General Discussion chapter brings together and summarises all the chapters in my thesis. Key findings are used to draw conclusions and propose informed modifications to, as well as support for, parts of the WHO guidelines for the control and elimination of schistosomiasis [63]. By application of statistical and modelling, my thesis explored the spatial and temporal transmission of schistosomiasis within SAC along the southern part of the lake shoreline in Mangochi district, Malawi. The thesis contributed insights regarding the drivers of *Schistosoma* infection, determinants of transmission, understanding SAC age prevalence profiles and snail distribution along the shoreline. This, in turn, helps to improve understanding on how to interrupt transmission and improve the application of control methods. The investigation of transmission of *Bulinus* spp. snails (UGS) and the unexpected presence of *Biomphalaria* sp. snails (IS) along the southern shoreline had already been started prior to my PhD research through parasitology surveys and malacological studies. This thesis has analysed this data added a quantitative dimension to the overall research programme. The situation in Lake Malawi acts as a good example of the changing dynamics, albeit natural or influenced by anthropogenic factors, that many disease control programmes face and must overcome. To help bring forward a synthesis I provide an oversight of the findings of each chapter and attempt to draw together a coherent output rationale.

In Chapter 2, I analysed the parasitological data collected on the SAC attending schools in the Mangochi District on the southern part of Lake Malawi to gain novel insight on age-stratified infection prevalence during a newly emerging IS infection focus alongside the existing transmission of UGS. Using GAMs, we were able to find evidence that SAC peak of IS and co-infection of IS and UGS prevalence was around 11 years old given our data. No clear age-infection pattern was

found for UGS. Baseline age-prevalence heterogeneity between schools was noted and suggested a need for further investigation into determinants and dynamics of transmission. To our knowledge, Chapter 2 is novel, as no other studies have analysed age-prevalence profiles of IS within SAC during a newly emerging focus of infection alongside existing background transmission of UGS. Further, we were able to identify the age prevalence profiles for IS, UGS and co-infection along the southern shoreline of Lake Malawi. My analysis was published in *Parasite Epidemiology and Control* and hopes to signpost others to the importance and power of inference that applied more detailed analytical techniques can bring.

Next, in Chapter 3, I carried out a geospatial analysis using the malacological data collected on the *Schistosoma* intermediate host snail spp. along the southern part of Lake Malawi to find the distribution of snails along the shoreline. A Bayesian Poisson latent Gaussian process model was fitted to the observed abundance data of *Bulinus* spp. and *Biomphalaria* sp. intermediate host snails of *Schistosoma* infection. This method allows us to predict snail abundance at locations along the lakes shoreline by interpolating between our sampling sites, allowing us to gain insight into the spatial scale at which snail populations vary. Substantive heterogeneities in snail distributions along the shoreline were found. Further investigation was required to improve the identification of likely snail habitats posing a high risk for schistosomiasis transmission. I highlighted that interpolation of snail survey data was possible and helped to define an appropriate scale at which it was informative.

Chapter 3 is the first study carried out in Lake Malawi, to identify possible areas with increased risk of *Schistosoma* infection along the shoreline, where SAC are most likely to be exposed to schistosomiasis using a remote sensor data and observed snail counts. Further, the method used in this chapter, to my knowledge is novel, where we take the 2D shoreline and convert this into 1D shoreline by interpolating the snail abundance, through smoothing the observed estimates and remote sensor data for a certain number of equally spaced predictions points. In turn, my analysis was published in *Parasites and Vectors* and allowed us to measure the intermediate snails' habitats "hotspots" along the shoreline, based on the snail abundance predictions and their associated environmental conditions that increase or decrease snail abundance.

Next, Chapter 4 developed a dynamical SEIRS model, the first of its kind integrating new compartmentalised variables, using the Chapter 2 and Chapter 3 outcomes as a cross-sectional single-point time study to find the determinants of infection and improve understanding of age-infection dynamics. Further, it sought to understand whether school distance from the shoreline affects *Schistosoma* transmission, in turn, improving the application of control methods within SAC along the southern shoreline of Lake Malawi. We developed a SEIRS ODE model with age-related immunity to model the transmission dynamics of *Schistosoma* infection. We fitted the SEIRS model over prevalence

data via maximum likelihood estimation, and compared four different models using AIC and MSE to assess the fit. Given our data, we found no effect on SAC exposure to *Schistosoma* infection risk based on the school's distance from the shoreline. Heterogeneity in transmission between schools was apparent, though differences in transmission rate between schools were not significant. However, we were able to reproduce the age-prevalence profiles found in Chapter 2, with the insight that age-related immunity is responsible for the pattern of prevalence seen even without explicitly accounting for age-related water exposure. Further we estimated a lower duration of the infectious period than previous studies [53]. The main limitation of the study was that the data was single-point in time, and longitudinal panel studies would be required to gain further evidence of causality between snail exposure and schistosomiasis incidence in our study locations. Further, lack of data could have led to shorter estimated exposed and infectious periods in study.

Chapter 4's modelling approach is similar to other studies, as we want to improve control programme implication and we use a compartment model to set up our model. However our study is different to most studies as in our study we did not measure the dynamics of the pathogen lifecycle as a function of interaction between the human and snail populations, as is the case in many other studies [114] [115]. We did not have data on whether the snails were infected or not. Instead, we computed a hazard rate given a SAC's susceptibility to *Schistosoma* infection and their school's distance from the shoreline. Further, Kanyi *et al.* 2021 [111], Lopez *et al.* 2024 [120] and Tabo *et al.* 2023 [193] studies were performed to establish the endemic equilibrium points and their threshold value to find the existence of the equilibrium points and their stabilities (e.g. bifurcation). This was different to this thesis as we used a dynamical model which includes knowledge of the disease process to replicate the age prevalence pattern that we see in our empirical data instead of the whole population. Furthermore, a dynamical model focusing only on the endemic (output) was useful as we wanted to study the effects of immunity on reproducing the age profiles that we see. Further other studies usually find the endemic equilibrium through setting the derivatives to zero and then use known parameters values from literature [78, 117, 193]. An adaption of this is Graham *et al.* 2021 [114] who produced a stochastic individual based model, also known as SCHISTOX using known parameters values from literature to model the dynamics and also adding age contact rates. The ability to study age and intensity of infection or the whether infection is acute or chronic has been investigated [114, 118], whereas we study age prevalence profiles. All these other studies are different to our thesis as we do not know the parameters values for our model so we identify the output of the steady state to estimate our parameter values, whereas other studies investigate the dynamics over time with known literature values.

In addition, we are different to other studies by using the concept that the children are a small fraction of the total population. Therefore, we were not considering all forms of transmission

occurring. As there is no reason given this epidemic assumption, that the infection should come and go, we treat it as a constant reservoir of infection. If we were modelling the whole population then you mostly likely would not be able to do this; but as we have other adults and children who would be causing infection who we do not consider in our model, we consider them a constant reservoir of infection. Furthermore, our model allowed us to explore the SAC's schools distance based on proximity to the shoreline from associated known risk factors for *Schistosoma* infection transmission instead of just looking at risk factors themselves. Our model had an age-structured format with age-related immunity as a novel implementation, where we explore immunity with respect to age in our SEIRS model.

5.2 Conclusion of this thesis

Upon broader appraisal of all chapters, my thesis found the peak age-prevalence profiles given our data to be around 11 years for IS and co-infection focus. Throughout the literature there are varied age-prevalence profiles reported with them mostly between early- to mid-adolescence [10–15 years] [17, 65, 76, 132, 133]. Furthermore we created a way to reconstruct these age prevalence curves using a simple SEIRS model, which shows the dynamics of SAC age structure with age-related immunity considered. However, we found an earlier peak in prevalence profile than other studies which is discussed in Chapter 2 and 4. We think this might be a result of the newly established transmission potential of IS (presence of *Biomphalaria* sp. snails along the shoreline) locally alongside growing acquired immunity in exposed children. However, it was also noted in our thesis that the age prevalence profiles were extremely noisy given our single-point in time data, with small sample size data which could have been the reason for the difference in outcomes to other studies. Further the thesis focused on the trace positive results instead of trace negative results, which allows for low intensity infections to be considered, but can lead to overestimation of the prevalence of infections [63, 129]. Furthermore a limitation of my thesis was intensity of infection was not considered due to lack of KK data as stool samples were only taken for CCA-positive SAC.

The thesis gained insight into the “peak shift” phrase association with infection epidemiology of schistosomiasis, which suggests that transmission can oscillate over time and can vary between schistosome species despite the same infection pathway [123]. The thesis modelled age structure with age-related immunity, allowing us to model that SAC start off with hardly any immunity, and it increases over time, essentially building an assumption of partial immunity. This allowed us to estimate the other important parameters for *Schistosoma* infection in the model with consideration of age-related immunity. We noted that Kura *et al.* 2021 [178], reported that the presence of

those with acquired immunity in MDA treatment programmes decreased the effectiveness of treatment compared to no acquired immunity individuals [178]. Hence, the thesis stresses the need for improved understanding of age-related immunity to aid control programmes' effectiveness. The thesis brought attention to changes in the peak of prevalence over time; this is also thought to be due to many site-specific factors, including environmental, socio-economic, demographics and water exposure. In particular our thesis emphasised the importance of understanding the microhabitat for schistosome specific species to find the associated snail abundance of focal areas [168, 175]. Our study only revealed the transmission dynamics along the shoreline and opens up future work into applying our analysis to other areas where snail habitats are found such as pools adjacent to the lake or rivers, ponds and streams.

My thesis improved the identification of potential “hotspots” of snail activity and areas most at risk to *Schistosoma* infection. Our study could lead to improved, more localised snail control. Supporting WHO guidelines recommendation of more focal snail control with molluscicides. Furthermore, our study improved our capability to measure at high-resolution the habitats of the intermediate host snails which live below the surface of the water. We found remote sensing seems a promising tool with which to measure intermediate schistosome snail species habitats, however, it was not the best approach to finding the fine-scale spatial microhabitat differences and their associated snail abundance due to its lack of resolution. The thesis promotes the need to take regular equally spaced environment parameters measurements all along the shoreline; perhaps using towed arrays of sensors behind a boat sailing close to the water's edge to gain a more realistic understanding of environmental conditions over time and space.

In turn, this would improve the understanding of the microhabitat of schistosome species intermediate snails and the potential hotspots of snail activity. This provides information on which areas of the shoreline to steer people away from to reduce transmission, again interrupting transmission. In turn, supporting WHO guidelines recommendation on behaviour change interventions required to reduce transmission of *Schistosoma* spp. in endemic areas [63]. The thesis draws attention to the complexity of the snails microhabitats with many more physical, chemical and environmental factors (abiotic and biotic) which could be impacting *Schistosoma* snails' habitat and their relative snail abundance, not considered in our model, due to time constraints or non-accessible data [163, 164, 174]. Further, other factors such as food source, pollution, parasitism and even the competition for snail habitat with other organisms within an area were not considered in our model [174]. Overall, the findings of the thesis promote eliminating schistosome transmission by snail control alongside MDA programmes [63].

Furthermore, it is known from the literature that age-infection profiles are dependent on transmission rates and focality [123]. The thesis explored optimisation of parameters for each stage of

Schistosoma infection transmission through our SEIRS model with age-related immunity which allowed us to draw out heterogeneities in transmission between schools. This allowed us to identify the SAC hazard rate given that they are susceptible based on their school's location. In addition, we also found SAC's school distance from the shoreline did not have an effect on SAC exposure given our data. We concluded that exposure is not as simple as distance from the lake shoreline and involves a more complicated approach to measure exposure. This could be due to individual human spatial movement more likely to have predictable patterns, where they are more likely to visit more accessible parts of the shoreline, although not always the case. Further, this emphasises the need for future studies, using a quantitative social study to look at the environmental and socio-economic factors which could aid in reducing transmission. These factors are known to affect human behaviour and their related exposure to the shoreline, for example, education, wealth, trade, or health could affect how often, where and when SAC are exposed to *Schistosoma* infection through the intermediate host snails, which would affect their *Schistosoma* infection risk [192]. In turn, the thesis highlights the difficulty of establishing human movement patterns to locate where acquired infections occur, with future need for global positioning system (GPS) monitors to help aid this.

My thesis highlights the forthcoming challenge to refine geospatial sampling frameworks with future opportunities to map schistosomiasis within actual or predicted snail distributions which might better reveal the environmental transmission possibilities. In addition, the thesis illustrates the complexity of the *Schistosoma* infection transmission dynamics and allows us insight into how we might interpret transmission, to meet WHO new guidelines to target elimination of schistosomiasis and aid appropriate application of control methods [63]. As mentioned previously, this shows the ability to identify locations where freshwater snails are most abundant and can aid targeted control methods, preventing re-infection and hence, helping eliminate or reduce transmission [67, 148, 149]. Further, the thesis shows demand to build on the models created in this thesis particularly the snail abundance and age prevalence modelling, using them to create rather more careful experimental designs that are targeted towards a spatial understanding of snail abundance along the lake shoreline. Ideally, we would use a rather more nuanced longitudinal understanding of how prevalence changes with respect to age, whereas only a cross-sectional study was considered in our study with a lack of repeated measurements of the same SAC over time.

The thesis's main limitation is the lack of data availability, which was constricted by the Covid-19 pandemic, funding, and time. The gifted data had small numbers of children sampled at each school, not the same number of children collected at each school and not the same school-children were collected. The data also had a small sample size for number of snails collected for each year with a different number of locations sampled in year. We used cross-sectional single-point in time

data where every time you go back to the school you end up testing a different cohort of children. Ideally, we would have the same panel of children, following up each year, testing them each year. Furthermore, we would want to know their immune status, where they visited the shoreline and what activities they do as they grow up. Therefore, we could get those repeat measurements and control in-between child variation which was observed in our study. This strengthens the demand to carry out a projected cohort study where we follow these children as they grow up. As from the model point of view, this could give the models used in this study the best chance to really identify some of these parameters which are otherwise quite difficult to estimate. Longitudinal cohort studies are required to attain more data; recruiting a panel of children and following them up each year. This will allow us to improve our understanding of age-prevalence curves and how *Schistosoma* transmission dynamics change over time and space. Further, this would allow us to better understand the sensitivity and specificity of tests used in our study which could improve our ability to correctly identify whether the SAC were infected with *Schistosoma* infection at any point in time and improve our understanding of age prevalence profiles over time.

In conclusion, my thesis used multiple statistical approaches to better understand the schistosomiasis transmission in the Mangochi district on the south shoreline of Lake Malawi. In this thesis the heterogeneities in transmission between schools, the focal spatial distribution of the intermediate host snail spp., environmental conditions and the socio-economic factors shows the complexity of transmission dynamics. This work further emphasises the need for further longitudinal cohort studies, environmental parameters collected at a finer spatial scale and socio-economic factors to be considered when studying transmission dynamics.

Bibliography

- [1] WHO. *Schistosomiasis (Bilharzia)*. 2023. URL: https://www.who.int/health-topics/schistosomiasis{\#}tab=tab{_}1 (visited on 07/21/2023).
- [2] Manuela Ciddio *et al.* “The temporal patterns of disease severity and prevalence in schistosomiasis”. eng. In: *Chaos* 25.3 (2015), p. 036405. ISSN: 10541500. DOI: 10.1063/1.4908202.
- [3] Makaula P. *et al.* “Schistosomiasis in Malawi: A systematic review”. In: *Parasites and Vectors* 7.1 (2014), no pagination. URL: <http://www.parasitesandvectors.com/content/pdf/s13071-014-0570-y.pdf{\%}5Cnhttp://www.parasitesandvectors.com/{\%}5Cnhttp://ovidsp.ovid.com/ovidweb.cgi?T=JS{\&}PAGE=reference{\&}D=emed18b{\&}NEWS=N{\&}AN=612350859>.
- [4] Mohamed Badawy Abdel-Naser *et al.* “Schistosomiasis (bilharziasis) and male infertility”. In: *Andrologia* 51.1 (2019), e13165. ISSN: 03034569. DOI: 10.1111/and.13165. URL: <http://doi.wiley.com/10.1111/and.13165>.
- [5] Donald P McManus *et al.* “Schistosomiasis”. In: (2018), pp. 1–19. DOI: 10.1038/s41572-018-0013-8.
- [6] Atupele P. Kapito-Tembo *et al.* “Prevalence distribution and risk factors for *Schistosoma hematobium* infection among school children in Blantyre, Malawi”. In: *PLoS Neglected Tropical Diseases* 3.1 (2009). DOI: 10.1371/journal.pntd.0000361.
- [7] WHO. *Schistosomiasis*. 2019. DOI: <https://www.who.int/schistosomiasis/disease/en/>. URL: <https://www.who.int/schistosomiasis/disease/en/> (visited on 11/19/2019).

- [8] David S Brown. *Freshwater Snails Of Africa And Their Medical Importance*. Second Edi. CRC Press, 1994. URL: <https://www.routledge.com/Freshwater-Snails-Of-Africa-And-Their-Medical-Importance/Brown/p/book/9780748400263>.
- [9] Horrall S. Lackey, EK. “Schistosomiasis (*Schistosoma haematobium*)”. In: *StatPearls* (2020). URL: <https://www.ncbi.nlm.nih.gov/books/NBK554434/>.
- [10] WHO. *Current estimated total number of individuals with morbidity and mortality due to Schistosomiasis haematobium and S. mansoni infection in Sub-Saharan Africa*. 2020. URL: <https://www.who.int/schistosomiasis/epidemiology/table/en/>.
- [11] Mohammad H. Alharbi *et al.* “*Biomphalaria pfeifferi* snails and intestinal schistosomiasis, Lake Malawi, Africa, 2017-2018”. In: *Emerging Infectious Diseases* 25.3 (2019), pp. 613–615. ISSN: 10806059. DOI: 10.3201/eid2503.181601. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6390747/>.
- [12] Kosala GAD Weerakoon *et al.* “Advances in the Diagnosis of Human Schistosomiasis”. In: (2015). DOI: 10.1128/CMR.00137-14. URL: <http://cmr.asm.org/>.
- [13] L. Bonnie Webster *et al.* “Praziquantel treatment of school children from single and mixed infection foci of intestinal and urogenital schistosomiasis along the Senegal River Basin: Monitoring treatment success and re-infection patterns”. In: *Acta Tropica* 128.2 (2013), pp. 292–302. ISSN: 0001706X. DOI: 10.1016/j.actatropica.2012.09.010.
- [14] WHO. *Schistosomiasis*. 2020. URL: <https://www.who.int/en/news-room/fact-sheets/detail/schistosomiasis>.
- [15] CDC. *Parasites-Schistosomiasis*. 2019. URL: <https://www.cdc.gov/parasites/schistosomiasis/biology.html> (visited on 04/27/2020).
- [16] Ralph Muller, Derek. Wakelin, and Ralph Muller. *Worms and human disease*. CABI, 2002, p. 300. ISBN: 9780851995168.
- [17] Daniel G. Colley *et al.* “Human schistosomiasis”. In: *The Lancet* 383.9936 (2014), pp. 2253–2264. ISSN: 01406736. DOI: 10.1016/S0140-6736(13)61949-2. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0140673613619492>.

- [18] Simon Brooker. “Spatial epidemiology of human schistosomiasis in Africa: risk models, transmission dynamics and control”. In: *Transactions of the Royal Society of Tropical Medicine and Hygiene* 101.1 (2007), pp. 1–8. ISSN: 00359203. DOI: 10.1016/j.trstmh.2006.08.004.
- [19] Azalech Tefera, Tariku Belay, and Mitiku Bajiro. “Epidemiology of *Schistosoma mansoni* infection and associated risk factors among school children attending primary schools nearby rivers in Jimma town, an urban setting, Southwest Ethiopia”. In: *PLOS ONE* 15.2 (2020). Ed. by David Joseph Diemert, e0228007. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0228007. URL: <https://dx.plos.org/10.1371/journal.pone.0228007>.
- [20] Maria E. Bavia *et al.* “Geographic information systems and the environmental risk of schistosomiasis in Bahia, Brazil”. In: *American Journal of Tropical Medicine and Hygiene* 60.4 (1999), pp. 566–572. ISSN: 00029637. DOI: 10.4269/ajtmh.1999.60.566.
- [21] Julien Kincaid-Smith *et al.* *Emerging Schistosomiasis in Europe: A Need to Quantify the Risks*. 2017. DOI: 10.1016/j.pt.2017.04.009.
- [22] J. Russell Stothard and Albis Francesco Gabrielli. *Schistosomiasis in African infants and preschool children: to treat or not to treat?* 2007. DOI: 10.1016/j.pt.2007.01.005.
- [23] J. Russell Stothard *et al.* *Schistosomiasis in African infants and preschool children: Let them now be treated!* 2013. DOI: 10.1016/j.pt.2013.02.001.
- [24] Hlengiwe Sacolo, Moses Chimbari, and Chester Kalinda. “Knowledge, attitudes and practices on Schistosomiasis in sub-Saharan Africa: A systematic review”. In: *BMC Infectious Diseases* 18.1 (2018), pp. 1–17. ISSN: 14712334. DOI: 10.1186/s12879-017-2923-6. URL: <https://link.springer.com/articles/10.1186/s12879-017-2923-6>
<https://link.springer.com/article/10.1186/s12879-017-2923-6>.
- [25] Donald P. McManus *et al.* *Schistosomiasis—from immunopathology to vaccines*. 2020. DOI: 10.1007/s00281-020-00789-x.
- [26] M Ng *et al.* *Spatial distribution and co-infection with urogenital and intestinal schistosomiasis among primary school children in Migori County, Kenya*. Tech. rep. 10. 2016, pp. 22–31. DOI: 10.4314/eamj.v93i10.. URL: <https://www.ajol.info/index.php/eamj/article/view/150689>.

- [27] Nmorsi Ukwandu. *Urinary tract pathology in some Schistosoma haematobium infected Nigerians Neglected Tropical Diseases (NTDs) in Nigeria: The Burden of Schistosoma haematobium infection in Nigeria View project*. Tech. rep. 2007. URL: <https://www.researchgate.net/publication/6297462>.
- [28] C. H. King *et al.* “Urinary tract morbidity in schistosomiasis haematobia: Associations with age and intensity of infection in an endemic area of Coast Province, Kenya”. In: *American Journal of Tropical Medicine and Hygiene* 39.4 (1988), pp. 361–368. ISSN: 00029637. DOI: 10.4269/ajtmh.1988.39.361.
- [29] Alex N. Wamachi *et al.* “Increased Ratio of Tumor Necrosis Factor- α to Interleukin-10 Production Is Associated with *Schistosoma haematobium* –Induced Urinary-Tract Morbidity”. In: *The Journal of Infectious Diseases* 190.11 (2004), pp. 2020–2030. ISSN: 0022-1899. DOI: 10.1086/425579. URL: <https://academic.oup.com/jid/article-lookup/doi/10.1086/425579>.
- [30] Eyrun F. Kjetland, Peter D.C. Leutscher, and Patricia D. Ndhlovu. *A review of female genital schistosomiasis*. 2012. DOI: 10.1016/j.pt.2011.10.008.
- [31] Sekeleghe Kayuni *et al.* *A systematic review with epidemiological update of male genital schistosomiasis (MGS): A call for integrated case management across the health system in sub-Saharan Africa*. 2018. DOI: 10.1016/j.parepi.2018.e00077.
- [32] Julie Balen *et al.* “Morbidity due to *Schistosoma mansoni*: an epidemiological assessment of distended abdomen syndrome in Ugandan school children with observations before and 1-year after anthelmintic chemotherapy”. In: *Transactions of the Royal Society of Tropical Medicine and Hygiene* 100.11 (2006), pp. 1039–1048. ISSN: 00359203. DOI: 10.1016/j.trstmh.2005.12.013. URL: <https://academic.oup.com/trstmh/article-lookup/doi/10.1016/j.trstmh.2005.12.013>.
- [33] J. Richter *et al.* “Sonographic prediction of variceal bleeding in patients with liver fibrosis due to *Schistosoma mansoni*”. In: *Tropical Medicine & International Health* 3.9 (1998), pp. 728–735. ISSN: 1360-2276. DOI: 10.1046/j.1365-3156.1998.00285.x. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-3156.1998.00285.x>.
- [34] Adebayo J. Molehin. *Schistosomiasis vaccine development: Update on human clinical trials*. 2020. DOI: 10.1186/s12929-020-0621-y.

- [35] Michael J. Doenhoff, Peter L. Chiodini, and Joanne V. Hamilton. *Specific and sensitive diagnosis of schistosome infection: Can it be done with antibodies?* 2004. DOI: 10.1016/j.pt.2003.10.019.
- [36] Marta G. Cavalcanti *et al.* *Schistosomiasis in areas of low endemicity: A new era in diagnosis.* 2013. DOI: 10.1016/j.pt.2012.11.003.
- [37] Rosie Christiansen. “A parasitological survey to ascertain the prevalence of intestinal schistosomiasis in school-aged children around a new focus of *Biomphalaria* in Lake Malawi”. In: *Msc Biology and Control of Parasites and Disease vectors, Liverpool School Tropical Medicine* (2018).
- [38] Angus More O’Ferrall. “The changing epidemiological landscape of schistosomiasis in Lake Malawi, Mangochi District: prevalence and morbidity associated with urogenital schistosomiasis in school children.” In: *Msc Biology and Control of Parasites and Disease vectors, Liverpool School Tropical Medicine* (2019).
- [39] WHO. *Deworming school-age children Helminth control in school-age children Second edition A guide for managers of control programmes.* 2011. ISBN: 9789241548267. URL: http://www.who.int/neglected{_}diseases/en.
- [40] J. E. McMahon. “Circadian rhythm in *Schistosoma haematobium* egg excretion”. In: *International Journal for Parasitology* 6.5 (1976), pp. 373–377. ISSN: 00207519. DOI: 10.1016/0020-7519(76)90021-7.
- [41] J. Utzinger *et al.* *New diagnostic tools in schistosomiasis.* 2015. DOI: 10.1016/j.j.cmi.2015.03.014.
- [42] S. J. De Vlas *et al.* “Validation of a chart to estimate true *Schistosoma mansoni* prevalences from simple egg counts”. In: *Parasitology* 114.2 (1997), pp. 113–121. ISSN: 00311820. DOI: 10.1017/S0031182096008207.
- [43] Dan Dan Lin *et al.* “Routine Kato-Katz technique underestimates the prevalence of *Schistosoma japonicum*: A case study in an endemic area of the People’s Republic of China”. In: *Parasitology International* 57.3 (2008), pp. 281–286. ISSN: 13835769. DOI: 10.1016/j.parint.2008.04.005.
- [44] Nega Berhe *et al.* “Variations in helminth faecal egg counts in Kato-Katz thick smears and their implications in assessing infection status with *Schistosoma mansoni*”. In: *Acta Tropica* 92.3 (2004), pp. 205–212. ISSN: 0001706X. DOI: 10.1016/j.actatropica.2004.06.011.

- [45] A. Kongs *et al.* “The unreliability of the Kato-Katz technique limits its usefulness for evaluating *S. mansoni* infections”. In: *Tropical Medicine and International Health* 6.3 (2001), pp. 163–169. ISSN: 1360-2276. DOI: 10.1046/j.1365-3156.2001.00687.x. URL: <http://doi.wiley.com/10.1046/j.1365-3156.2001.00687.x>.
- [46] Sabrina Menezes da Frota *et al.* “Combination of Kato-Katz faecal examinations and ELISA to improve accuracy of diagnosis of intestinal schistosomiasis in a low-endemic setting in Brazil”. In: *Acta Tropica* 120.SUPPL. 1 (2011), S138–S141. ISSN: 0001706X. DOI: 10.1016/j.actatropica.2010.05.007.
- [47] Hermann Feldmeier and Gabriele Poggensee. “Diagnostic techniques in schistosomiasis control. A review”. In: *Acta Tropica* 52.4 (1993), pp. 205–220. ISSN: 0001706X. DOI: 10.1016/0001-706X(93)90009-Z.
- [48] Collins Okoyo *et al.* “Comparing the performance of circulating cathodic antigen and Kato-Katz techniques in evaluating *Schistosoma mansoni* infection in areas with low prevalence in selected counties of Kenya: A cross-sectional study”. In: *BMC Public Health* 18.1 (2018), p. 478. ISSN: 14712458. DOI: 10.1186/s12889-018-5414-9. URL: <https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-018-5414-9>.
- [49] World Health Organisation. “Prevention and control of schistosomiasis and soil-transmitted helminthiasis : report of a WHO expert committee”. In: (2002). URL: https://apps.who.int/iris/bitstream/handle/10665/42588/WHO_{_}TRS_{_}912.pdf?sequence=1{\&}isAllowed=y.
- [50] R.J. Pitchford and P.S. Visser. “A simple and rapid technique for quantitative estimation of helminth eggs in human and animal excreta with special reference to *Schistosoma* sp.” In: *Transactions of the Royal Society of Tropical Medicine and Hygiene* 69.3 (1975), pp. 318–322. ISSN: 00359203. DOI: 10.1016/0035-9203(75)90126-1. URL: [https://academic.oup.com/trstmh/article-lookup/doi/10.1016/0035-9203\(75\)90126-1](https://academic.oup.com/trstmh/article-lookup/doi/10.1016/0035-9203(75)90126-1).
- [51] Michael D. French *et al.* “School-based control of urinary schistosomiasis on Zanzibar, Tanzania: Monitoring micro-haematuria with reagent strips as a rapid urological assessment”. In: *Journal of Pediatric Urology* 3.5 (2007), pp. 364–368. ISSN: 14775131. DOI: 10.1016/j.jpuro.2007.01.198.

- [52] Emmanuel Emukah *et al.* “Urine heme dipsticks are useful in monitoring the impact of Praziquantel treatment on *Schistosoma haematobium* in sentinel communities of Delta State, Nigeria”. In: *Acta Tropica* 122.1 (2012), pp. 126–131. ISSN: 0001706X. DOI: 10.1016/j.actatropica.2012.01.002.
- [53] Bruno Gryseels *et al.* *Human schistosomiasis*. 2006. DOI: 10.1016/S0140-6736(06)69440-3.
- [54] Lisette Van Lieshout *et al.* *Analysis of Worm Burden Variation in Human Schistosoma mansoni Infections by Determination of Serum Levels of Circulating Anodic Antigen and Circulating Cathodic Antigen*. Tech. rep. 1995. URL: <https://academic.oup.com/jid/article-abstract/172/5/1336/950422>.
- [55] Berhanu Erko *et al.* “Evaluation of urine-circulating cathodic antigen (Urine-CCA) cassette test for the detection of *Schistosoma mansoni* infection in areas of moderate prevalence in Ethiopia”. In: *Tropical Medicine & International Health* 18.8 (2013), pp. 1029–1035. ISSN: 13602276. DOI: 10.1111/tmi.12117. URL: <http://doi.wiley.com/10.1111/tmi.12117>.
- [56] Oliver Bärenbold *et al.* “Translating preventive chemotherapy prevalence thresholds for *Schistosoma mansoni* from the Kato-Katz technique into the point-of-care circulating cathodic antigen diagnostic test”. In: *PLOS Neglected Tropical Diseases* 12.12 (2018). Ed. by Jean-Philippe Chippaux, e0006941. ISSN: 1935-2735. DOI: 10.1371/journal.pntd.0006941. URL: <https://dx.plos.org/10.1371/journal.pntd.0006941>.
- [57] Victor C.W. Tsang and Patricia P. Wilkins. “Emmunodiagnosis of Schistosomiasis”. In: *Immunological Investigations* 26.42006 (1997), pp. 175–188. ISSN: 15324311. DOI: 10.3109/08820139709048925.
- [58] Tamer Elbaz and Gamal Esmat. *Hepatic and Intestinal Schistosomiasis: Review*. 2013. DOI: 10.1016/j.jare.2012.12.001.
- [59] Robert J. ten Hove *et al.* “Multiplex real-time PCR for the detection and quantification of *Schistosoma mansoni* and *S. haematobium* infection in stool samples collected in northern Senegal”. In: *Transactions of the Royal Society of Tropical Medicine and Hygiene* 102.2 (2008), pp. 179–185. ISSN: 00359203. DOI: 10.1016/j.trstmh.2007.10.011. URL: <https://academic.oup.com/trstmh/article-lookup/doi/10.1016/j.trstmh.2007.10.011>.

- [60] Jan Clerinx and Alfons Van Gompel. “Schistosomiasis in travellers and migrants”. In: *Travel Medicine and Infectious Disease* 9.1 (2011), pp. 6–24. ISSN: 14778939. DOI: 10.1016/j.tmaid.2010.11.002.
- [61] Donato Cioli *et al.* *Schistosomiasis control: Praziquantel forever?* 2014. DOI: 10.1016/j.molbiopara.2014.06.002.
- [62] Amaya L. Bustinduy *et al.* “Population pharmacokinetics and pharmacodynamics of praziquantel in Ugandan children with intestinal schistosomiasis: Higher dosages are required for maximal efficacy”. In: *mBio* 7.4 (2016). ISSN: 21507511. DOI: 10.1128/mBio.00227-16.
- [63] WHO. *WHO GUIDELINE on control and elimination of human schistosomiasis*. 2022, p. 142. ISBN: 978 92 4 004160 8. URL: <https://www.who.int/publications/i/item/9789240041608>.
- [64] Livia Pica-Mattocchia and Donato Cioli. “Sex and stage-related sensitivity of *Schistosoma mansoni* to in vivo and in vitro praziquantel treatment”. In: *International Journal for Parasitology* 34.4 (2004), pp. 527–533. ISSN: 00207519. DOI: 10.1016/j.ijpara.2003.12.003.
- [65] R. Alan Wilson. “Schistosomiasis then and now: What has changed in the last 100 years?” In: *Parasitology* 147.5 (2020), pp. 507–515. ISSN: 14698161. DOI: 10.1017/S0031182020000049.
- [66] P. G. Fallon and M. J. Doenhoff. “Drug-resistant schistosomiasis: Resistance to praziquantel and oxamniquine induced in *Schistosoma mansoni* in mice is drug specific”. In: *American Journal of Tropical Medicine and Hygiene* 51.1 (1994), pp. 83–88. ISSN: 00029637. DOI: 10.4269/ajtmh.1994.51.83.
- [67] WHO. “Ending the neglect to attain the Sustainable Development Goals: a road map for neglected tropical diseases 2021–2030.” In: *World Health Organization* (2020), p. 196. URL: <https://www.who.int/publications/i/item/9789240010352>.
- [68] Nathan C. Lo *et al.* *Review of 2022 WHO guidelines on the control and elimination of schistosomiasis*. 2022. DOI: 10.1016/S1473-3099(22)00221-3. URL: <http://www.thelancet.com/article/S1473309922002213/fulltext><http://www.thelancet.com/article/S1473309922002213/abstract>[https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(22\)00221-3/abstract](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(22)00221-3/abstract).

- [69] H. Madsen. “Ecological studies on the intermediate host snails and the relevance to schistosomiasis control.” In: *Memórias do Instituto Oswaldo Cruz* 87 Suppl 4 (1992), pp. 249–253. ISSN: 00740276. DOI: 10.1590/S0074-02761992000800039. URL: <https://pubmed.ncbi.nlm.nih.gov/1343904/>.
- [70] CDC. *CDC - Schistosomiasis - Prevention & Control*. 2018. URL: <https://www.cdc.gov/parasites/schistosomiasis/prevent.html> (visited on 05/22/2020).
- [71] Martin C. Arostegui *et al.* “Potential biological control of schistosomiasis by fishes in the lower Senegal river basin”. In: *American Journal of Tropical Medicine and Hygiene* 100.1 (2019), pp. 117–126. ISSN: 00029637. DOI: 10.4269/ajtmh.18-0469.
- [72] Charles H. King and David Bertsch. “Historical Perspective: Snail Control to Prevent Schistosomiasis”. In: *PLoS Neglected Tropical Diseases* 9.4 (2015). ISSN: 19352735. DOI: 10.1371/journal.pntd.0003657. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4408102/>.
- [73] Susanne H. Sokolow *et al.* *To Reduce the Global Burden of Human Schistosomiasis, Use ‘Old Fashioned’ Snail Control*. 2018. DOI: 10.1016/j.pt.2017.10.002.
- [74] United Nations. *Malawi Country Profile - United Nations*. 2014. URL: <http://mw.one.un.org/country-profile/> (visited on 06/02/2020).
- [75] National Statistical Office. *MALAWI POPULATION AND HOUSING CENSUS REPORT-2018 2018 Malawi Population and Housing Main Report*. Tech. rep. 2019.
- [76] Henry Madsen *et al.* *Schistosomiasis in Lake Malawi villages*. 2011. DOI: 10.1007/s10393-011-0687-9.
- [77] C H Teesdale and L Chitsulo. “Schistosomiasis in Malawi—a review.” In: *Tropical medicine and parasitology : official organ of Deutsche Tropenmedizinische Gesellschaft and of Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ)* 36.1 (1985), pp. 1–6. ISSN: 0177-2392. URL: <http://www.ncbi.nlm.nih.gov/pubmed/4001764>.
- [78] Sekeleghe A. Kayuni *et al.* “Male Genital Schistosomiasis Along the Shoreline of Lake Malawi: Baseline Prevalence and Associated Knowledge, Attitudes and Practices Among Local Fishermen in Mangochi District, Malawi”. In: *Frontiers in Public Health* 9 (2021), p. 590695. ISSN: 22962565. DOI: 10.3389/FPUBH.2021.590695/BIBTEX.

- [79] Bwanali A.N. Munthali L.E. Mphepo M. Chumbi G.D. Kangoma M. Matola Y. Kaonga B. Lubanga A.F. and C.S. Moyo. “Exploring the Role of Community Involvement in Reducing the Burden of Schistosomiasis and Other Neglected Tropical Diseases in Malawi: Where are We in the Fight Against Neglected Tropical Diseases?” In: (2024), pp. 51–58.
- [80] *Find a map — Global Atlas of Helminth Infections*. 2020. URL: <http://www.thiswormyworld.org/maps/find-a-map> (visited on 04/26/2020).
- [81] J. Russell Stothard *et al.* “Towards interruption of schistosomiasis transmission in sub-Saharan Africa: Developing an appropriate environmental surveillance framework to guide and to support ‘end game’ interventions”. In: *Infectious Diseases of Poverty* 6.1 (2017), pp. 1–11. ISSN: 20499957. DOI: 10.1186/s40249-016-0215-9. URL: <https://link.springer.com/articles/10.1186/s40249-016-0215-9>
<https://link.springer.com/article/10.1186/s40249-016-0215-9>.
- [82] Bert Van Bocxlaer, Christian Albrecht, and Jay R. Stauffer. *Growing population and ecosystem change increase human schistosomiasis around Lake Malaŵi*. 2014. DOI: 10.1016/j.pt.2014.02.006.
- [83] Selpha Opisa *et al.* “Malacological survey and geographical distribution of vector snails for schistosomiasis within informal settlements of Kisumu City, western Kenya”. In: *Parasites and Vectors* 4.1 (2011), p. 226. ISSN: 17563305. DOI: 10.1186/1756-3305-4-226. URL: <http://parasitesandvectors.biomedcentral.com/articles/10.1186/1756-3305-4-226>.
- [84] *IUCN Red List of Threatened Species*. 2020. URL: <https://www.iucnredlist.org/search?query=Bulinus{\&}searchType=species> (visited on 09/09/2020).
- [85] Tayo Alex Adekiya *et al.* “The effect of climate change and the snail-schistosome cycle in transmission and bio-control of schistosomiasis in sub-saharan africa”. In: *International Journal of Environmental Research and Public Health* 17.1 (2020). ISSN: 16604601. DOI: 10.3390/ijerph17010181. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6981654/>.
- [86] K. N. De Kock, C. T. Wolmarans, and M. Bornman. “Distribution and habitats of *Biomphalaria pfeifferi*, snail intermediate host of *Schistosoma mansoni*, in South Africa”.

- In: *Water SA* 30.1 (2004), pp. 29–36. ISSN: 03784738. DOI: 10.4314/wsa.v30i1.5023.
- [87] Yvonne Walz *et al.* *Risk profiling of schistosomiasis using remote sensing: Approaches, challenges and outlook*. 2015. DOI: 10.1186/s13071-015-0732-6. URL: [/pmc/articles/PMC4406176/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/264406176/)<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4406176/>.
- [88] Samuel Nii Ardey Codjoe and Reuben Tete Larbi. “Climate change/variability and schistosomiasis transmission in Ga district, Ghana”. In: *Climate and Development* 8.1 (2016), pp. 58–71. ISSN: 1756-5529. DOI: 10.1080/17565529.2014.998603. URL: <http://www.tandfonline.com/doi/full/10.1080/17565529.2014.998603>.
- [89] A. S. Stensgaard *et al.* “Modeling freshwater snail habitat suitability and areas of potential snail-borne disease transmission in Uganda.” In: *Geospatial health* 1.1 (2006), pp. 93–104. ISSN: 19707096. DOI: 10.4081/gh.2006.284. URL: <https://pubmed.ncbi.nlm.nih.gov/18686235/><https://pubmed.ncbi.nlm.nih.gov/18686235/?dopt=Abstract>.
- [90] Candia Rowel *et al.* “Environmental epidemiology of intestinal schistosomiasis in Uganda: Population dynamics of *Biomphalaria* (Gastropoda: Planorbidae) in Lake Albert and Lake Victoria with observations on natural infections with digenetic trematodes”. In: *BioMed Research International* 2015 (2015). ISSN: 23146141. DOI: 10.1155/2015/717261.
- [91] Sarah Levitz *et al.* “Environmental epidemiology of intestinal schistosomiasis and genetic diversity of *Schistosoma mansoni* infections in snails at Bugoigo village, Lake Albert”. In: *Acta Tropica* 128.2 (2013), pp. 284–291. ISSN: 0001706X. DOI: 10.1016/j.actatropica.2012.10.003. URL: <https://pubmed.ncbi.nlm.nih.gov/23085327/>.
- [92] Matthew R. Behrend *et al.* “Modelling for policy: The five principles of the Neglected Tropical Diseases Modelling Consortium”. In: *PLOS Neglected Tropical Diseases* 14.4 (2020), e0008033. ISSN: 1935-2735. DOI: 10.1371/JOURNAL.PNTD.0008033. URL: <https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0008033>.
- [93] Grant Brown and Marie Ozanne. “Statistical Models for Infectious Diseases: A Useful Tool for Practical Decision-Making”. In: *The American Journal of Tropical Medicine and Hygiene* 101.1 (2019), p. 1. DOI: 10.4269/AJTMH.19-0354. URL:

- /pmc/articles/PMC6609166/https:
//www.ncbi.nlm.nih.gov/pmc/articles/PMC6609166/.
- [94] Liang Shi *et al.* “Development of New Technologies for Risk Identification of Schistosomiasis Transmission in China”. In: *Pathogens* 2022, Vol. 11, Page 224 11.2 (2022), p. 224. ISSN: 2076-0817. DOI: 10.3390/PATHOGENS11020224. URL: <https://www.mdpi.com/2076-0817/11/2/224/htmhttps://www.mdpi.com/2076-0817/11/2/224>.
- [95] Simon Brooker, Simon I. Hay, and Don A.P. Bundy. “Tools from ecology: useful for evaluating infection risk models?” In: *Trends in Parasitology* 18.2 (2002), pp. 70–74. ISSN: 1471-4922. DOI: 10.1016/S1471-4922(01)02223-1.
- [96] J.A. Nelder and R.W Wedderburn. “Generalized linear models”. In: 135 (3).
- [97] T J Hastie and R J Tibsh. *Generalized Additive Models - 1st Edition*. Chapman & Hall/CRC, 1990, p. 352. ISBN: 978-0-412-34390-2. DOI: <https://doi.org/10.1201/9780203753781>. URL: <https://www.routledge.com/Generalized-Additive-Models/Hastie-Tibshirani/p/book/9780412343902>.
- [98] In: ().
- [99] E Diggle, P and Giorgi. *Model-based Geostatistics for Global Public Health - Research Portal — Lancaster University*. 2019. URL: [http://www.research.lancs.ac.uk/portal/en/publications/modelbased-geostatistics-for-global-public-health\(f5bb2161-03b4-40c7-9a9d-ca465cb36557\).html](http://www.research.lancs.ac.uk/portal/en/publications/modelbased-geostatistics-for-global-public-health(f5bb2161-03b4-40c7-9a9d-ca465cb36557).html) (visited on 07/05/2022).
- [100] Alan E. Gelfand and Sudipto Banerjee. “Bayesian Modeling and Analysis of Geostatistical Data”. In: *Annual review of statistics and its application* 4 (2017), p. 245. ISSN: 2326831X. DOI: 10.1146/ANNUREV-STATISTICS-060116-054155. URL: [/pmc/articles/PMC5790124//pmc/articles/PMC5790124/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5790124/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5790124/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5790124/).
- [101] Ross R. “The prevention of malaria”. In: (191), pp. 651–86.
- [102] McKendrick AG Kermack WO. “Contribution to the mathematical theory of epidemics”. In: ().
- [103] McKendrick AG Kermack WO. “Contributions to the mathematical theory of epidemics, part II”. In: (1932).

- [104] McKendrick AG Kermack WO. “Contributions to the mathematical theory of epidemics, part III”. In: (1933), pp. 94–112.
- [105] Frost W.H. “Some conceptions of epidemics in general”. In: (1976), pp. 141–51.
- [106] Nelson G. Hairston. “Population Ecology and Epidemiological Problems”. In: *John Wiley & Sons Ltd. John Wiley & Sons, Ltd, 1962*, pp. 36–80. DOI: 10.1002/9780470719312.ch4. URL: <https://onlinelibrary.wiley.com/doi/full/10.1002/9780470719312.ch4><https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470719312.ch4><https://onlinelibrary.wiley.com/doi/10.1002/9780470719312.ch4>.
- [107] N. G. Hairston. “On the mathematical analysis of schistosome populations.” In: *Bulletin of the World Health Organization* 33.1 (1965), pp. 45–62. ISSN: 00429686. URL: <https://pubmed.ncbi.nlm.nih.gov/abstract/PMC2475813/><https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2475813/>.
- [108] George Macdonald. “The dynamics of helminth infections, with special reference to schistosomes”. In: *Transactions of the Royal Society of Tropical Medicine and Hygiene* 59.5 (1965), pp. 489–506. ISSN: 18783503. DOI: 10.1016/0035-9203(65)90152-5.
- [109] R. M. Anderson and R. M. May. “Herd immunity to helminth infection and implications for parasite control”. In: *Nature* 315.6019 (1985), pp. 493–496. ISSN: 00280836. DOI: 10.1038/315493a0. URL: <https://www.nature.com/articles/315493a0>.
- [110] Roy M Anderson *et al.* “Advances in parasitology. Volume ninety four, Mathematical models neglected tropical diseases : essential tools for control and elimination, Part B”. In: *Elsevier Science and Technology* 94 (2016), pp. 199–246. URL: [https://books.google.com/books?id=jD3jCwAAQBAJ{\&pg=PA208{\&lpg=PA208{\&dq=The+aggregation+parameter+k+varies+inversely+with+the+degree+of+aggregation+or+clumping,{\&source=bl{\&ots=Y1XmP-Ox1{_}{\&sig=U5xqW5MbUasbvK2IblxDyyc9XiE{\&hl=en{\&sa=X{\&ved=0ahUKEwjssYzH37LXAhVHZCYK](https://books.google.com/books?id=jD3jCwAAQBAJ&pg=PA208&lpg=PA208&dq=The+aggregation+parameter+k+varies+inversely+with+the+degree+of+aggregation+or+clumping,&source=bl&ots=Y1XmP-Ox1{_}{\&sig=U5xqW5MbUasbvK2IblxDyyc9XiE{\&hl=en{\&sa=X{\&ved=0ahUKEwjssYzH37LXAhVHZCYK).
- [111] Ebrima Kanyi, Ayodeji Sunday Afolabi, and Nelson Owuor Onyango. “Mathematical Modeling and Analysis of Transmission Dynamics and Control of Schistosomiasis”. In: *Journal of Applied Mathematics* 2021 (2021). ISSN: 16870042. DOI: 10.1155/2021/6653796.

- [112] W. Evan Secor *et al.* “Insights from quantitative and mathematical modelling on the proposed WHO 2030 goal for schistosomiasis”. In: *Gates Open Research* 3 (2019). ISSN: 25724754. DOI: 10.12688/GATESOPENRES.13052.2. URL: /pmc/articles/PMC6820450//pmc/articles/PMC6820450/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6820450/.
- [113] M. S. Chan *et al.* “The development of an age structured model for schistosomiasis transmission dynamics and control and its validation for *Schistosoma mansoni*”. In: *Epidemiology & Infection* 115.2 (1995), pp. 325–344. ISSN: 1469-4409. DOI: 10.1017/S0950268800058453. URL: https://www.cambridge.org/core/journals/epidemiology-and-infection/article/development-of-an-age-structured-model-for-schistosomiasis-transmission-dynamics-and-control-and-its-validation-for-schistosoma-mansoni/8E89442C2C08BFF0C5AA37EC80413E06.
- [114] Matthew Graham *et al.* “SCHISTOX: An individual based model for the epidemiology and control of schistosomiasis”. In: *Infectious Disease Modelling* 6 (2021), p. 438. ISSN: 24680427. DOI: 10.1016/J.IDM.2021.01.010. URL: /pmc/articles/PMC7897994//pmc/articles/PMC7897994/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7897994/.
- [115] David Gurarie *et al.* “Modelling control of *Schistosoma haematobium* infection: predictions of the long-term impact of mass drug administration in Africa”. In: *Parasites and Vectors* 8.1 (2015), pp. 1–14. ISSN: 17563305. DOI: 10.1186/S13071-015-1144-3/FIGURES/8. URL: https://parasitesandvectors.biomedcentral.com/articles/10.1186/s13071-015-1144-3.
- [116] Anna Borlase, Joaquin M. Prada, and Thomas Crellen. *Modelling morbidity for neglected tropical diseases: the long and winding road from cumulative exposure to long-term pathology*. 2023. DOI: 10.1098/rstb.2022.0279. URL: https://royalsocietypublishing.org/doi/10.1098/rstb.2022.0279.
- [117] Suryanto A. Nur W. and W.M. Kusumawinahyu. “Mathematical model of schistosomiasis with health education and molluscicide intervention”. In: 1821 (1 2021), 012033).
- [118] Chazuka Z. Onwubuya I.O. Fatmawati F. Madubueze C.E. and C.W. Chukwu. “On the mathematical modeling of schistosomiasis transmission dynamics with heterogeneous intermediate host”. In: (2022).

- [119] Shen D. Zhang W. Chen Z. Zou L. and Ruan S. “Mathematical modelling and control of schistosomiasis in Hubei Province, China”. In: 115 (1-2 2010), 119–125.
- [120] Majid S Syed R. Rychtar J. Lopez S and D. Taylor. “Mathematical model of voluntary vaccination against schistosomiasis”. In: 12 (2024).
- [121] Patrice A. Mawa *et al.* *Schistosomiasis morbidity hotspots: roles of the human host, the parasite and their interface in the development of severe morbidity*. 2021. DOI: 10.3389/fimmu.2021.635869. URL: www.frontiersin.org.
- [122] Anouk N. Gouvras *et al.* “The impact of single versus mixed *Schistosoma haematobium* and *S. mansoni* infections on morbidity profiles amongst school-children in Taveta, Kenya”. In: *Acta Tropica* 128.2 (2013), pp. 309–317. ISSN: 0001706X. DOI: 10.1016/j.actatropica.2013.01.001.
- [123] M. E.J Woolhouse. “Patterns in parasite epidemiology: the peak shift”. In: *Parasitology Today* 14.10 (1998), pp. 428–434. URL: [https://doi.org/10.1016/S0169-4758\(98\)01318-0](https://doi.org/10.1016/S0169-4758(98)01318-0).
- [124] Salwa Dawaki *et al.* “Prevalence and risk factors of schistosomiasis among Hausa communities in Kano state, Nigeria”. In: *Revista do Instituto de Medicina Tropical de Sao Paulo* 58.1 (2016). ISSN: 16789946. DOI: 10.1590/S1678-9946201658054. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4964323/>.
- [125] O. G. Oso and A. B. Odaibo. “Human water contact patterns in active schistosomiasis endemic areas”. In: *Journal of Water and Health* 18.6 (2020), pp. 946–955. ISSN: 19967829. DOI: 10.2166/wh.2020.147. URL: <http://iwaponline.com/jwh/article-pdf/18/6/946/824717/jwh0180946.pdf>.
- [126] Nupur Kittur *et al.* “Defining persistent hotspots: Areas that fail to decrease meaningfully in prevalence after multiple years of mass drug administration with praziquantel for control of schistosomiasis”. In: *American Journal of Tropical Medicine and Hygiene* 97.6 (2017), pp. 1810–1817. ISSN: 00029637. DOI: 10.4269/ajtmh.17-0368. URL: <https://score.uga.edu/>.
- [127] Sekeleghe A. Kayuni *et al.* “An outbreak of intestinal schistosomiasis, alongside increasing urogenital schistosomiasis prevalence, in primary school children on the shoreline of Lake Malawi, Mangochi District, Malawi”. In: *Infectious Diseases of Poverty* 9.1 (2020). ISSN: 20499957. DOI: 10.1186/s40249-020-00736-w. URL: [/pmc/articles/PMC7456765/](https://pmc/articles/PMC7456765/)
<https://pmc/articles/PMC7456765/?report=>

- abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7456765/.
- [128] WHO. “The control of schistosomiasis. Second report of the WHO Expert Committee.” In: *World Health Organization - Technical Report Series* 830 (1993), pp. 1–86. ISSN: 05123054. DOI: 10.1016/0035-9203(93)90322-h.
- [129] Cavalcanti MG Peralta JM. “Is POC-CCA a truly reliable test for schistosomiasis diagnosis in low endemic areas? The trace results controversy.” In: 11 (12 2018).
- [130] Simon Wood. “Package ‘mgcv’ Mixed GAM computation vehicle with automatic smoothness estimation”. In: 1.8-31 (2019). URL: <https://cran.r-project.org/web/packages/mgcv/mgcv.pdf>.
- [131] Andrew Gelman and Jennifer Hill. *Data analysis using regression and multilevel/hierarchical models*. 2007. URL: [https://books.google.co.uk/books?hl=en&lr=&id=c9xLKzZWoz4C&oi=fnd&pg=PR17&dq={\%}5B10:52{\%}5D+Fronterre,+Claudio+Gelman,+A.,+and+Hill,+J.+ \(2007\).+Data+analysis+using+regression+and+multilevel/hierarchical+models.+Cambridge{\%}3B+New+York:+Cambridge+University+Pre \(visited on 11/21/2022\)](https://books.google.co.uk/books?hl=en&lr=&id=c9xLKzZWoz4C&oi=fnd&pg=PR17&dq={\%}5B10:52{\%}5D+Fronterre,+Claudio+Gelman,+A.,+and+Hill,+J.+ (2007).+Data+analysis+using+regression+and+multilevel/hierarchical+models.+Cambridge{\%}3B+New+York:+Cambridge+University+Pre (visited on 11/21/2022)).
- [132] Kiran Singh and D Muddasiru. “Epidemiology of schistosomiasis in school aged children in some riverine areas of Sokoto, Nigeria”. In: *Journal of Public Health and Epidemiology Full Length Research Paper* 6.6 (2014), pp. 197–201. ISSN: 2006-9723. DOI: 10.5897/JPHE2013.0580. URL: <http://www.academicjournals.org/JPHE>.
- [133] Humphrey Deogratias Mazigo *et al.* “Prevalence, infection intensity and geographical distribution of schistosomiasis among pre-school and school aged children in villages surrounding Lake Nyasa, Tanzania”. In: *Scientific Reports* 11.1 (2021), p. 295. ISSN: 20452322. DOI: 10.1038/s41598-020-80317-x. URL: <https://doi.org/10.1038/s41598-020-80317-x>.
- [134] Mohammad H. Alharbi *et al.* “First molecular identification of *Bulinus africanus* in Lake Malawi implicated in transmitting *Schistosoma* parasites”. In: *Tropical medicine and infectious disease* 7.8 (2022). ISSN: 2414-6366. DOI: 10.3390/TROPICALMED7080195. URL: <https://pubmed.ncbi.nlm.nih.gov/36006287/>.

- [135] Mohammad H. Alharbi *et al.* “*Biomphalaria pfeifferi* (Gastropoda: Planorbidae) in Lake Malawi and Upper Shire River, Mangochi District, Malawi: Distribution, Genetic Diversity and Pre-Patent Schistosome Infections”. In: *Tropical Medicine and Infectious Disease* 2023, Vol. 8, Page 126 8.2 (2023), p. 126. ISSN: 2414-6366. DOI: 10.3390/TROPICALMED8020126. URL: <https://www.mdpi.com/2414-6366/8/2/126>.
- [136] Pauline N.M. Mwinzi *et al.* “Predictive value of school-aged children’s schistosomiasis prevalence and egg intensity for other age groups in Western Kenya”. In: *American Journal of Tropical Medicine and Hygiene* 93.6 (2015), pp. 1311–1317. ISSN: 00029637. DOI: 10.4269/ajtmh.15-0467. URL: <https://www.ajtmh.org/view/journals/tpmd/93/6/article-p1311.xml>.
- [137] Michał Laskowski *et al.* “On the usefulness of the generalised additive model for mean path loss estimation in body area networks”. In: *IEEE Access* 8 (2020), pp. 376873–376882. ISSN: 21693536. DOI: 10.1109/ACCESS.2020.3025118.
- [138] Henry Madsen and Jay Richard Stauffer. “Schistosomiasis Control Under Changing Ecological Settings in Lake Malawi”. In: *EcoHealth* 19.3 (2022), pp. 320–323. ISSN: 16129210. DOI: 10.1007/s10393-022-01606-7/METRICS. URL: <https://link.springer.com/article/10.1007/s10393-022-01606-7>.
- [139] Richard Ogutu-Ohwayo and John S. Balirwa. “Management challenges of freshwater fisheries in Africa”. In: *Lakes & Reservoirs: Research & Management* 11.4 (2006), pp. 215–226. ISSN: 1440-1770. DOI: 10.1111/J.1440-1770.2006.00312.X. URL: <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1440-1770.2006.00312.x>
<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1440-1770.2006.00312.x>
<https://onlinelibrary.wiley.com/doi/10.1111/j.1440-1770.2006.00312.x>.
- [140] Jay Stauffer *et al.* “A one health approach to reducing schistosomiasis transmission in Lake Malawi”. In: *Preventive Medicine and Community Health Prev Med Commun Health* 1.3 (2018), pp. 1–4. ISSN: 2516-7073. DOI: 10.15761/PMCH.1000115. URL: <https://doi.org/10.15761/PMCH.1000115>.
- [141] Harold W T Mapoma and Xianjun Xie. “Basement and alluvial aquifers of Malawi: An overview of groundwater quality and policies”. In: *African Journal of Environmental Science and Technology* 8.3 (2014), pp. 190–202. ISSN: 1996-0786. DOI:

- 10.4314/ajest.v8i3.. URL:
<https://www.ajol.info/index.php/ajest/article/view/110403>.
- [142] Peter P. Mkandawire. “Groundwater resources of Malawi”. In: *isarm.org* (2004). URL:
<https://isarm.org/sites/default/files/resources/files/ManagingSharedAquiferResourcesinAfrica.pdf{\#}page=91>.
- [143] David H Eccles. “An outline of the physical limnology of Lake Malawi”. In: (1974). DOI:
10.4319/lo.1974.19.5.0730. URL: <https://aslopubs.onlinelibrary.wiley.com/doi/10.4319/lo.1974.19.5.0730>.
- [144] Ajay G. Bhave *et al.* “Lake Malawi’s threshold behaviour: A stakeholder-informed model to simulate sensitivity to climate change”. In: *Journal of Hydrology* 584 (2020). ISSN: 00221694. DOI: 10.1016/J.JHYDROL.2020.124671.
- [145] National Statistical Office and Icf Macro. “Malawi 2010 Demographic and Health Survey”. In: *NSO and ICF* (2011). URL:
http://www.nsomalawi.mw/images/stories/data{_}on{_}line/demography/MDHS2010/MDHS2010report.pdf.
- [146] S. A. Kayuni *et al.* “How can schistosome circulating antigen assays be best applied for diagnosing male genital schistosomiasis (MGS): an appraisal using exemplar MGS cases from a longitudinal cohort study among fishermen on the south shoreline of Lake Malawi”. In: *Parasitology* 146.14 (2019), pp. 1785–1795. ISSN: 0031-1820. DOI: 10.1017/S0031182019000969. URL: <https://www.cambridge.org/core/journals/parasitology/article/how-can-schistosome-circulating-antigen-assays-be-best-applied-for-diagnosing-male-genital-schistosomiasis-mgs-an-appraisal-using-exemplar-mgs-cases-from-a-longitudinal-cohort-study-among-fisher>.
- [147] GoM. *The Third National Communication of the republic of Malawi to the Conference of the Parties (COP) of the United Nations Framework Convention on Climate Change*. 2020. URL: <https://unfccc.int/documents/268340> (visited on 04/19/2023).
- [148] Fiona Allan *et al.* “Snail-Related Contributions from the Schistosomiasis Consortium for Operational Research and Evaluation Program Including Xenomonitoring, Focal Mollusciciding, Biological Control, and Modeling”. In: *The American Journal of Tropical Medicine and Hygiene* 103.1_Suppl (2020), pp. 66–79. ISSN: 0002-9637. DOI: 10.4269/AJTMH.19-0831. URL: <https://>

- //www.ajtmh.org/view/journals/tpmd/103/1{_}Suppl/article-p66.xml.
- [149] Charles H. King *et al.* “SCORE Studies on the Impact of Drug Treatment on Morbidity due to *Schistosoma mansoni* and *Schistosoma haematobium* Infection”. In: *The American Journal of Tropical Medicine and Hygiene* 103.1_Suppl (2020), pp. 30–35. ISSN: 0002-9637. DOI: 10.4269/AJTMH.19-0830. URL: https://www.ajtmh.org/view/journals/tpmd/103/1{_}Suppl/article-p30.xml.
- [150] Ross I. Maidment *et al.* “The 30 year TAMSAT African Rainfall Climatology And Time series (TARCAT) data set”. In: *Journal of Geophysical Research: Atmospheres* 119.18 (2014), pp. 10,619–10,644. ISSN: 2169-8996. DOI: 10.1002/2014JD021927. URL: <https://onlinelibrary.wiley.com/doi/full/10.1002/2014JD021927><https://onlinelibrary.wiley.com/doi/abs/10.1002/2014JD021927><https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2014JD021927>.
- [151] Ross I. Maidment *et al.* “A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa”. In: *Scientific Data* 4 (2017), p. 170063. ISSN: 20524463. DOI: 10.1038/SDATA.2017.63. URL: <https://pennstate.pure.elsevier.com/en/publications/a-new-long-term-daily-satellite-based-rainfall-dataset-for-operat>.
- [152] Elena Tarnavsky *et al.* “Extension of the TAMSAT Satellite-Based Rainfall Monitoring over Africa and from 1983 to Present”. In: *Journal of Applied Meteorology and Climatology* 53.12 (2014), pp. 2805–2822. ISSN: 15588432. DOI: 10.1175/JAMC-D-14-0016.1.
- [153] Z Wan, S Hook, and G Hulley. *MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1km Global SIN Grid V061 [Data set]*. 2021. URL: <https://lpdaac.usgs.gov/products/mod11a2v061/{\#}citation> (visited on 04/19/2023).
- [154] S Running *et al.* *MODIS/Terra Net Evapotranspiration Gap-Filled Yearly L4 Global 500m SIN Grid V061 [Data set]*. 2021. URL: <https://lpdaac.usgs.gov/products/mod16a3gfv061/> (visited on 04/19/2023).

- [155] K Didan. *MODIS/TERRA Vegetation Indices 16-Day L3 Global 1km SIN Grid V061*. 2021. URL: <https://lpdaac.usgs.gov/products/mod13a2v061/{\#}citation> (visited on 04/19/2023).
- [156] D.K Golon. *The Land Processes Distributed Active Archive Center (LP DAAC) — U.S. Geological Survey*. 2016. URL: <https://www.usgs.gov/publications/land-processes-distributed-active-archive-center-lp-daac> (visited on 04/19/2023).
- [157] J.A Dijkshoorn *et al.* *ISRIC Report 2016/01 Soil and Terrain database of the Republic of Malawi — ISRIC*. 2016. DOI: 10.17027/isric-wdcsoils.20160001. URL: <https://www.isric.org/documents/document-type/isric-report-201601-soil-and-terrain-database-republic-malawi> (visited on 04/18/2023).
- [158] Stan Development Team. *RStan: the R interface to Stan*. R package version 2.32.6. 2024. URL: <https://mc-stan.org/>.
- [159] *6.13 Gaussian Process Covariance Functions - Stan Functions reference*. URL: https://mc-stan.org/docs/2{_}29/functions-reference/gaussian-process-covariance-functions.html (visited on 05/05/2023).
- [160] Yanfeng Gong *et al.* “Three Gorges Dam: Differential determinants and spatial-temporal effects of the change of snail density”. In: (2023). DOI: 10.21203/RS.3.RS-2867328/V1. URL: <https://www.researchsquare.comhttps://www.researchsquare.com/article/rs-2867328/v1>.
- [161] Yingyan Zheng *et al.* “Environmental Determinants for Snail Density in Dongting Lake Region: An Ecological Study Incorporating Spatial Regression.” In: *The American Journal of Tropical Medicine and Hygiene* 107.6 (2022), pp. 1178–1184. ISSN: 0002-9637. DOI: 10.4269/AJTMH.22-0238. URL: <https://europepmc.org/article/med/36375461>.
- [162] Onyekachi Esther Nwoko, Tawanda Manyangadze, and Moses John Chimbari. “Spatial distribution, abundance, and infection rates of human schistosome-transmitting snails and related physicochemical parameters in KwaZulu-Natal (KZN) province, South Africa”. In: *Heliyon* 9.2 (2023), e12463. ISSN: 2405-8440. DOI: 10.1016/J.HELIYON.2022.E12463.

- [163] Sabiano Odero *et al.* “Distribution of *Biomphalaria* Snails in Associated Vegetations and Schistosome Infection Prevalence Along the Shores of Lake Victoria in Mbita, Kenya: A Cross-Sectional Study”. In: *The East African health research journal* 3.2 (2019), pp. 172–177. ISSN: 2520-5285. DOI: 10.24248/EAHRJ-D-19-00013. URL: <https://pubmed.ncbi.nlm.nih.gov/34308211/>.
- [164] Dana M. Woodhall *et al.* “Use of Geospatial Modeling to Predict *Schistosoma mansoni* Prevalence in Nyanza Province, Kenya”. In: *PLOS ONE* 8.8 (2013), e71635. ISSN: 1932-6203. DOI: 10.1371/JOURNAL.PONE.0071635. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0071635>.
- [165] Steady Chasimpha *et al.* “Patterns and risk factors for deaths from external causes in rural Malawi over 10 years: A prospective population-based study Health behavior, health promotion and society”. In: *BMC Public Health* 15.1 (2015), pp. 1–9. ISSN: 14712458. DOI: 10.1186/s12889-015-2323-Z/TABLES/4. URL: <https://link.springer.com/articles/10.1186/s12889-015-2323-z>
<https://link.springer.com/article/10.1186/s12889-015-2323-z>.
- [166] Matthew D. Tyler *et al.* “The epidemiology of drowning in low- and middle-income countries: a systematic review”. In: *BMC Public Health* 17.1 (2017), pp. 1–7. ISSN: 14712458. DOI: 10.1186/s12889-017-4239-2/FIGURES/3. URL: <https://bmcpublikealth.biomedcentral.com/articles/10.1186/s12889-017-4239-2>.
- [167] Anna Sofie Stensgaard *et al.* *Schistosomes, snails and climate change: Current trends and future expectations*. 2019. DOI: 10.1016/j.actatropica.2018.09.013.
- [168] Muriel Rabone *et al.* “Freshwater snails of biomedical importance in the Niger River Valley: Evidence of temporal and spatial patterns in abundance, distribution and infection with *Schistosoma* spp.” In: *Parasites and Vectors* 12.1 (2019), pp. 1–20. ISSN: 17563305. DOI: 10.1186/s13071-019-3745-8. URL: <https://link.springer.com/articles/10.1186/s13071-019-3745-8>
<https://link.springer.com/article/10.1186/s13071-019-3745-8>.
- [169] R. W. Tiner. “Ecology of Wetlands: Classification Systems”. In: *Encyclopedia of Inland Waters*. Academic Press, 2009, pp. 516–525. ISBN: 9780123706263. DOI: 10.1016/B978-012370626-3.00057-0.

- [170] Britannica. *Gleysol*. 2011. URL: <https://www.britannica.com/science/Gleysol>.
- [171] Alexandra V Kulinkina *et al.* “Improving spatial prediction of *Schistosoma haematobium* prevalence in southern Ghana through new remote sensors and local water access profiles”. In: *PLoS Neglected Tropical Diseases* 12.6 (2018). ISSN: 19352735. DOI: 10.1371/journal.pntd.0006517. URL: <https://doi.org/10.1371/journal.pntd.0006517>.
- [172] Laura Poggio *et al.* “SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty”. In: *SOIL* 7.1 (2021), pp. 217–240. ISSN: 2199398X. DOI: 10.5194/soil-7-217-2021.
- [173] J. C. Ernould *et al.* “Recent urban growth and urinary schistosomiasis in Niamey, Niger”. In: *Tropical Medicine and International Health* 5.6 (2000), pp. 431–437. ISSN: 13602276. DOI: 10.1046/j.1365-3156.2000.00577.x. URL: <https://europepmc.org/article/MED/10929143>.
- [174] C. C. Appleton and H. Madsen. “Human schistosomiasis in wetlands in southern Africa”. In: *Wetlands Ecology and Management* 20.3 (2012), pp. 253–269. ISSN: 09234861. DOI: 10.1007/s11273-012-9266-2/FIGURES/4. URL: <https://link.springer.com/article/10.1007/s11273-012-9266-2>.
- [175] Bahram Khazaei *et al.* “GLOBathy, the global lakes bathymetry dataset”. In: *Scientific Data* 9.1 (2022), pp. 1–10. ISSN: 20524463. DOI: 10.1038/s41597-022-01132-9. URL: <https://www.nature.com/articles/s41597-022-01132-9>.
- [176] Amber L. Reed *et al.* “Modelling the age-prevalence relationship in schistosomiasis: A secondary data analysis of school-aged-children in Mangochi District, Lake Malawi”. In: *Parasite Epidemiology and Control* 22 (2023), e00303. ISSN: 2405-6731. DOI: 10.1016/J.PAREPI.2023.E00303. URL: <https://linkinghub.elsevier.com/retrieve/pii/S240567312300020X>.
- [177] Fabian Reitzug, Julia Ledien, and Goylette F. Chami. “Associations of water contact frequency, duration, and activities with schistosome infection risk: A systematic review and meta-analysis”. In: *PLOS Neglected Tropical Diseases* 17.6 (2023), e0011377. ISSN: 1935-2735. DOI: 10.1371/JOURNAL.PNTD.0011377. URL: <https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0011377>.

- [178] Klodeta Kura *et al.* “What is the impact of acquired immunity on the transmission of schistosomiasis and the efficacy of current and planned mass drug administration programmes?” In: *PLOS Neglected Tropical Diseases* 15.12 (2021), e0009946. ISSN: 1935-2735. DOI: 10.1371/JOURNAL.PNTD.0009946. URL: <https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0009946>.
- [179] Jaspreet Toor *et al.* “The design of schistosomiasis monitoring and evaluation programmes: The importance of collecting adult data to inform treatment strategies for *Schistosoma mansoni*”. In: *PLOS Neglected Tropical Diseases* 12.10 (2018), e0006717. ISSN: 1935-2735. DOI: 10.1371/JOURNAL.PNTD.0006717. URL: <https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0006717>.
- [180] Charles H. King *et al.* *Transmission control for schistosomiasis - why it matters now*. 2006. DOI: 10.1016/j.pt.2006.09.006.
- [181] Olimpia Lamberti *et al.* “*Schistosoma mansoni* infection risk for school-aged children clusters within households and is modified by distance to freshwater bodies”. In: *PLOS ONE* 16.11 (2021), e0258915. ISSN: 1932-6203. DOI: 10.1371/JOURNAL.PONE.0258915. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0258915>.
- [182] Estelle Mezajou Mewamba *et al.* “Fine-scale mapping of *Schistosoma mansoni* infections and infection intensities in sub-districts of Makenene in the Centre region of Cameroon”. In: *PLoS Neglected Tropical Diseases* 16.10 (2022), e0010852. ISSN: 19352735. DOI: 10.1371/journal.pntd.0010852. URL: <https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0010852>.
- [183] Mark A. Deka. “Predictive Risk Mapping of Schistosomiasis in Madagascar Using Ecological Niche Modeling and Precision Mapping”. In: *Tropical Medicine and Infectious Disease* 7.2 (2022), p. 15. ISSN: 24146366. DOI: 10.3390/tropicalmed7020015. URL: <https://www.mdpi.com/2414-6366/7/2/15/htm><https://www.mdpi.com/2414-6366/7/2/15>.
- [184] Louis Albert Tchuem Tchuente *et al.* “Precision mapping: An innovative tool and way forward to shrink the map, better target interventions, and accelerate toward the elimination of schistosomiasis”. In: *PLoS Neglected Tropical Diseases* 12.8 (2018), e0006563. ISSN: 19352735. DOI: 10.1371/journal.pntd.0006563. URL: <https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0006563>.

- [185] Timothy Kamara, Michael Byamukama, and Martin Karuhanga. “Modelling the Role of Treatment, Public Health Education, and Chemical Control Strategies on Transmission Dynamics of Schistosomiasis”. In: *Journal of Mathematics* 2022 (2022). ISSN: 23144785. DOI: 10.1155/2022/2094979.
- [186] Larissa C. Anderson, Eric S. Loker, and Helen J. Wearing. “Modeling schistosomiasis transmission: the importance of snail population structure”. In: *Parasites and Vectors* 14.1 (2021), pp. 1–14. ISSN: 17563305. DOI: 10.1186/s13071-021-04587-8. URL: <https://parasitesandvectors.biomedcentral.com/articles/10.1186/s13071-021-04587-8>.
- [187] Chinwendu E Madubueze, Reuben I Gweryina, and Agatha Abokwara Joseph Sarwuan. “Mathematical Analysis and Optimal Control of Schistosomiasis Transmission Model”. In: *Biomath Communications* 9.1 (2022), pp. 2203071–2203071. ISSN: 2367-5241. DOI: 10.11145/BMC.2022.03.071. URL: <https://biomath.math.bas.bg/biomath/index.php/conference/article/view/bmc.2022.03.071>.
- [188] Al-Harbi M.H. Makaula P. Condemine C. Hesketh J. Archer J. Jones S. Kayuni S.A. Musaya J. Stanton M.C. Reed A.L. and J.R. Stothard. “A geospatial analysis of local intermediate snail host distributions provides insight into schistosomiasis risk within under-sampled areas of southern Lake Malawi”. In: 17 (2024), p. 272.
- [189] Richard H. Byrd *et al.* “A Limited Memory Algorithm for Bound Constrained Optimization”. In: *SIAM Journal on Scientific Computing* 16.5 (1995), pp. 1190–1208. ISSN: 1064-8275. DOI: 10.1137/0916069.
- [190] Marilyn Ronoh *et al.* “Modelling the spread of schistosomiasis in humans with environmental transmission”. In: *Applied Mathematical Modelling* 95 (2021), pp. 159–175. ISSN: 0307904X. DOI: 10.1016/j.apm.2021.01.046.
- [191] Benjamin S. Collyer *et al.* “How important is the spatial movement of people in attempts to eliminate the transmission of human helminth infections by mass drug administration?” In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 378.1887 (2023). ISSN: 14712970. DOI: 10.1098/rstb.2022.0273. URL: <https://royalsocietypublishing.org/doi/10.1098/rstb.2022.0273>.
- [192] Carlos A. Torres-Vitolasid *et al.* “Behaviour change interventions for the control and elimination of schistosomiasis: A systematic review of evidence from low-and middle-income countries”. In: *PLoS Neglected Tropical Diseases* 17.5 (2023), e0011315. ISSN: 19352735. DOI: 10.1371/journal.pntd.0011315. URL:

<https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0011315>.

- [193] Kalinda C. Breuer L. Tabo Z. and C. Albrecht. “Adapting strategies for effective schistosomiasis prevention: A mathematical modeling approach”. In: 11 (12 2023), p. 2609.

Appendices

Appendix A

Ethics



Applicant: Amber Reed
Supervisor: Chris Jewell
Department: Lancaster Medical School
FHMREC Reference: FHMREC19049

01 May 2020

Dear Amber

Re: Understanding the prevalence/intensity of *Schistosoma mansoni*, *Schistosoma haematobium* and co-infection in school-aged children in Mangochi region and how the distribution of snail habits along the shoreline of Lake Malawi, impacts disease transmission to these school-aged children.

Thank you for submitting your research ethics application for the above project for review by the **Faculty of Health and Medicine Research Ethics Committee (FHMREC)**. The application was recommended for approval by FHMREC, and on behalf of the Chair of the Committee, I can confirm that approval has been granted for this research project.

As principal investigator your responsibilities include:

- ensuring that (where applicable) all the necessary legal and regulatory requirements in order to conduct the research are met, and the necessary licenses and approvals have been obtained;
- reporting any ethics-related issues that occur during the course of the research or arising from the research to the Research Ethics Officer at the email address below (e.g. unforeseen ethical issues, complaints about the conduct of the research, adverse reactions such as extreme distress);
- submitting details of proposed substantive amendments to the protocol to the Research Ethics Officer for approval.

Please contact me if you have any queries or require further information.

Tel:- 01542 593987

Email:- fhmresearchsupport@lancaster.ac.uk

Yours sincerely,

A handwritten signature in black ink that reads "R. E. Case".

Becky Case
Research Ethics Officer, Secretary to FHMREC.

Figure A.1: Lancaster University Ethics approval 2020

Appendix B

Chapter 2

B.1 Raw data

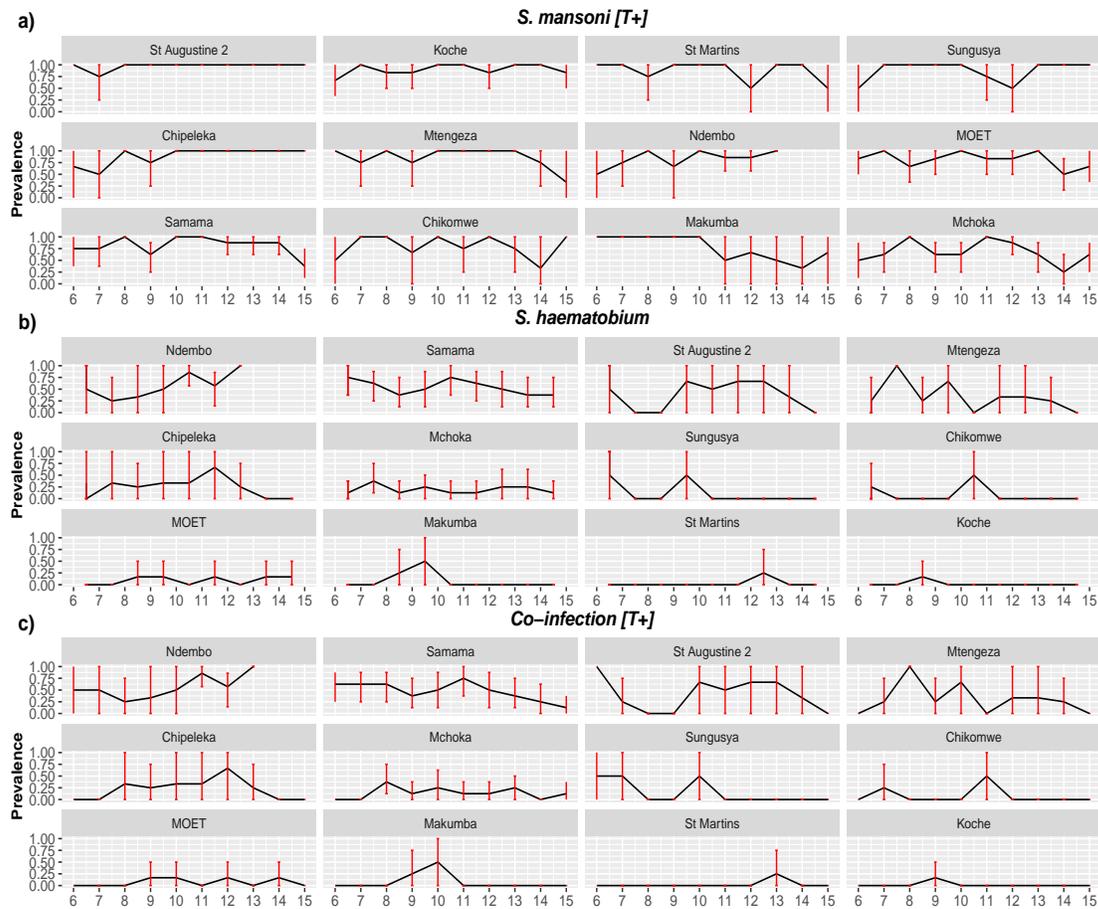


Figure B.1: Raw data plot showing the age of the children vs school prevalence for a) *S. mansoni* [T+], b) co-infection [T+] and c) *S. haematobium*. Order of schools on heatmap was by highest to lowest prevalence.

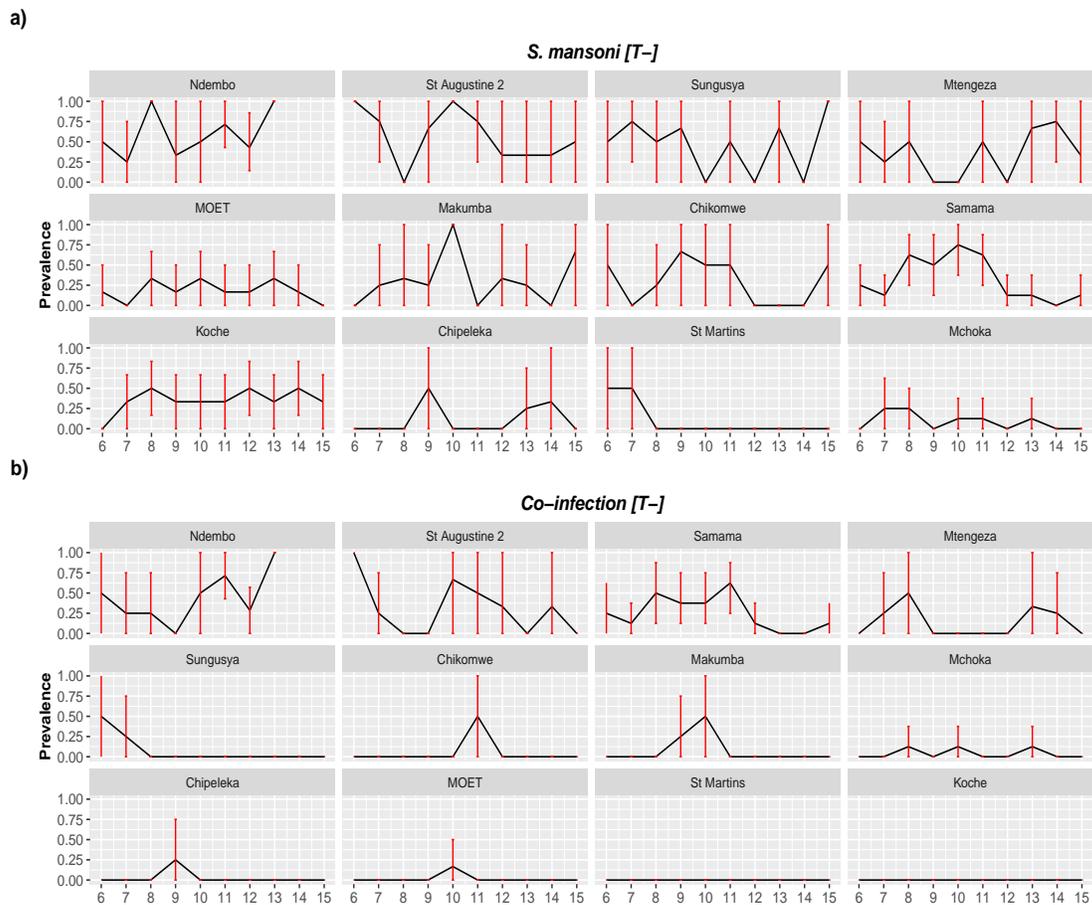


Figure B.2: Raw data plot showing the age of the children vs school prevalence for a) *S. mansoni* [T-], b) co-infection [T-] and c) Order of schools on heatmap was by highest to lowest prevalence.

B.2 Prevalence heatmaps

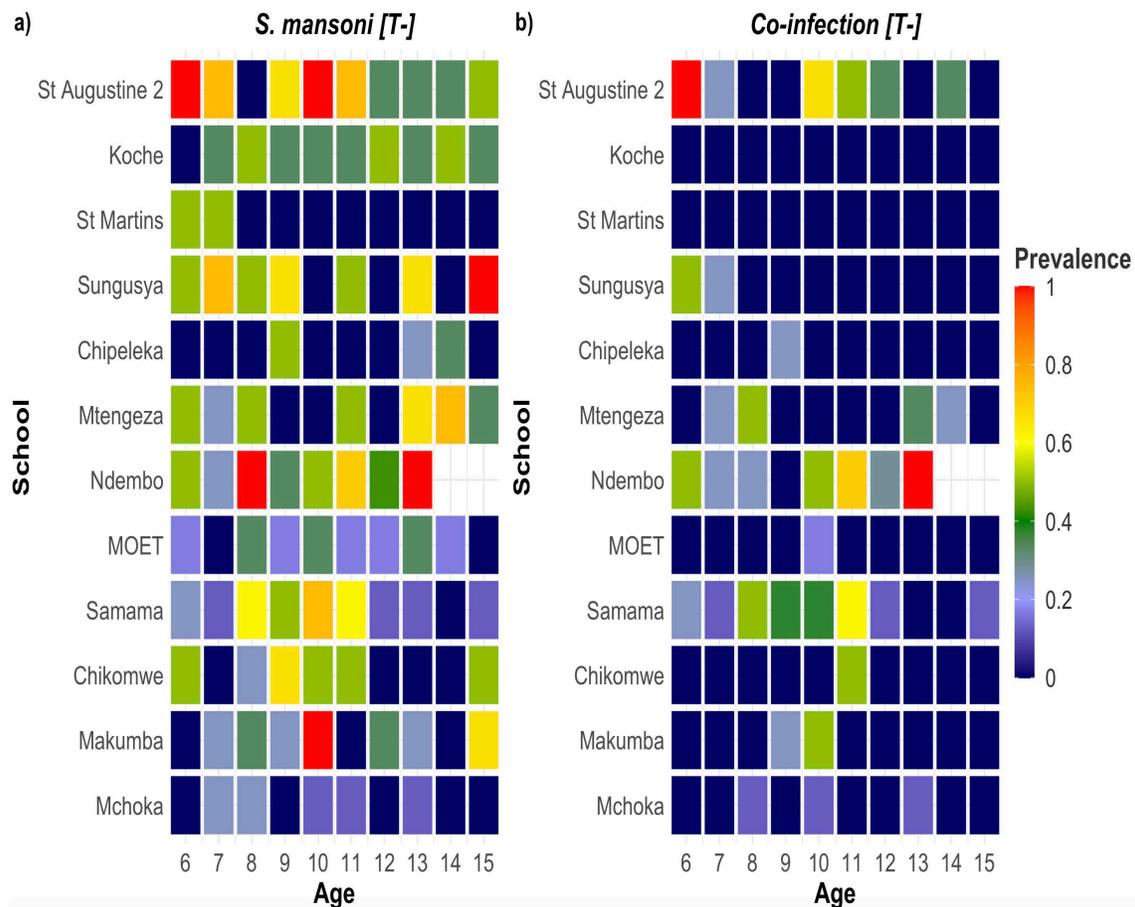


Figure B.3: Heatmap showing the age of the children vs school prevalence for a) *S. mansoni* [T-] and b) co-infection [T-]. Order of schools on heatmap was by highest to lowest prevalence.

B.3 Generalised additive models

Table B.1: Summary of prevalence of *S. mansoni* [T-] and co-infection [T-]

	<i>S. mansoni</i> [T+]		<i>S. haematobium</i>		Co-infection [T+]	
	95% CI		95% CI		95% CI	
Smooth term (p-value)						
Age	8.45 × 10 ⁻⁴ ***		0.114		7.81 × 10 ⁻³ **	
Factor (estimated coefficient)						
School						
Samama	0.767*	(0.0206, 1.51)	1.75***	(1.03, 2.47)	1.81***	(1.01, 2.59)
MOET	0.796.	(-0.0251, 1.62)	-0.940.	(-0.202, 0.138)	-0.815	(-2.02, 0.391)
Koche	1.52**	(0.540, 2.50)	-2.62*	(-4.68, -0.567)	-2.26*	(-4.34, -0.180)
St Augustine 2	2.63*	(0.576, 4.69)	1.19*	(0.271, 2.12)	1.43**	(0.440, 2.41)
Ndembo	0.621	(-0.465, 1.71)	1.74***	(0.807, 2.67)	1.89***	(0.920, 2.86)
Sungusya	1.48*	(0.186, 2.78)	-0.143	(-1.26, 0.975)	-4.33e-2	(-1.28, 1.20)
St Martins	1.47*	(0.169, 2.76)	-1.92.	(-3.99, 0.155)	-1.57	(-3.67, 0.532)
Chikomwe	0.634	(-0.393, 1.66)	-0.745	(-2.07, 0.578)	-0.387	(-1.75, 0.97)
Chipeleka	1.47*	(0.168, 2.76)	0.445	(-0.546, 1.43)	0.628	(-0.444, 1.70)
Makumba	0.484	(-0.503, 1.47)	-1.16	(-2.70, 0.386)	-0.788	(-2.37, 0.791)
Mtengza	0.928.	(-0.160, 2.02)	0.656	(-0.314, 1.63)	1.05*	(0.0228, 2.07)
Mchoka	0	-	0	-	0	-

*Significance $p < 0.05$, **Significance $p < 0.01$, ***Significance $p < 0.001$, . Significance at $p < 0.1$

Table B.2: GAM with smooth term age adjusted for school

	<i>S. mansoni</i> [T-]		Co-infection [T-]	
	95% CI		95% CI	
Smooth term (p-value)				
Age	0.111		32.0 × 10 ⁻² *	
Factor (estimated coefficient)				
School				
Samama	1.63***	(0.718, 2.54)	2.21**	(0.939, 3.48)
MOET	0.856.	(-0.162, 1.87)	-0.838	(-3.13, 1.46)
Koche	1.74***	(0.797, 2.68)	-28.3	(-178, 178)
St Augustine 2	2.61***	(1.55, 3.68)	2.43***	(1.01, 3.84)
Ndembo	2.53***	(1.47, 3.60)	2.65***	(1.27, 4.03)
Sungusya	2.21***	(1.16, 3.28)	0.618	(-1.24, 2.47)
St Martins	0.132	(-1.29, 1.56)	-28.3	(-254, 254)
Chikomwe	1.33	(0.201, 2.45)	0.611	(-1.24, 2.46)
Chikomwe	0.461	(-0.851, 1.77)	-0.146	(-2.46, 2.17)
Makumba	1.51**	(0.408, 2.62)	0.681	(-1.17, 2.54)
Mtengza	1.67**	(0.580, 2.76)	1.45.	(-0.133, 3.03)
Mchoka	0	0	0	0

*Significance $p < 0.05$, **Significance $p < 0.01$, ***Significance $p < 0.001$, . Significance at $p < 0.1$

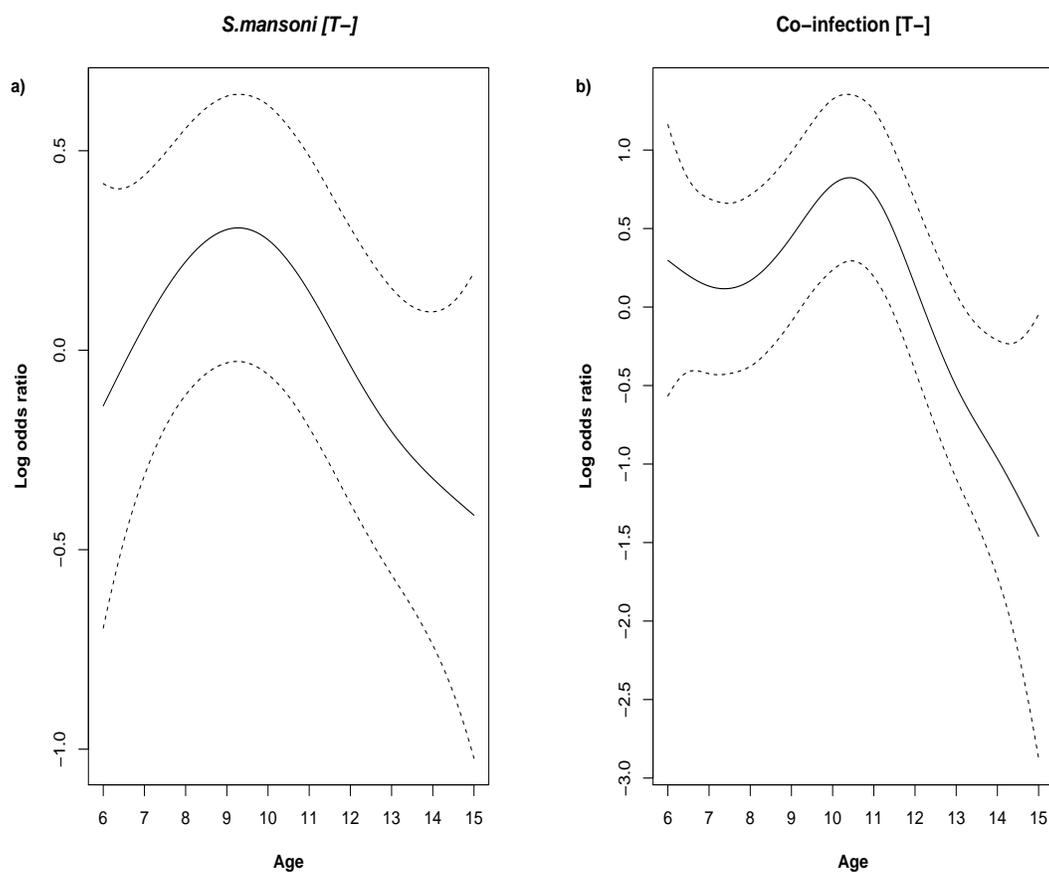


Figure B.4: Smooth age term plot for the GAM of *Schistosoma* association with age of SAC for a) *S. mansoni* [T-] and b) co-infection [T-].

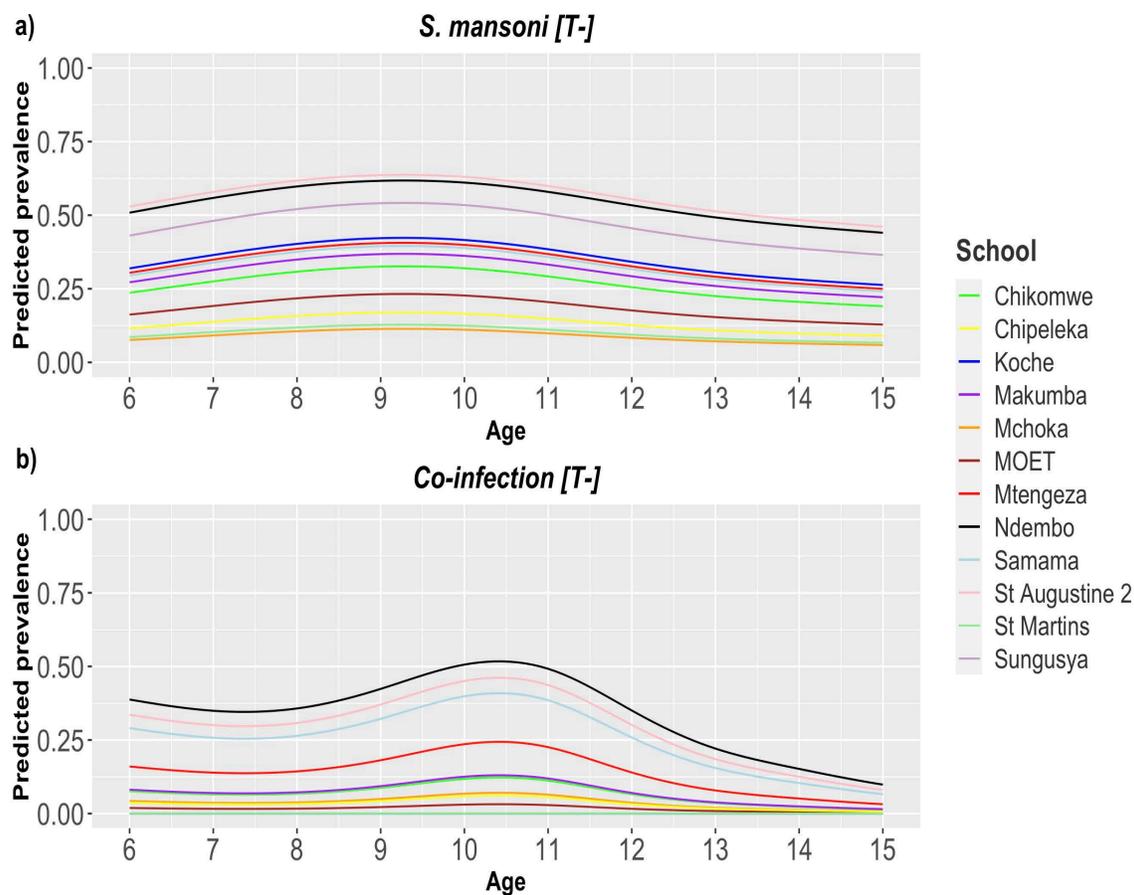


Figure B.5: Gam of *Schistosoma* association with age of SAC for each school. Invlogit of predicted fitted values versus age, a) *S. mansoni* [T-] and b) co-infection. Light Green: Chikomwe, Yellow: Chiipeleka, Dark Blue: Koche, Purple: Makumba, Orange: Mchoka, Brown: Moet Red: Mtengeza Black: Ndem-bo, Light Blue: Samama, Pink: St Augustine 2, Dark Green: St Martins, Mauve: Sungusya

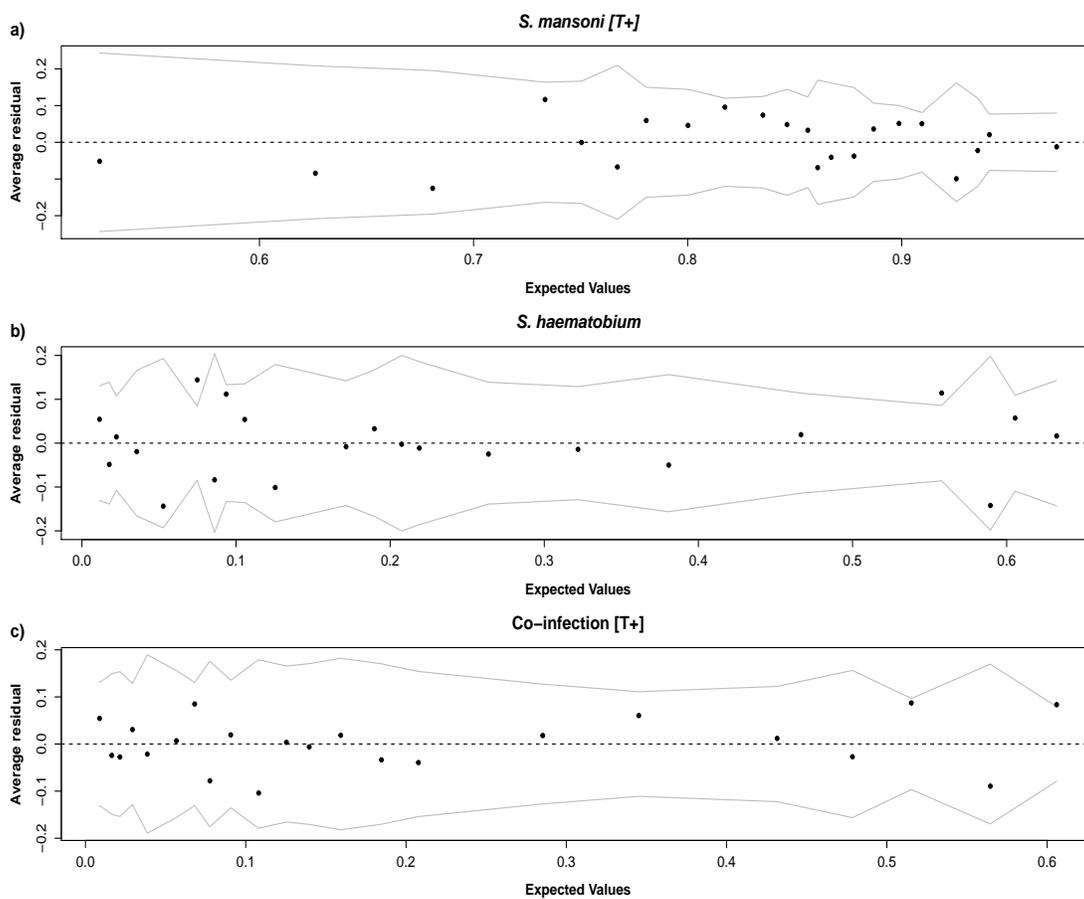


Figure B.6: Probability of the being positive with *Schistosoma* versus the average residuals, a) *S. mansoni* [T+], b) *S. haematobium* and c) co-infection [T+].

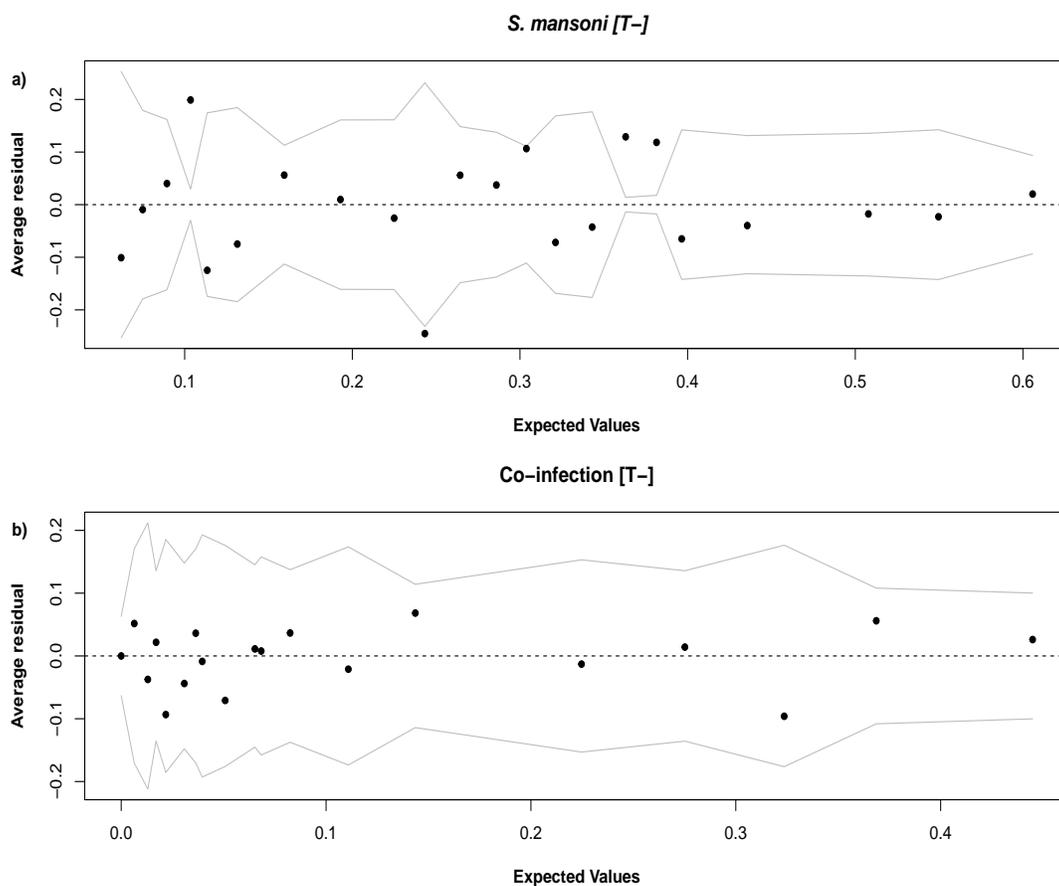


Figure B.7: Probability of the being positive with *Schistosoma* versus the average residuals a) *S. mansoni* [T-] and b) co-infection [T-].

B.4 Water contact

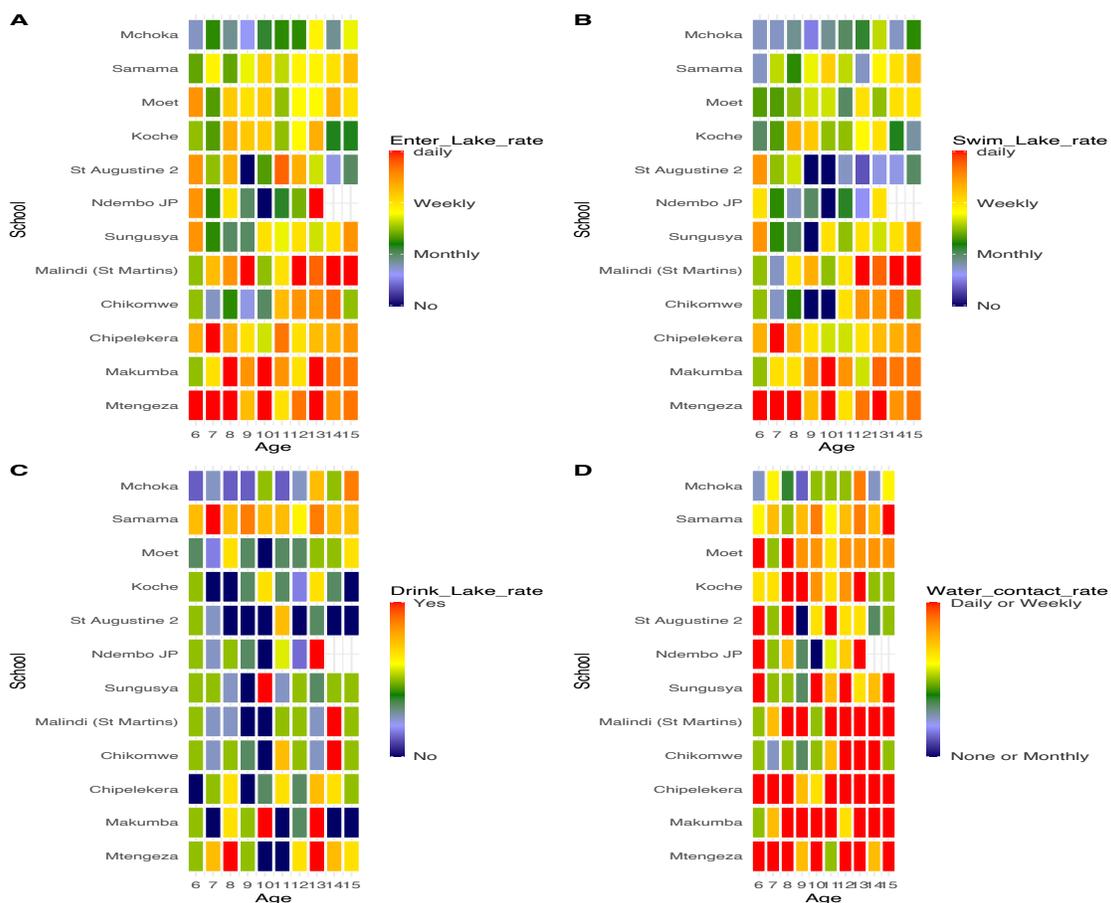


Figure B.8: Water contact rate calculated using the questionnaire answers from the SAC collected in the primary study. Age of SAC versus their school A) Enter Lake rate B) Swim Lake rate C) Drink late rate D) Water Contact rate.

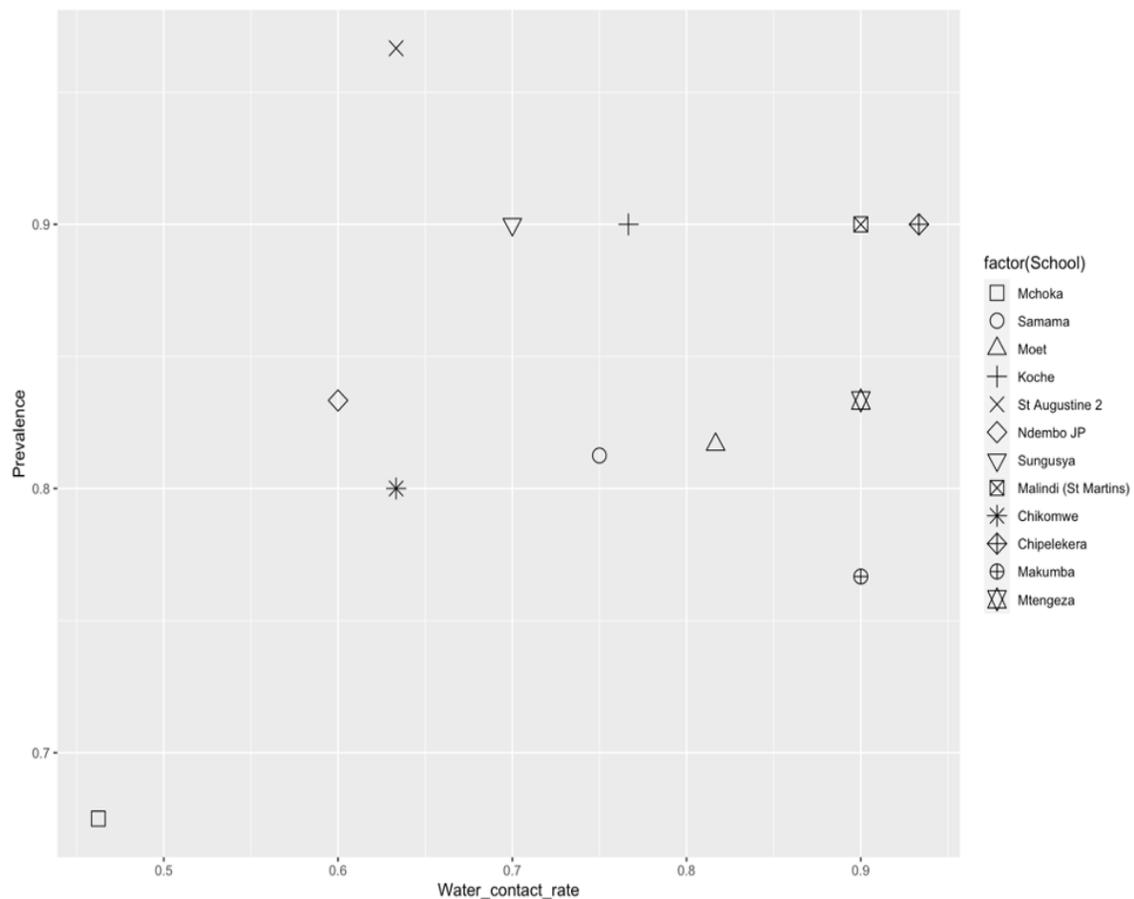


Figure B.9: Water contact rate versus prevalence of *Schistosoma* infection for each school. Legend: Different symbols represent each school

Appendix C

Chapter 3

C.1 Construction of 200 prediction points

C.2 1D extracted environmental data

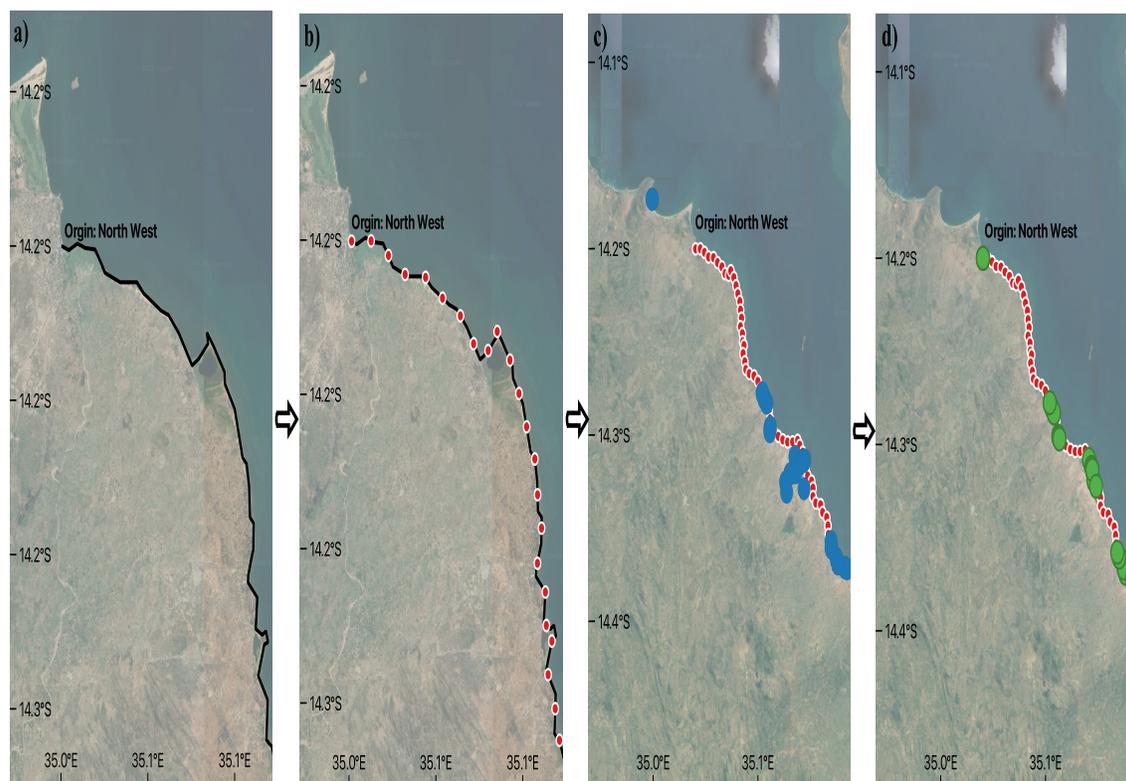
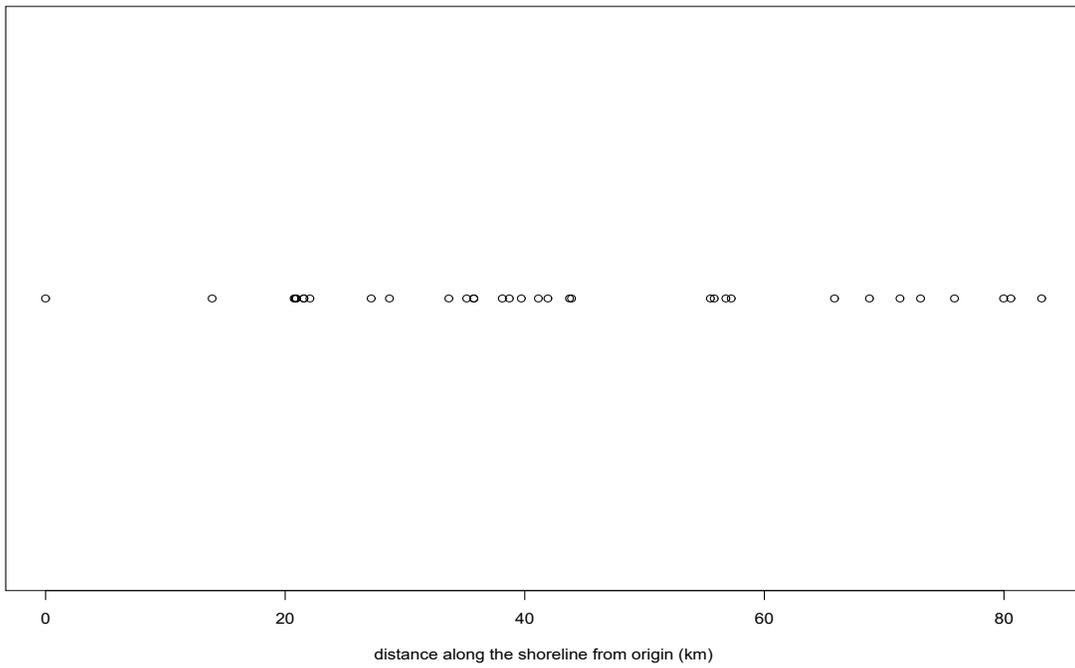


Figure C.1: Flow diagram showing the stages for constructing the 200 predictions along the shoreline. a) A 2-D linestring was drawn by hand following the shoreline as shown by Google Satellite imagery, b) the linestring was re-sampled to 4000 equally spaced vertices and resampled them to 200 equally intervals (red dots), c) observed sampling site locations (blue dots), d) each observed sampling site location was snapped to its nearest vertex (green dots).

(a) *Biomphalaria* sp.



(b) *Bulinus* spp.

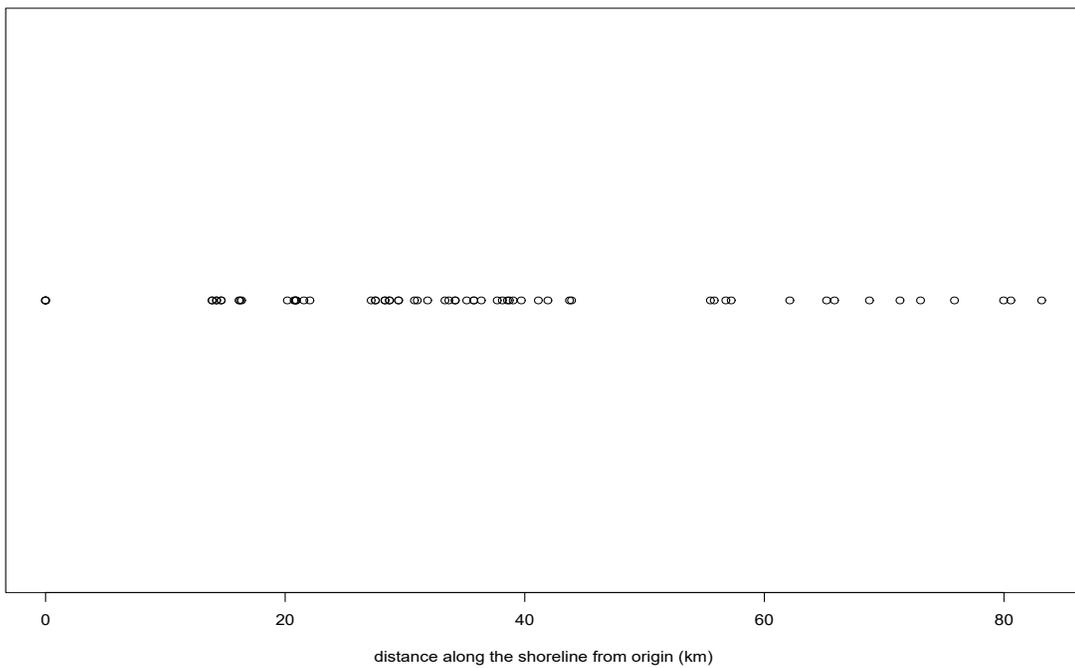


Figure C.2: The distance along the line from the origin (northwest-most vertex) to each of the snapper observed sampling site locations for each species a) *Biomphalaria* sp. b) *Bulinus* spp..

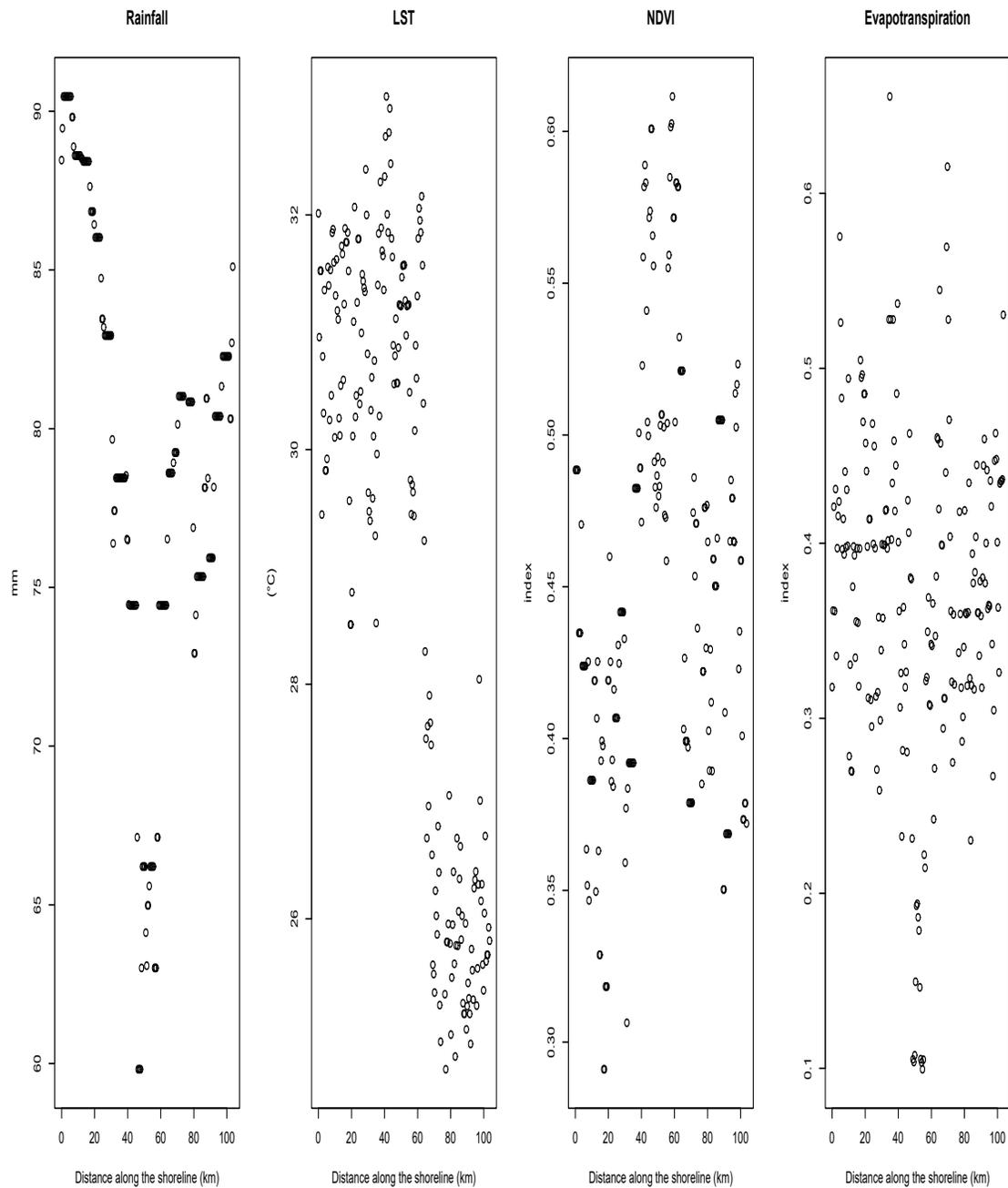


Figure C.3: Scatterplot of environmental data extracted versus distance along shoreline (km) for prediction points

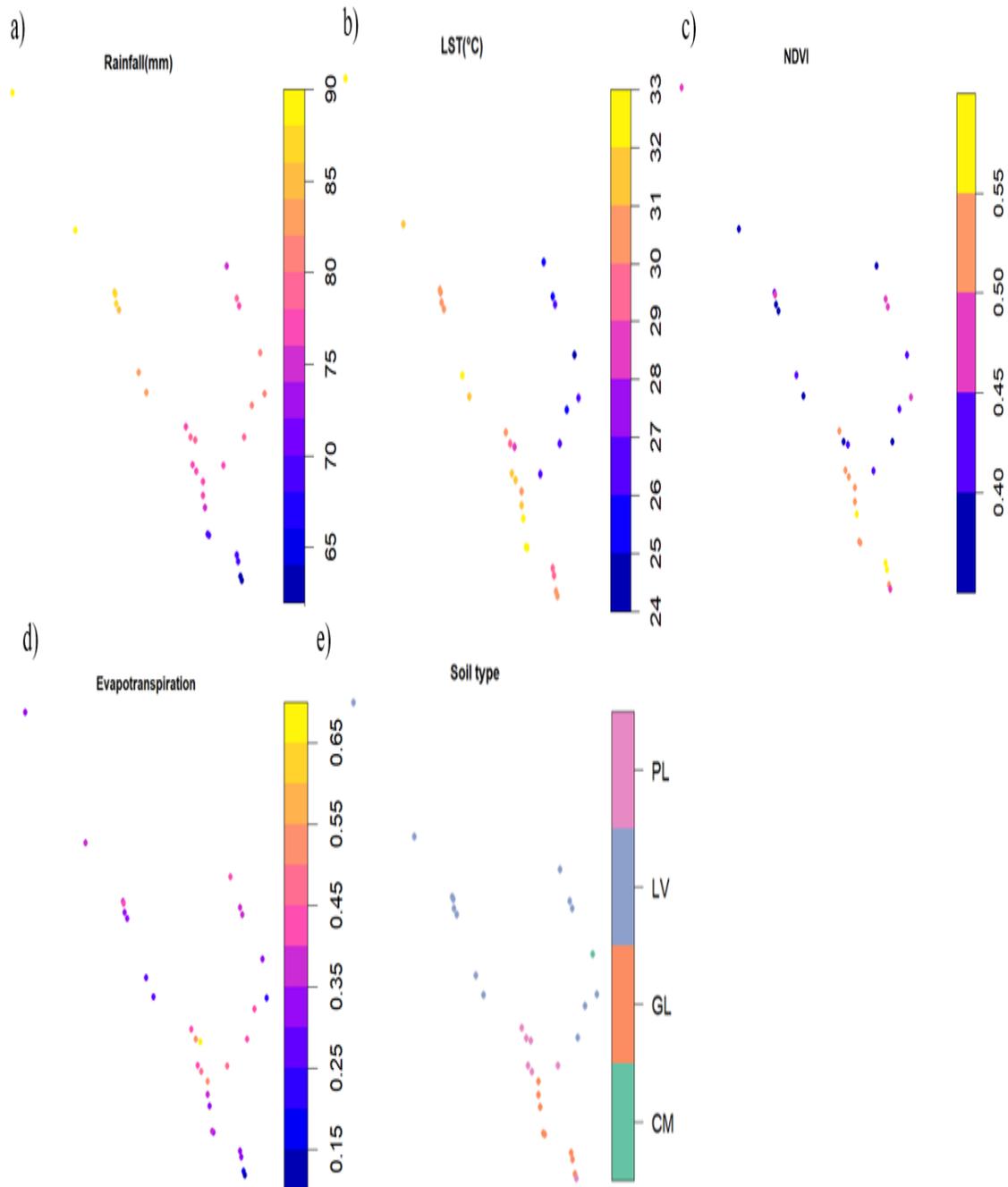


Figure C.4: *Biomphalaria* sp. observed sample points extracted environmental data a) Rainfall (mm) b) LST (°C) c) NDVI (index) d) Evapotranspiration (index) e) Soil type

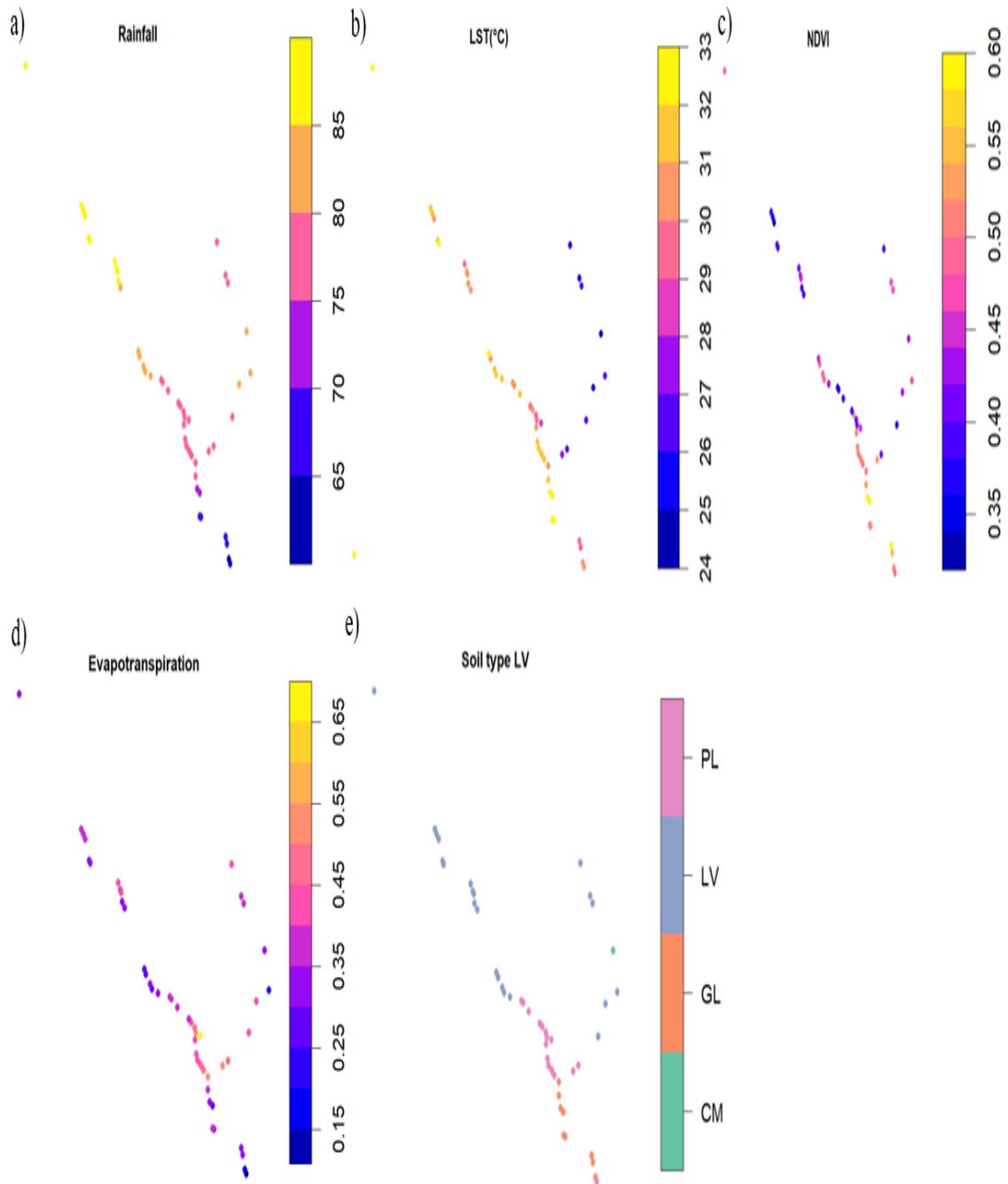
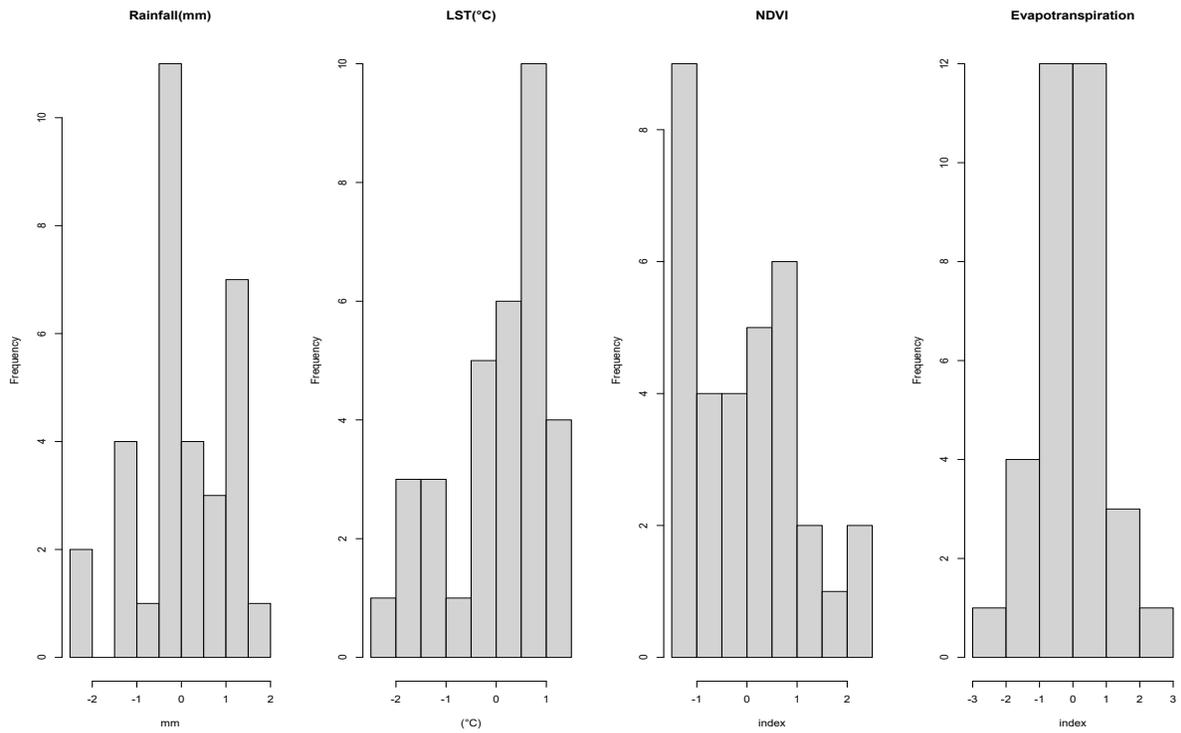


Figure C.5: *Bulinus* spp. observed sample points extracted environmental data. a) Rainfall (mm) b) LST (°C) c) NDVI (index) d) Evapotranspiration (index) e) Soil type

C.3 Center and scaling environment data

(a) *Biomphalaria* sp.



(b) *Bulinus* spp.

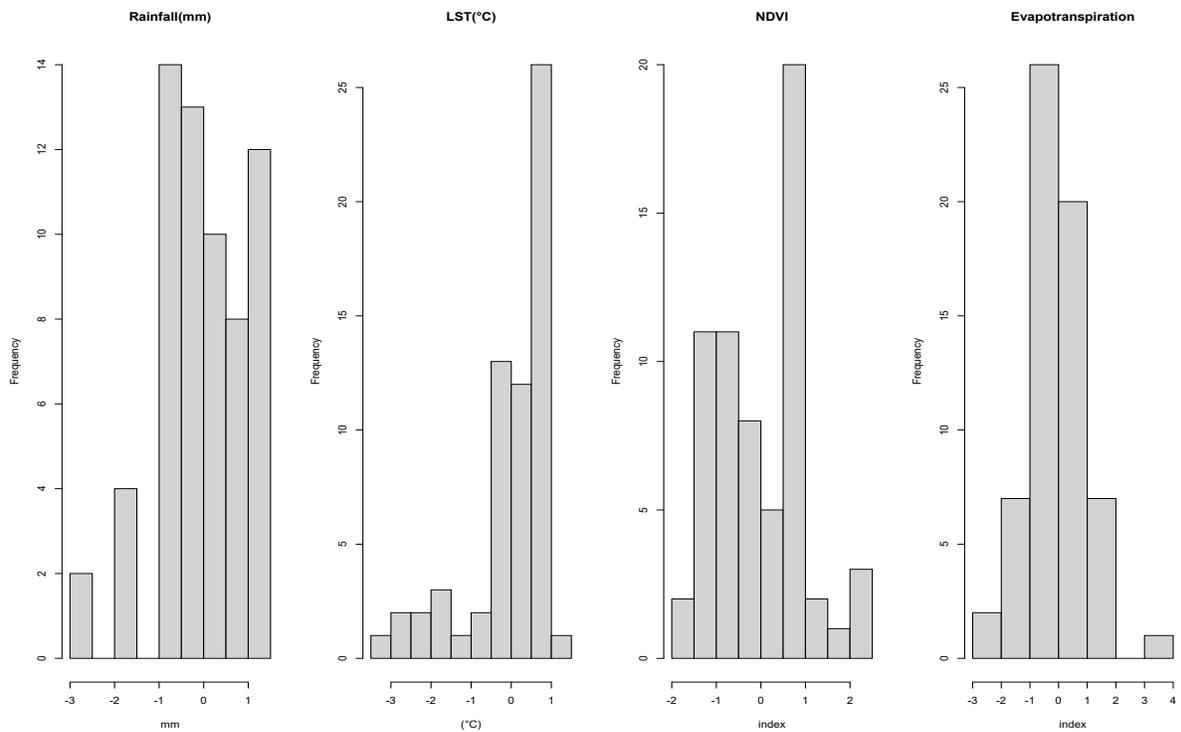
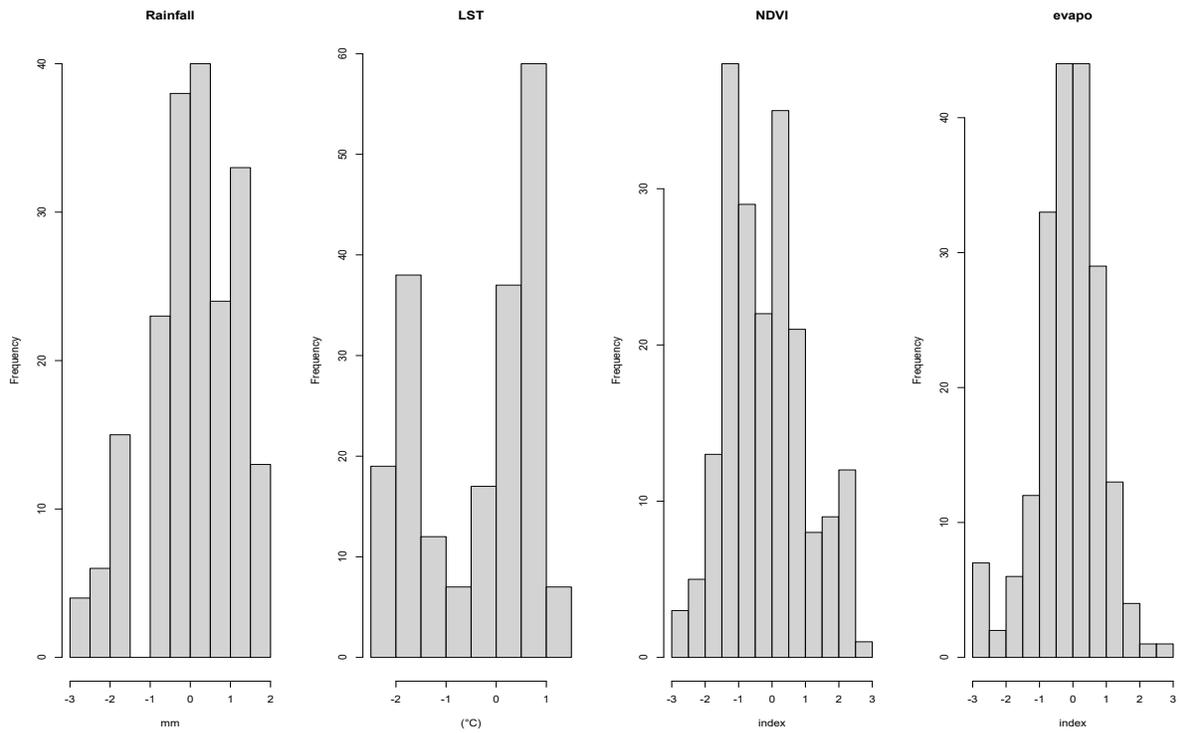


Figure C.6: Observed points values for each covariate centred and scaled a) *Biomphalaria* sp. b) *Bulinus* spp.

(a) *Biomphalaria* sp.



(b) *Bulinus* spp.

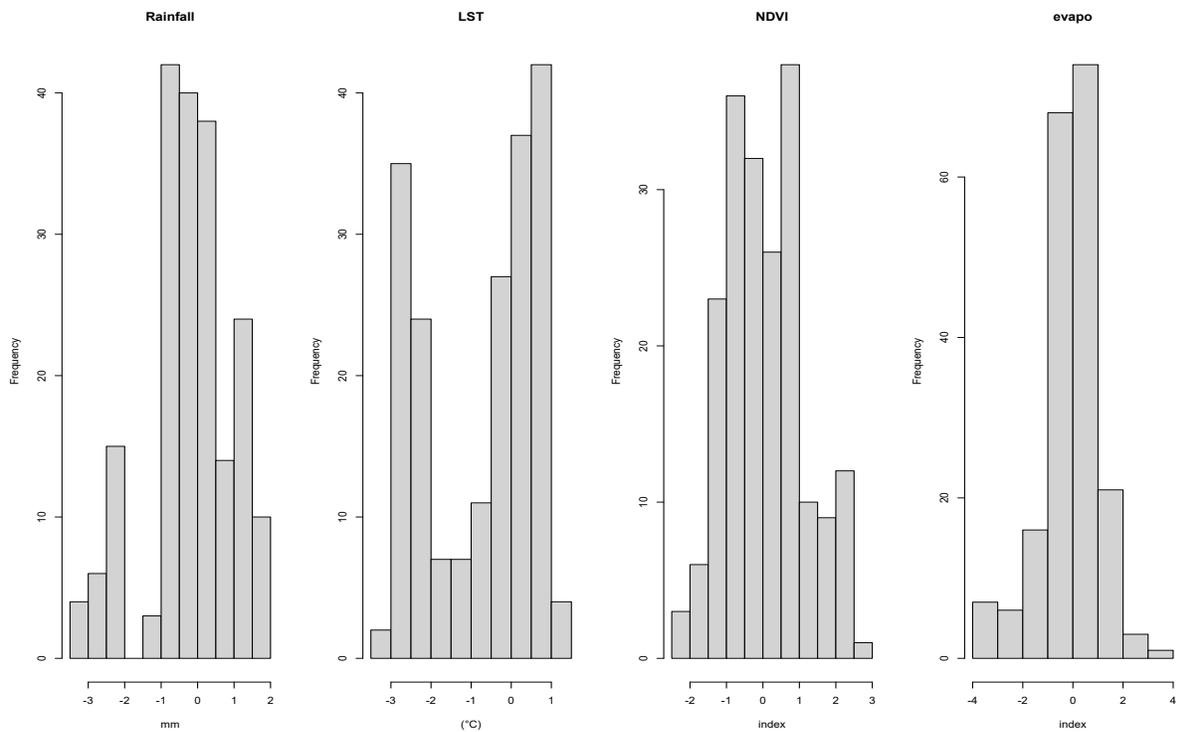
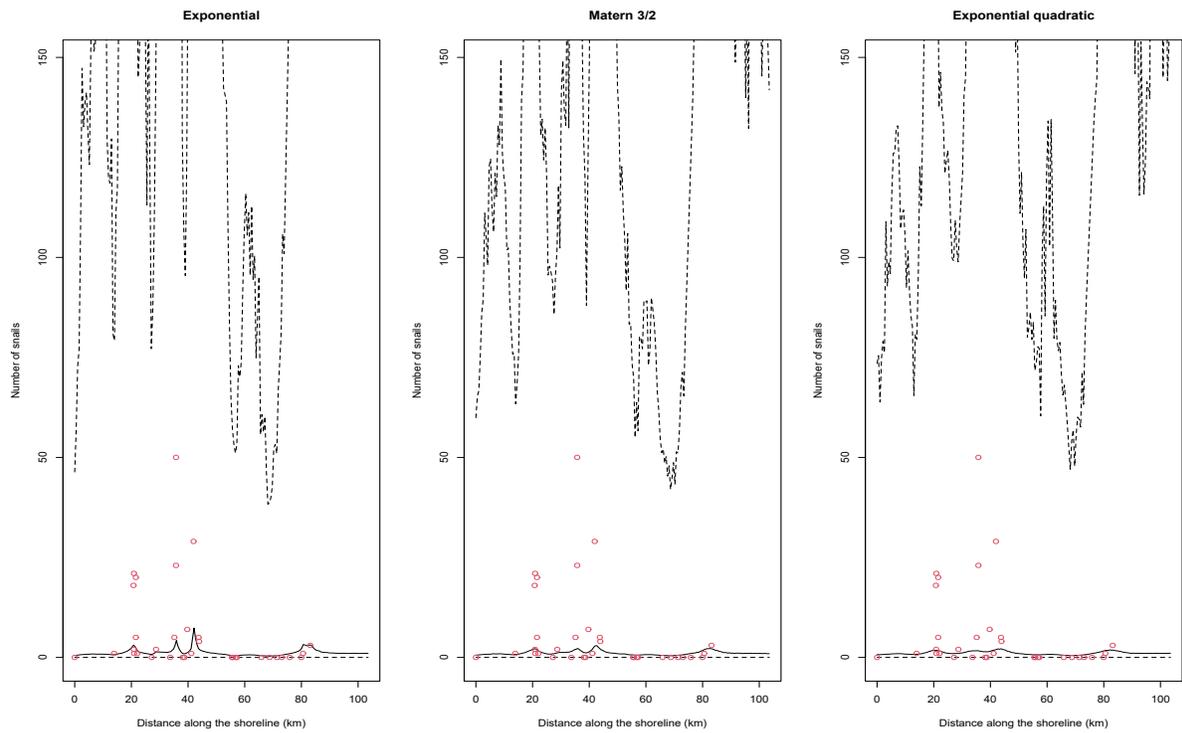


Figure C.7: Histograms of prediction points values for each covariate centred and scaled. a) *Biomphalaria* sp. b) *Bulinus* spp.

C.4 Covariance functions

(a) *Biomphalaria* sp.



(b) *Bulinus* spp.

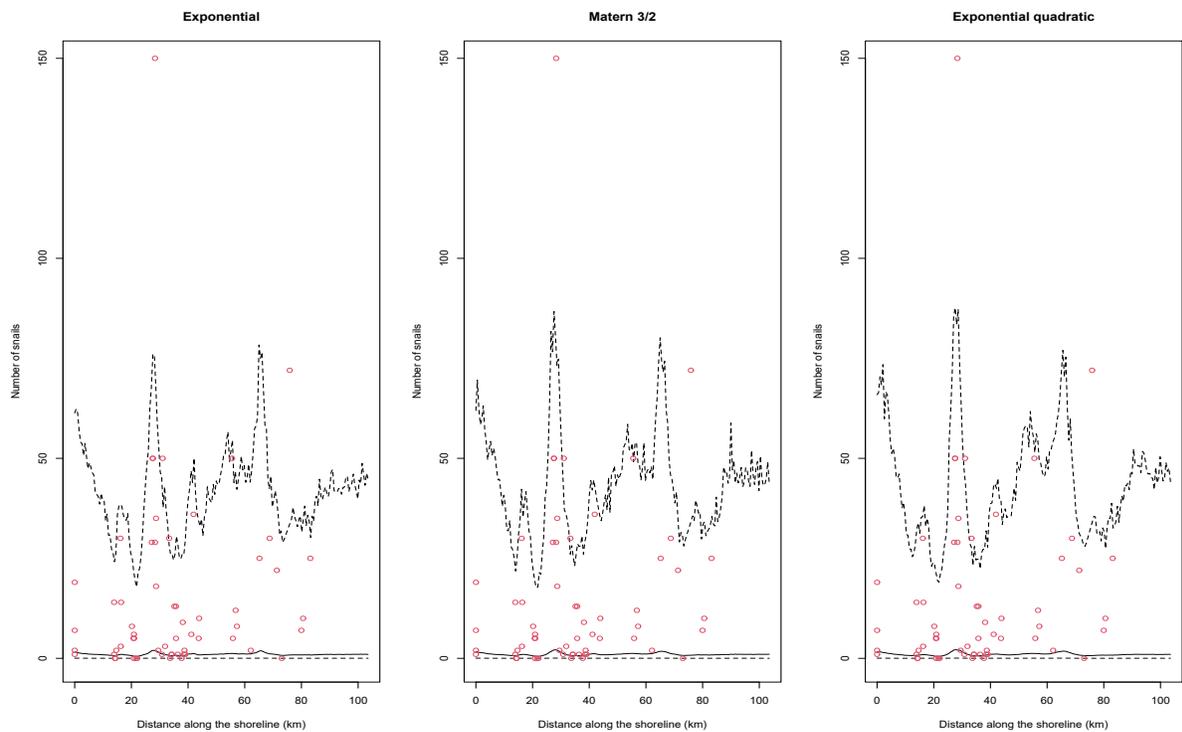
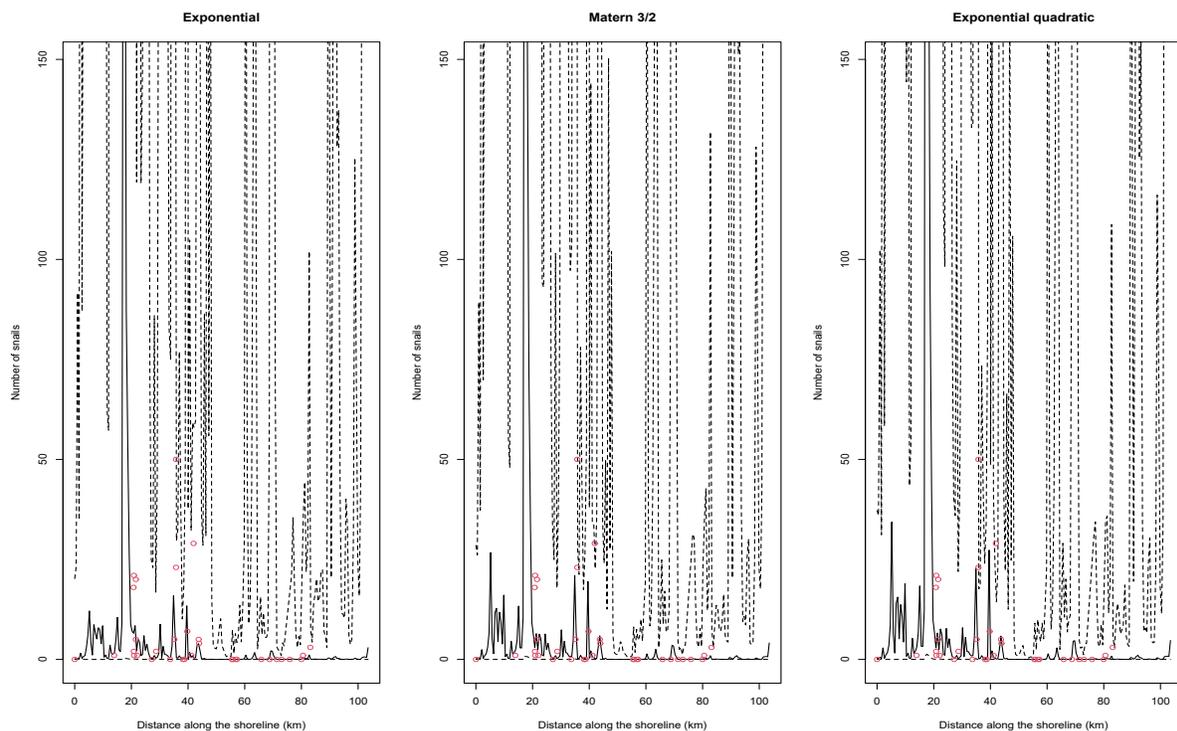


Figure C.8: Comparison of covariance functions for above snails mean abundance for a) *Biomphalaria* sp. b) *Bulinus* spp. against distance along (km).

(a) *Biomphalaria* sp.



(b) *Bulinus* spp.

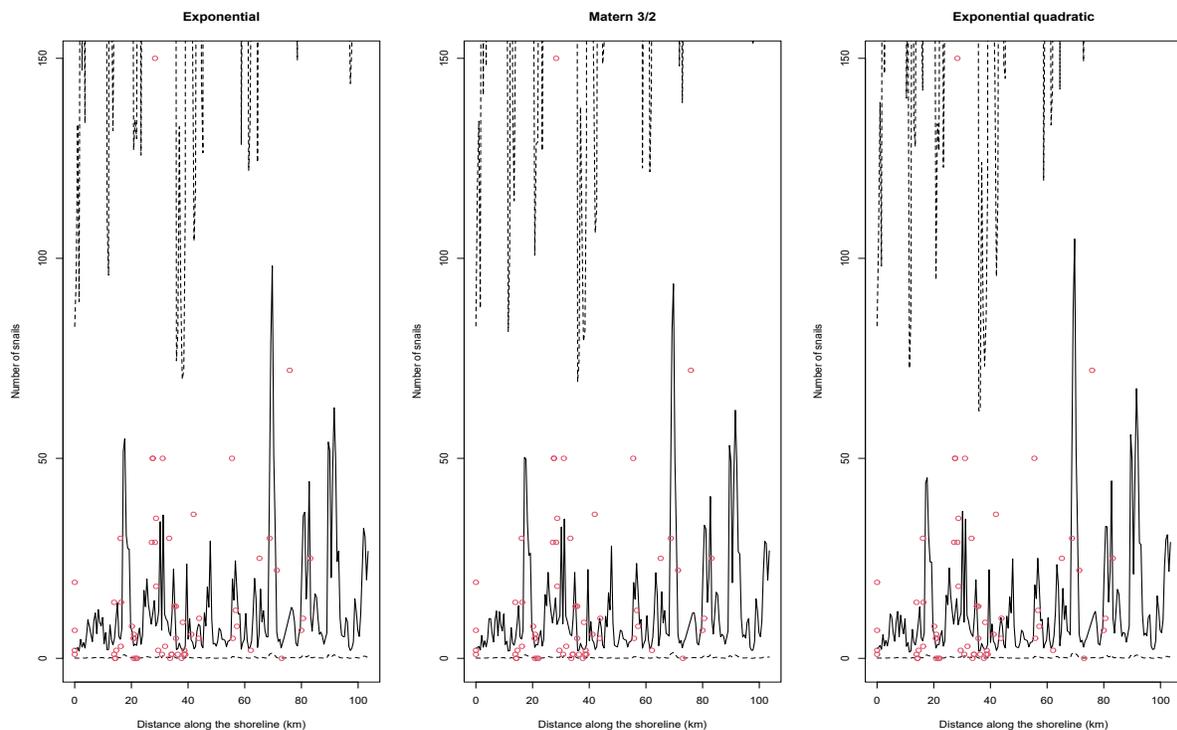


Figure C.9: Comparison of covariance functions for number of snails predicted for a) *Biomphalaria* sp. b) *Bulinus* spp. against distance along (km).

C.5 Convergence

C.6 Priors and Posteriors

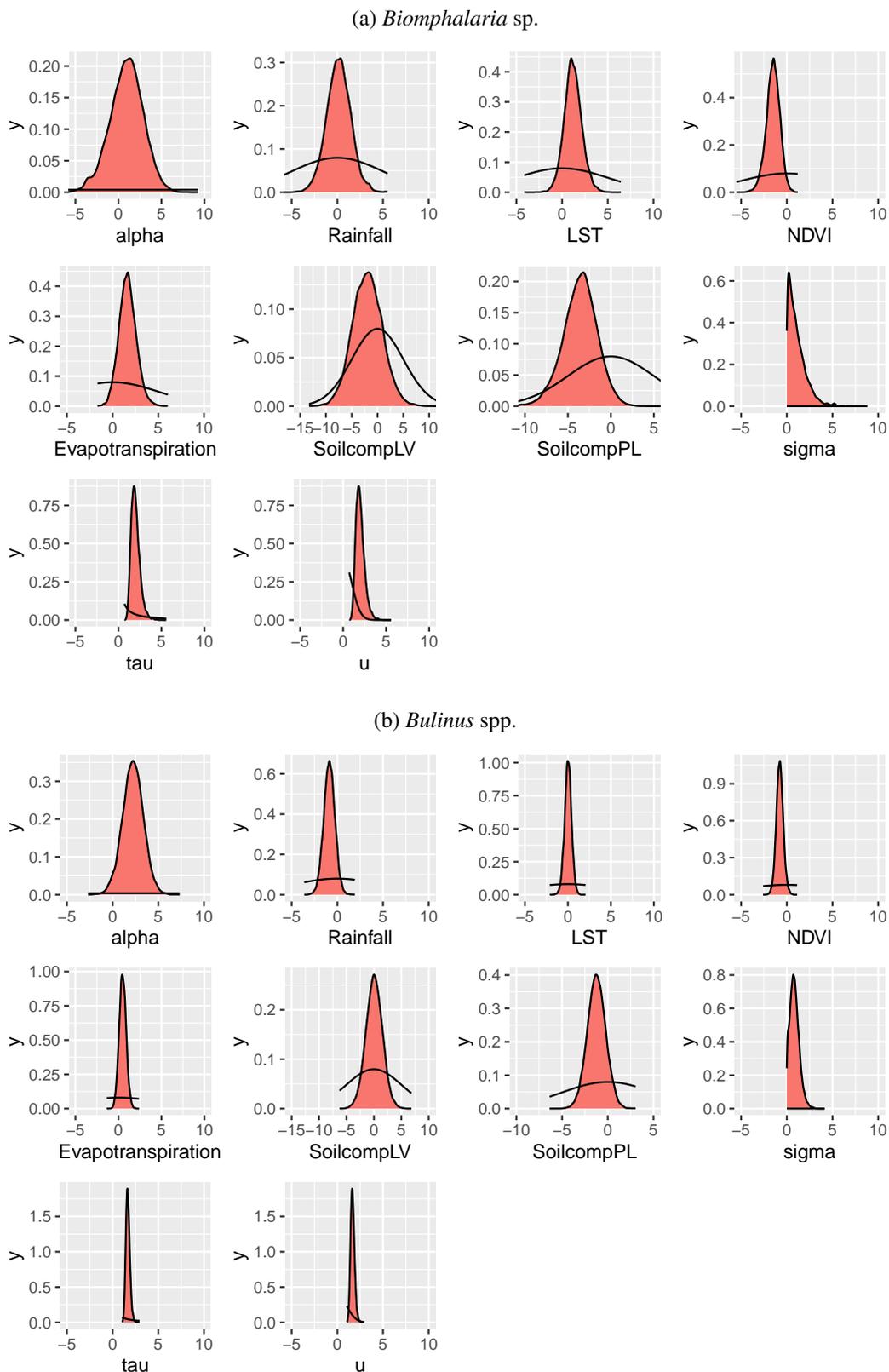


Figure C.10: Prior against the posterior distribution a) *Biomphalaria* sp. b) *Bulinus* spp.

C.7 1D result

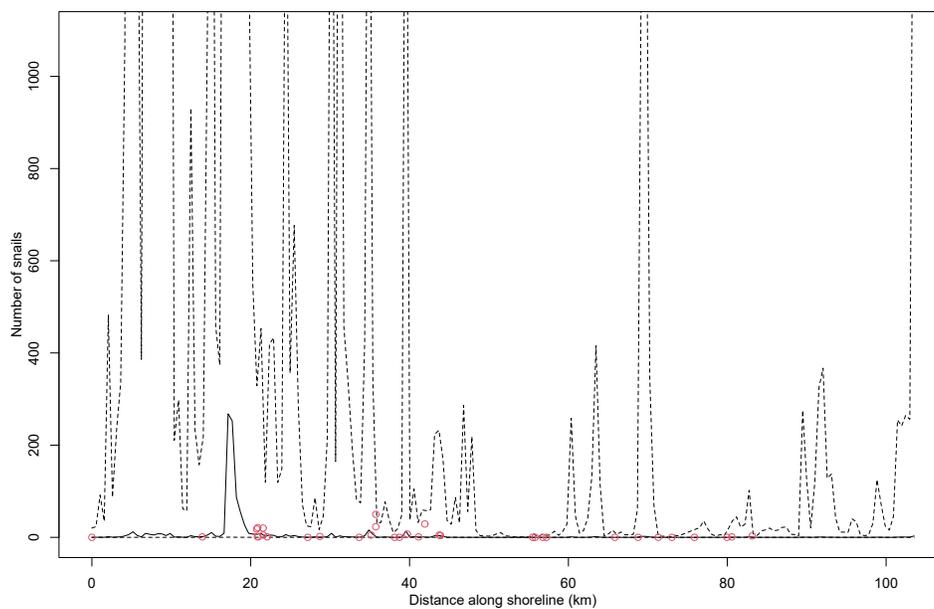
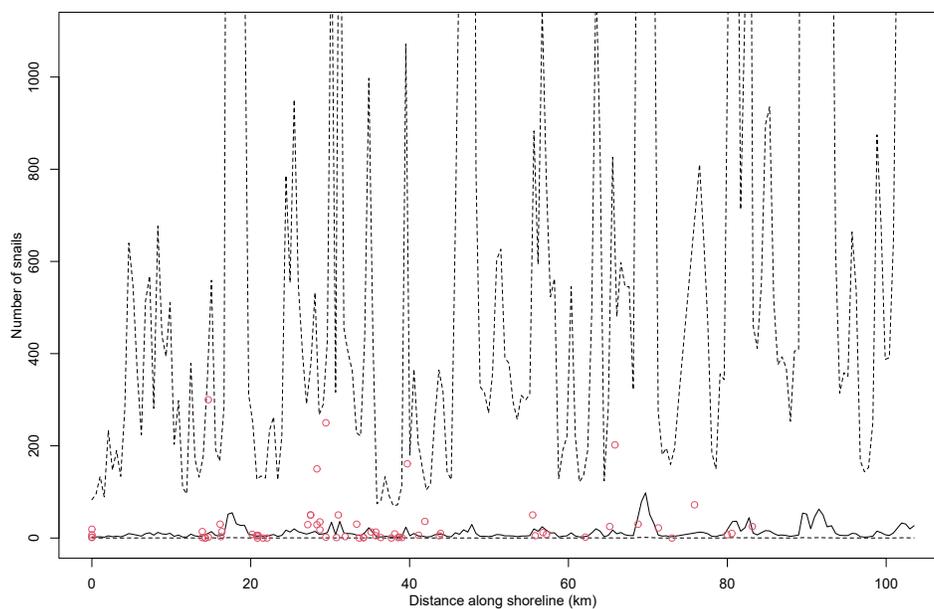
(a) *Biomphalaria* sp.(b) *Bulinus* spp.

Figure C.11: 1D mean GP prediction (exponential covariance function) of number of snails $\log(\hat{\mu}_i)$ against distance along the shoreline (km) a) *Biomphalaria* sp. b) *Bulinus* spp. Red crosses: observed number of snails at sampling locations along the shoreline. Black faded lines : SD of GP values

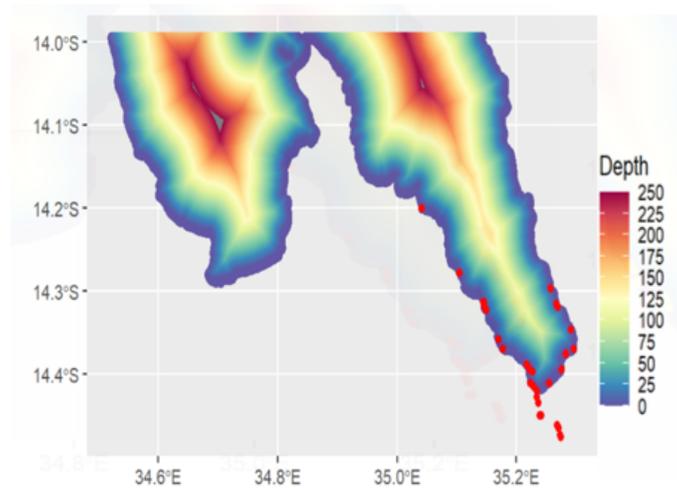
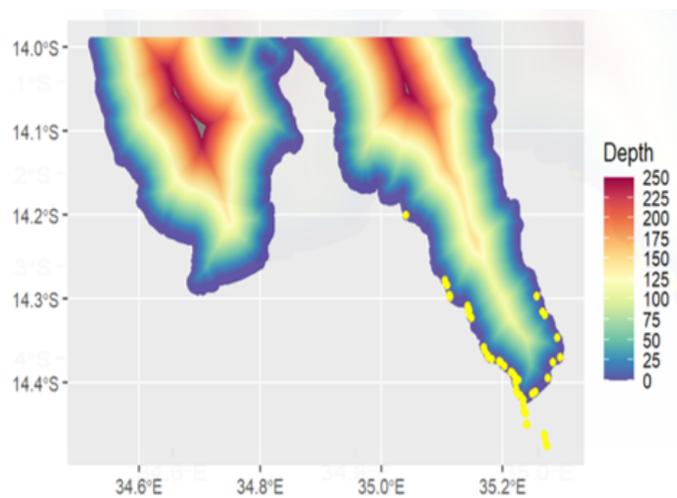
(a) *Biomphalaria* sp.(b) *Bulinus* spp.

Figure C.12: Bathymetric water depth (m) data for the shoreline with observed snails locations plotted a) *Biomphalaria* sp. b) *Bulinus* spp.

C.8 Bathymetric data

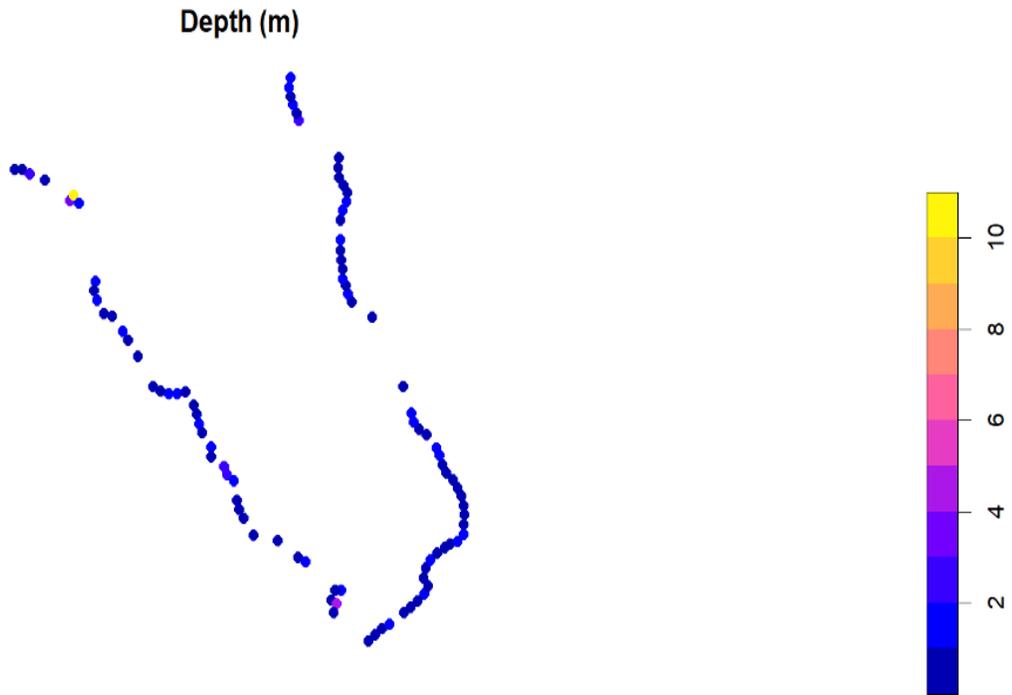


Figure C.13: 2D plot of depth (m) extracted for the prediction points with 100km buffer.

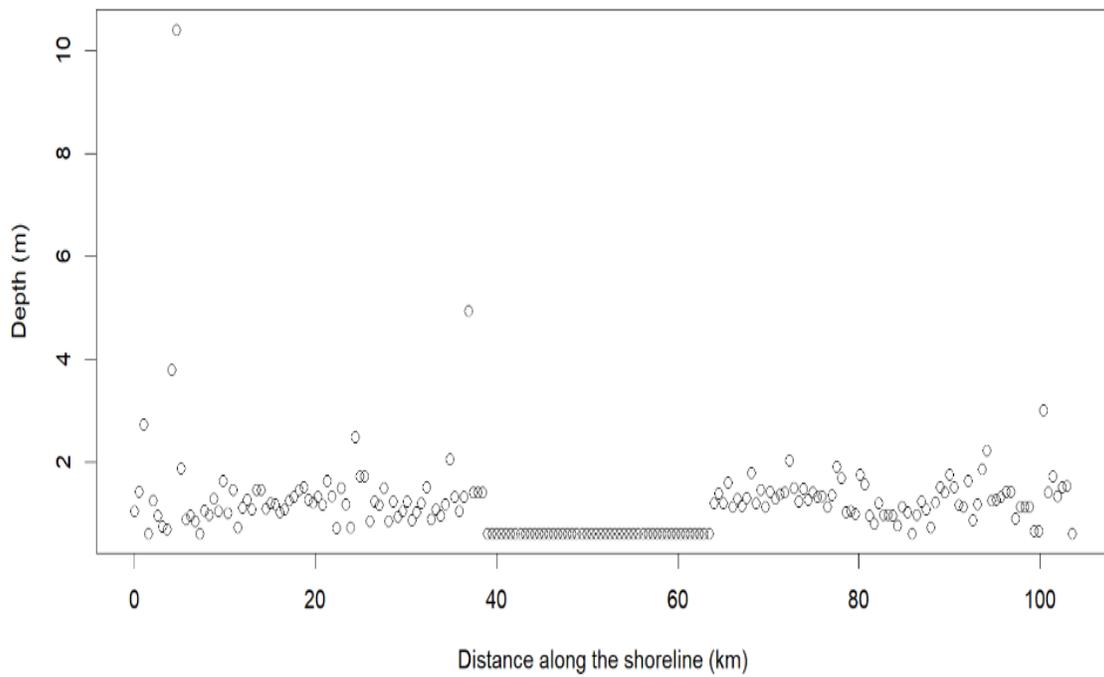


Figure C.14: Depth (m) against distance along (km) the shoreline with 100km buffer and fill in NAs for the shoreline.

Appendix D

Chapter 4

D.1 Sensitivity analysis

D.1.1 plots

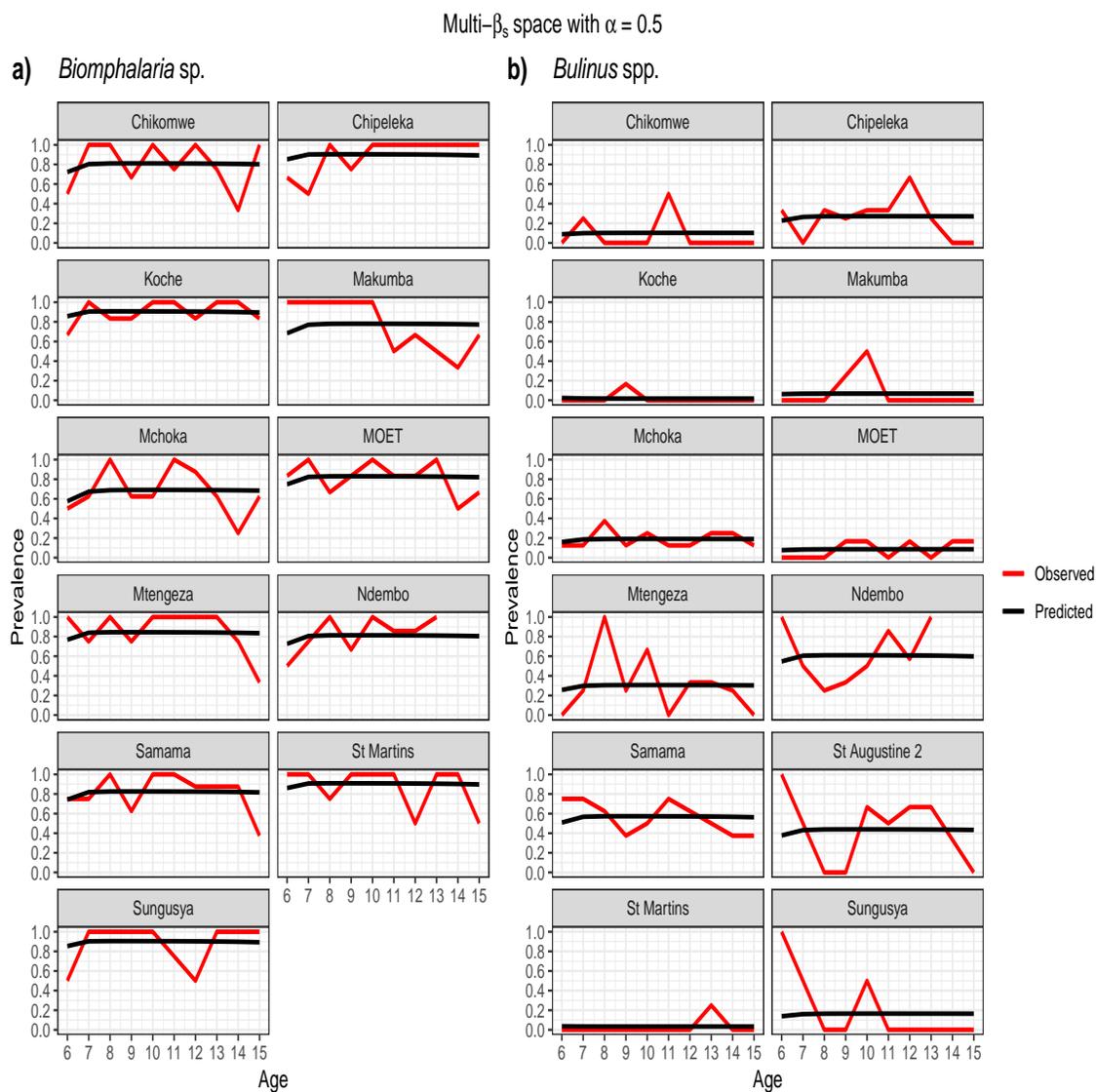


Figure D.1: Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Multi- β_s with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

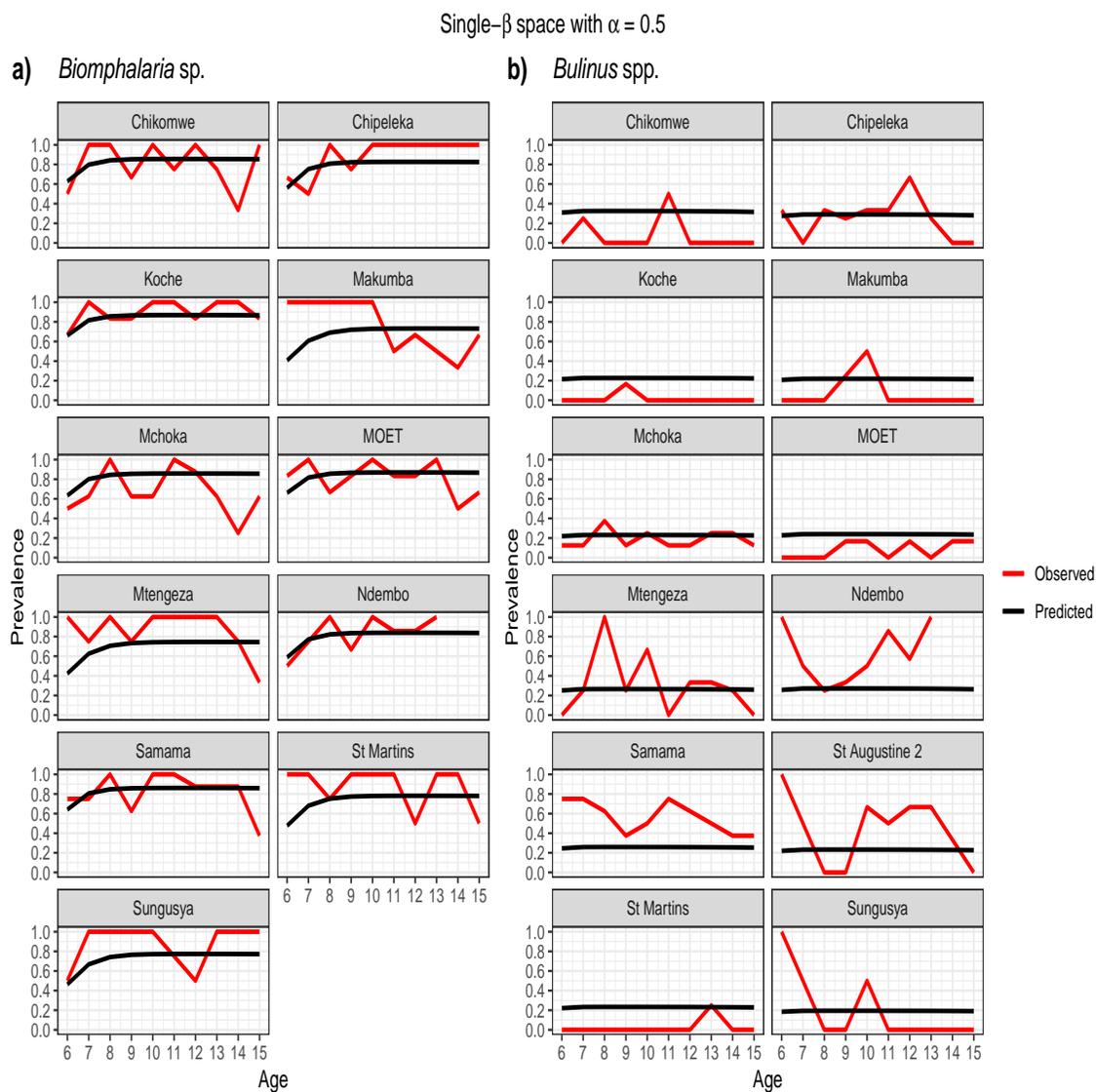


Figure D.2: Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Single- β with space effect model optimisation prevalence prediction (black line) and observed prevalence (black) against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

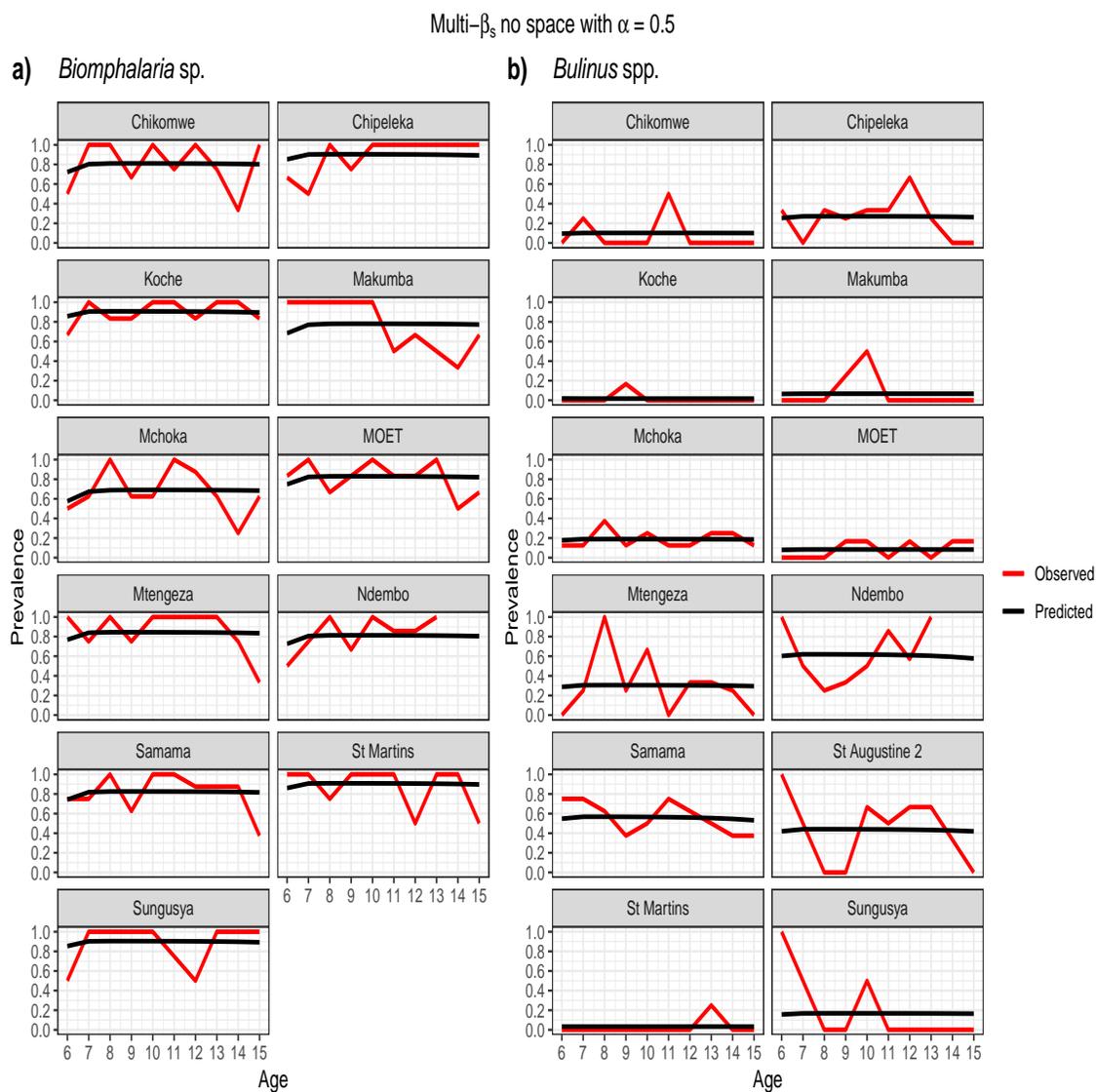


Figure D.3: Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Multi- β_s with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black) against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

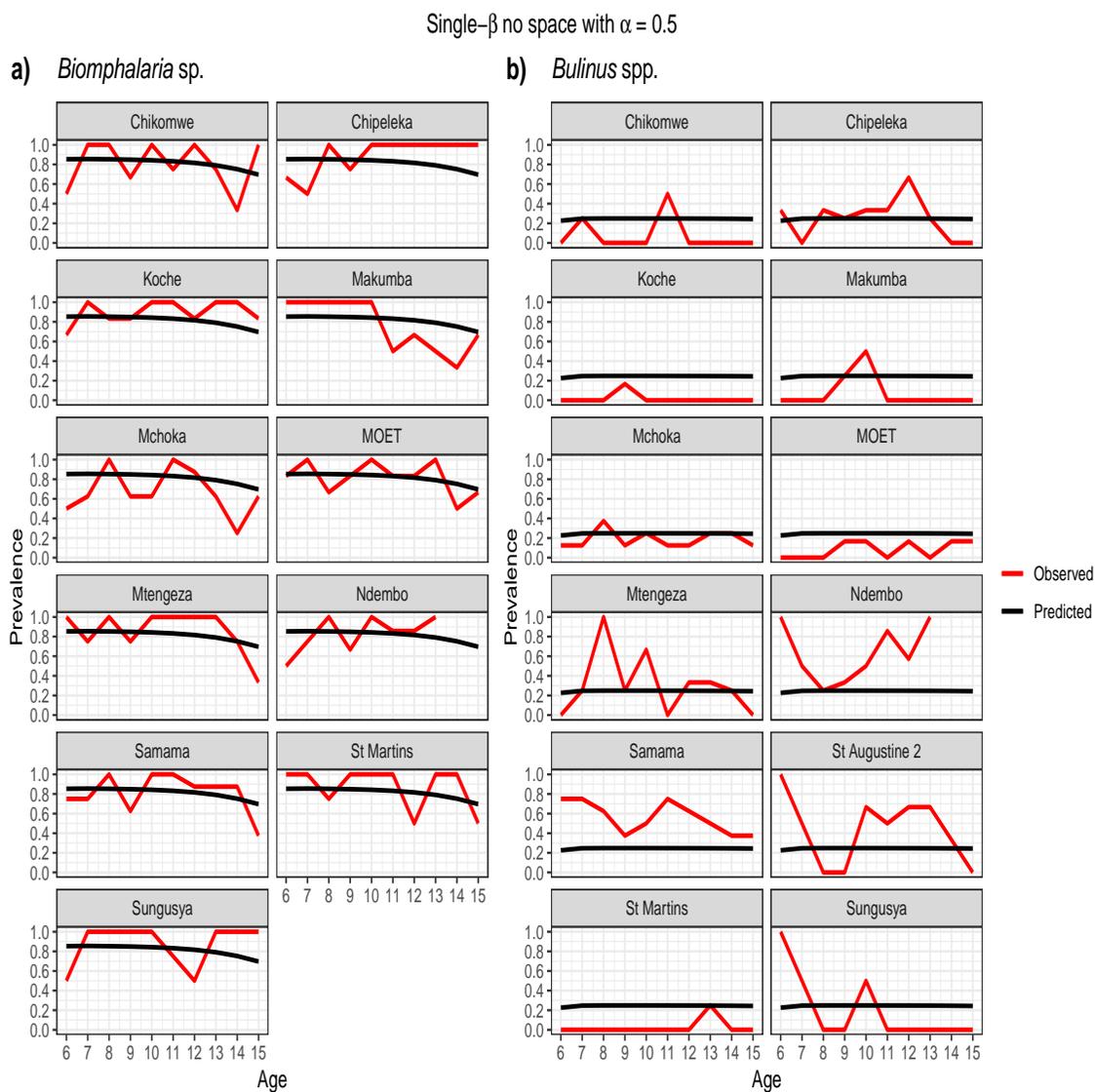


Figure D.4: Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Single- β with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

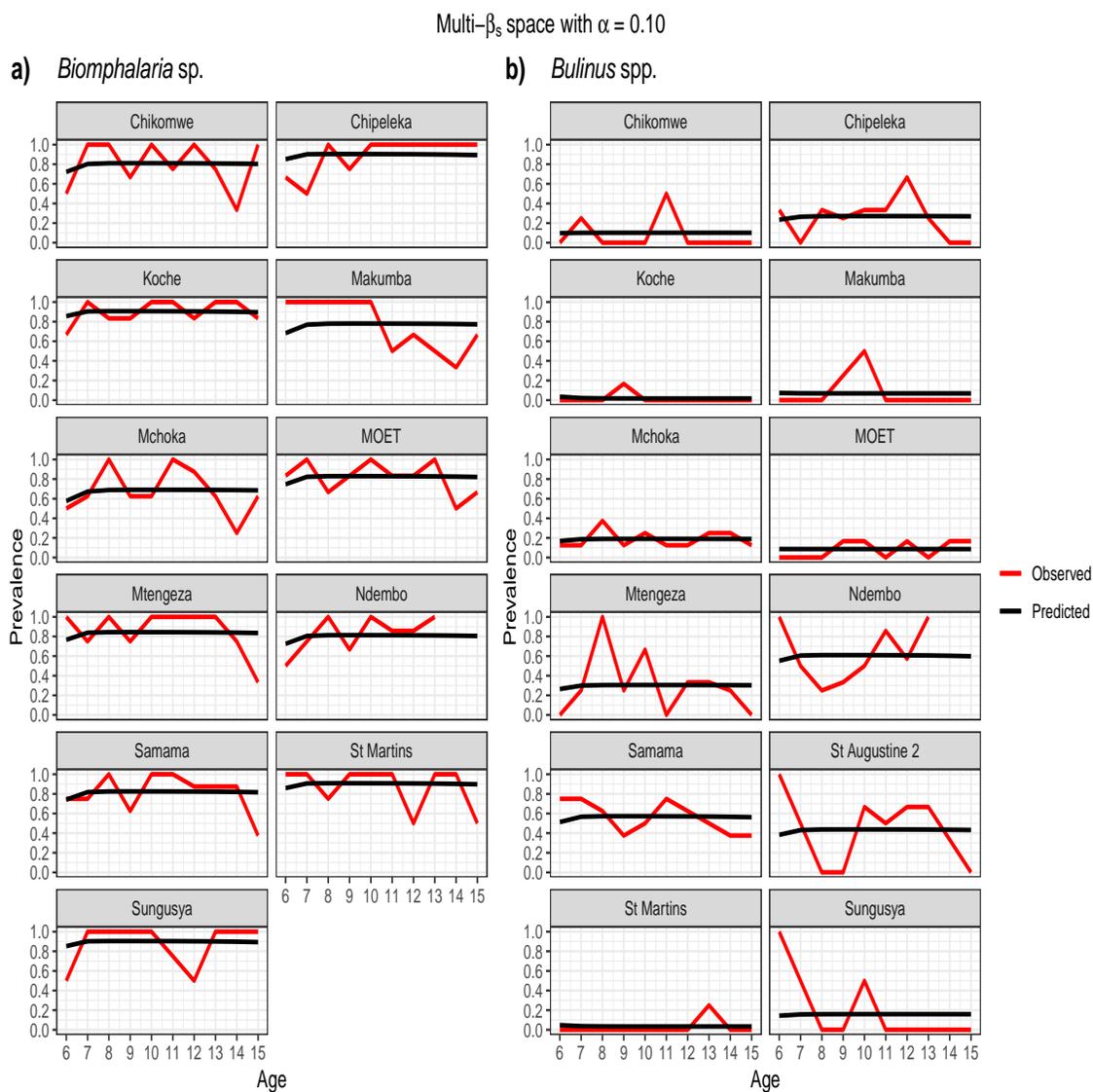


Figure D.5: Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Multi- β_s with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

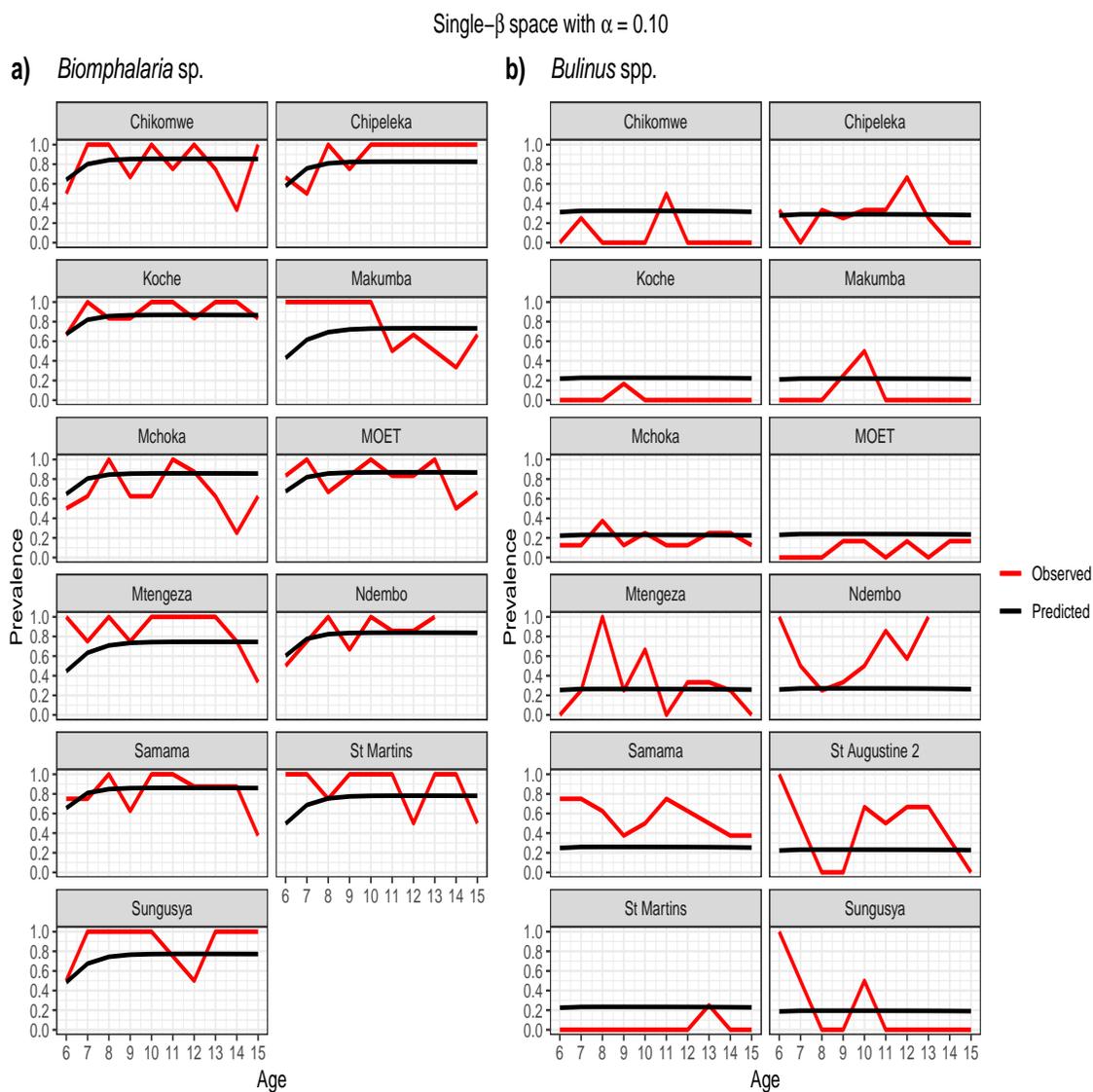


Figure D.6: Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Single- β with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

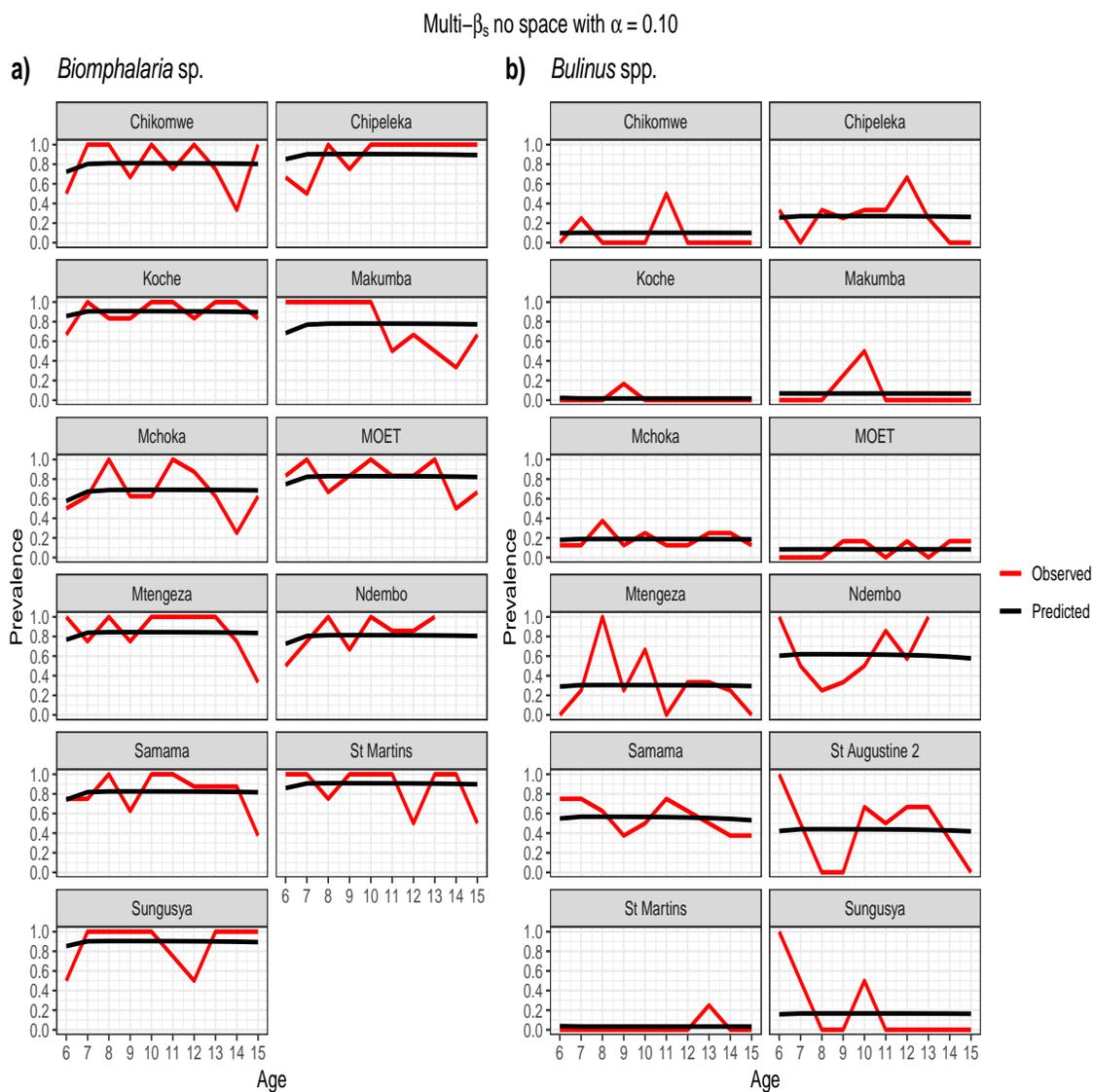


Figure D.7: Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Multi- β_s with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

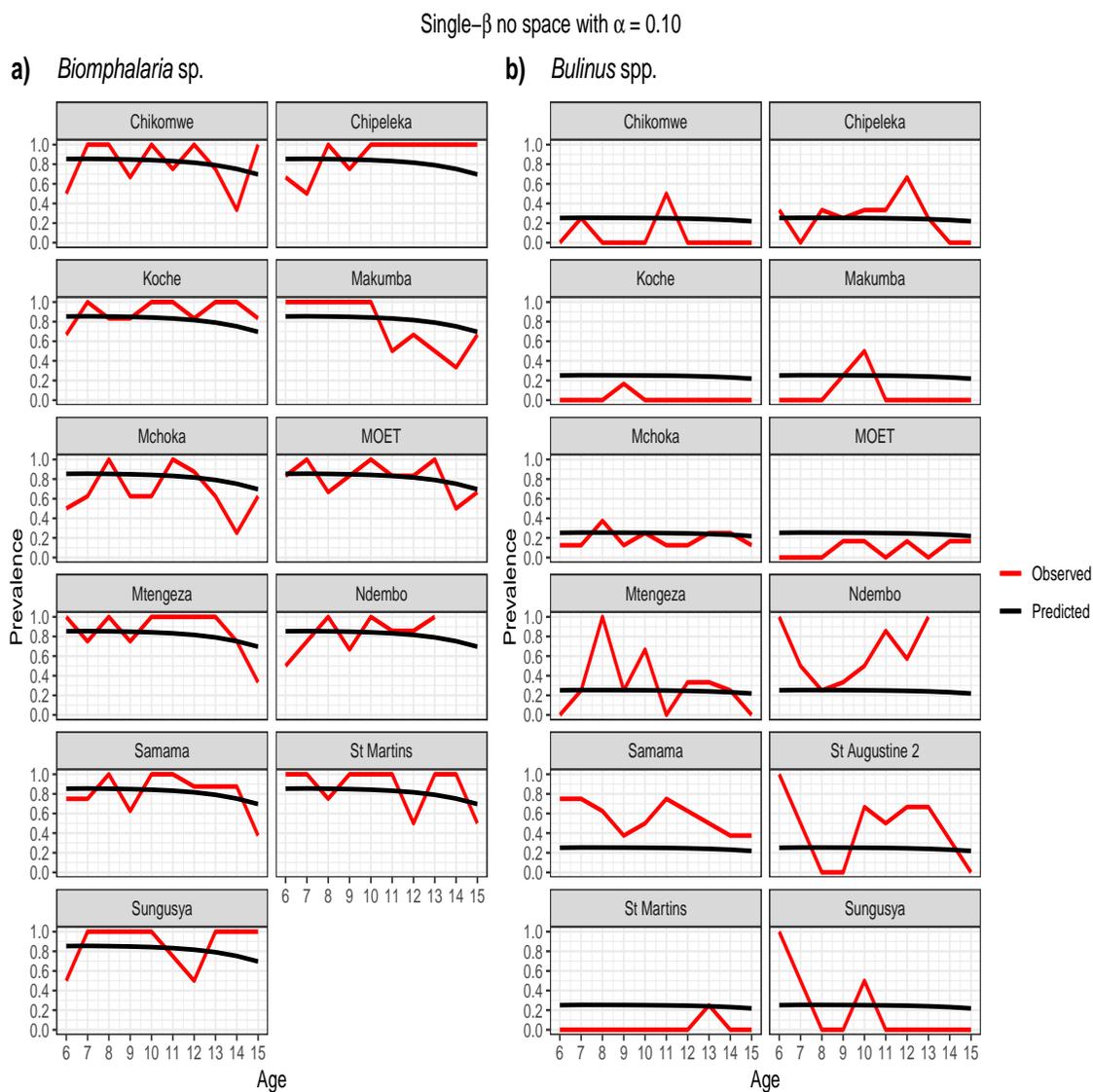


Figure D.8: Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Single- β with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

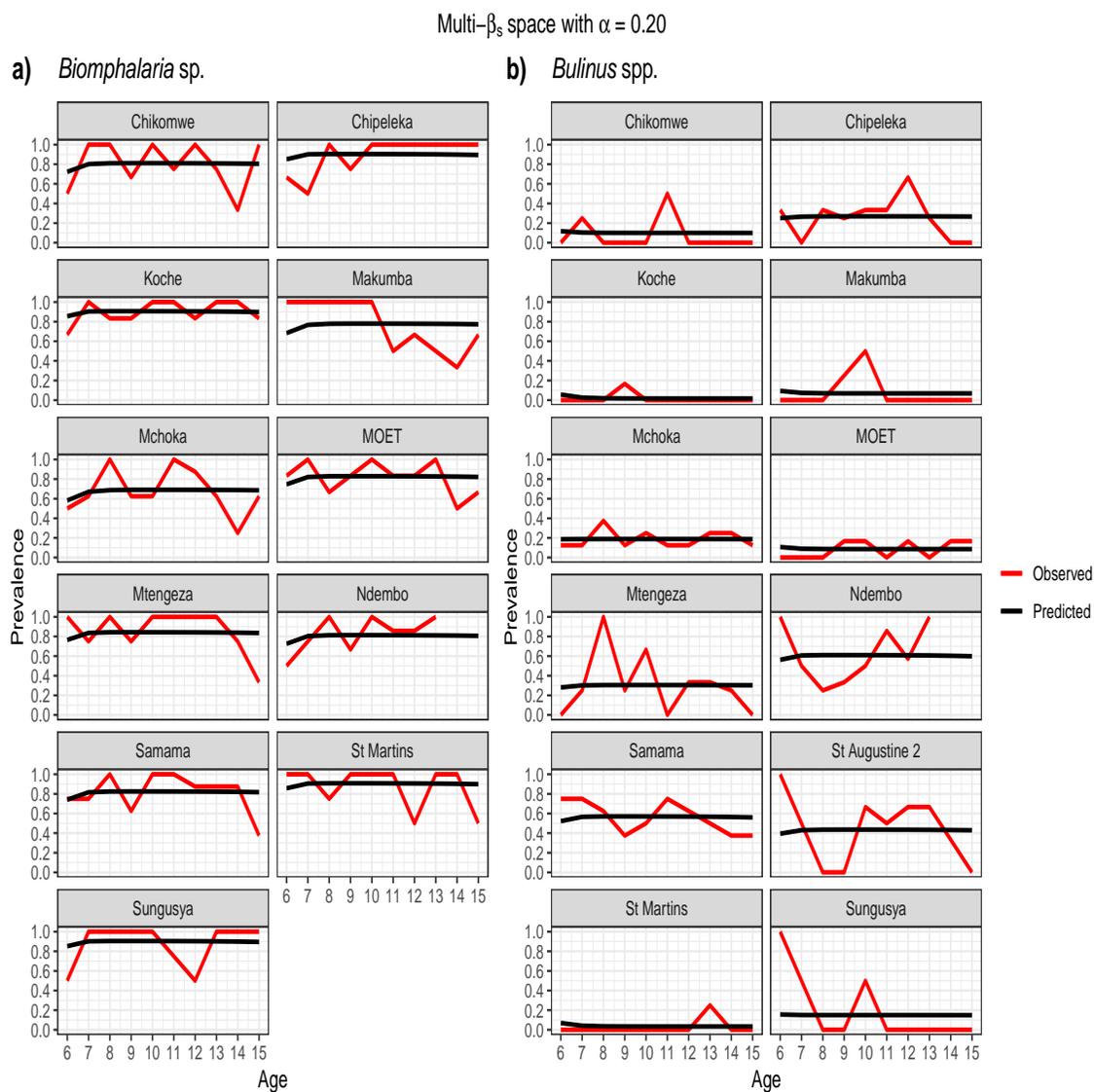


Figure D.9: Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Multi- β_s with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

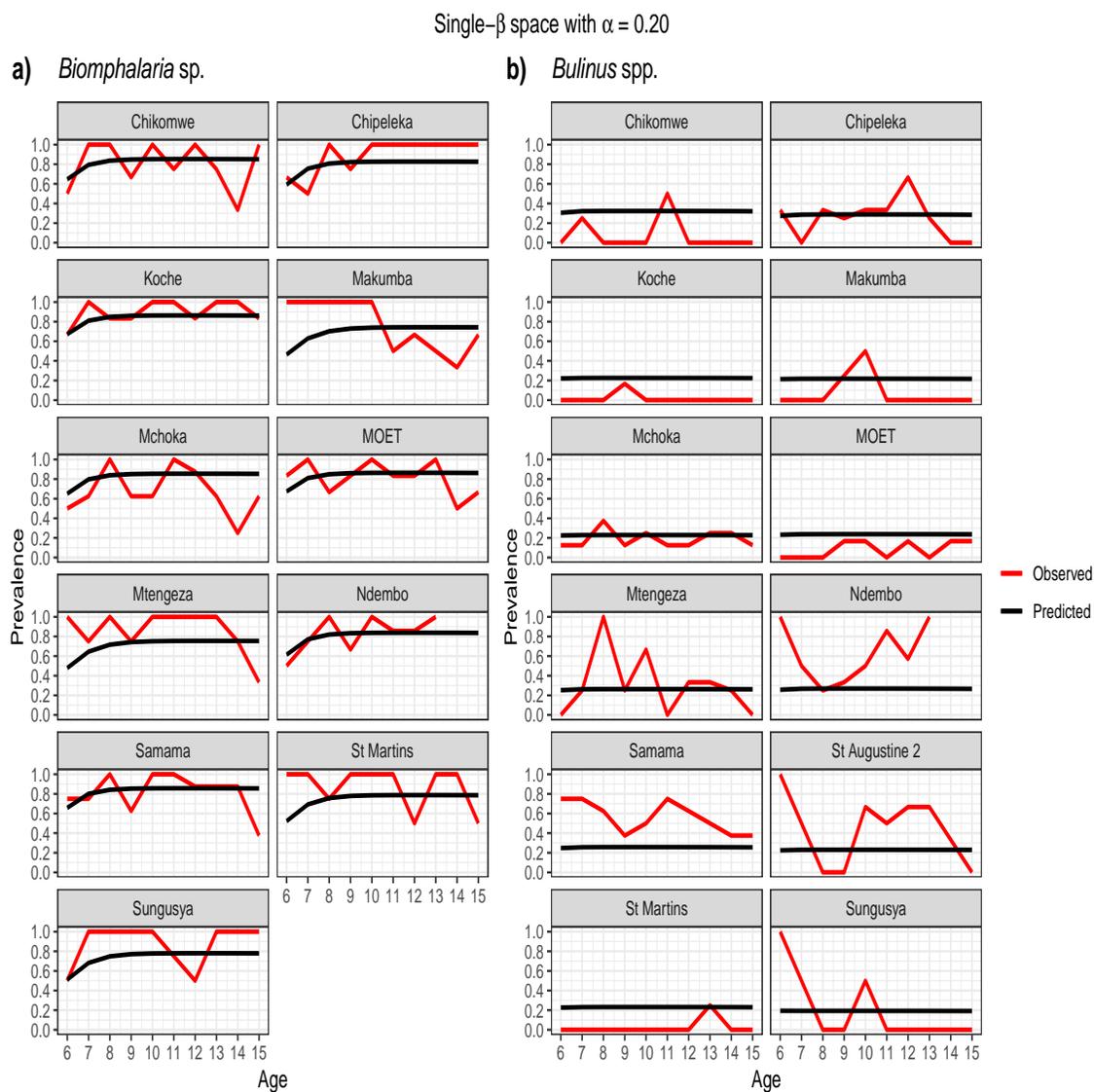


Figure D.10: Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Single- β with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

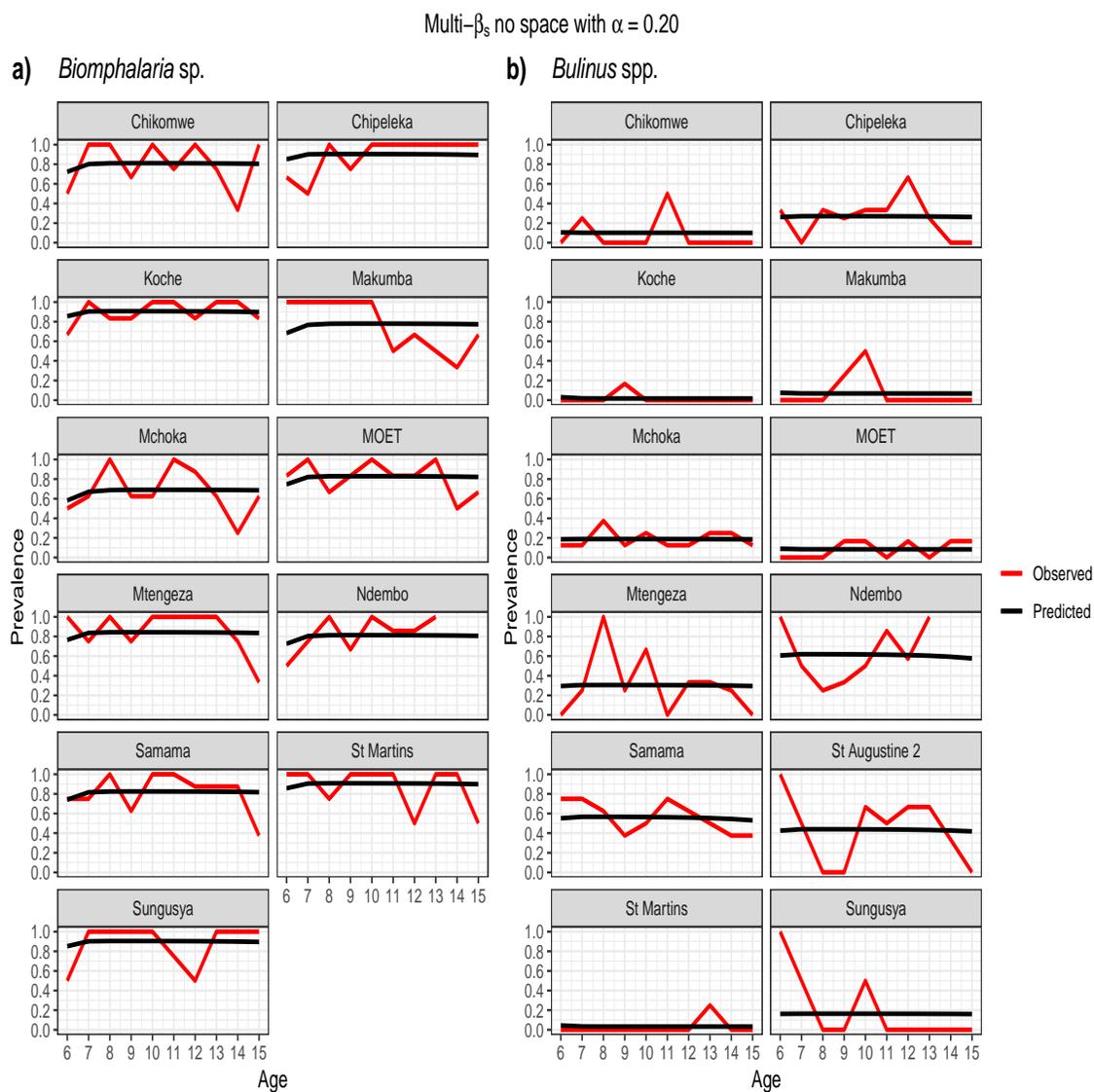


Figure D.11: Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Multi- β_s with space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

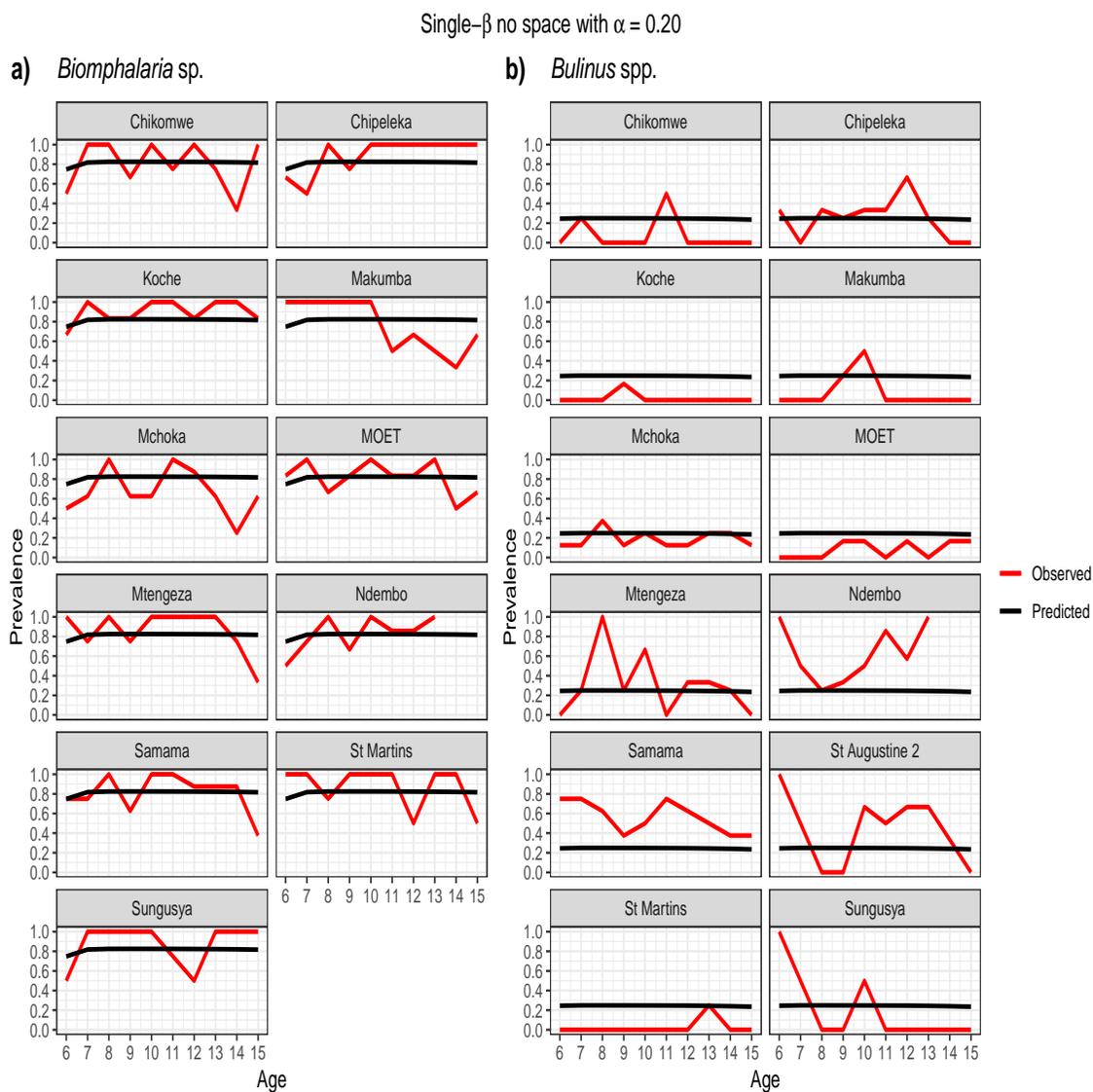


Figure D.12: Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Single- β with no space effect model optimisation prevalence prediction (black line) and observed prevalence (black against age of SAC carried out for each species a) *Biomphalaria* sp. b) *Bulinus* spp.

D.1.2 Tables

Multi- β_s space with $\alpha = 0.05$

Parameters/Schools [CI]	<i>Biomphalaria</i> AIC=258			MSE	<i>Bulinus</i> AIC=250			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-2.86 [-5.28, -0.438]	-2.66 [-9.01, 3.69]	-5.48 [-7.64, -3.31]	0.443	-6.31 [-8.47, -4.16]	-2.80 [-8.20, 2.60]	-4.65 [-6.65, 2.65]	0.0668
Samama	-2.23 [-5.77, 1.32]			0.340	-4.52 [-7.62, -1.43]			0.251
MOET	-4.73 [-8.39, -1.07]			0.253	-7.32 [-9.51, -5.12]			0.0675
Koche	-3.06 [-12.8, 6.67]			0.102	-8.95 [-11.8, -6.12]			0.0253
St Augustine 2					-5.03 [-7.63, -2.42]			1.14
Ndembo	-1.00 [-4.44, 2.43]			0.185	-4.38 [-7.79, -0.974]			0.653
Sungusya	1.58 [-7.69, 10.8]			0.379	-6.28 [-8.64, -3.91]			1.16
St Martins	1.62 [-9.27, 12.5]			0.420	-8.26 [-11.1, -5.46]			0.0573
Chikomwe	-1.85 [-5.24, 1.54]			0.487	-7.55 [-9.92, -5.18]			0.261
Chipeleka	0.679 [-8.55, 9.91]			0.290	-6.14 [-8.47, -3.82]			0.396
Makumba	0.356 [-2.70, 3.41]	0.673	-7.45 [-9.91, -4.98]	0.256				
Mtengeza	0.833 [-3.44, 5.11]	0.453	-5.84 [-8.19, -3.49]	0.874				

Figure D.13: Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species

D.1.3 Confidence intervals

Single- β , space with $\alpha = 0.05$

Parameters/Schools [CI]	Biomphalaria AIC=275			MSE	Bulinus AIC=341			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-1.00 [-1.67, -0.328]	-4.96 [-5.49, -4.42]	-6.86 [-7.45, -6.28]	0.677	-5.00 [-10.4, 0.350]	0.500 [-12.7, 13.7]	-3.55 [-8.84, 1.75]	0.0861
Samama				0.366				1.16
MOET				0.313				0.308
Koche				0.107				0.460
St Augustine 2								1.49
Ndembo				0.121				1.63
Sungusya				0.516				1.11
St Martins				0.777				0.483
Chikomwe				0.474				0.859
Chipeleka				0.298				0.400
Makumba				1.03				0.456
Mtengeza				0.865				0.911

Figure D.14: Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Parameter estimates for single- β spatial model for each species

Multi- β_s no space with $\alpha = 0.05$

Parameters/Schools [CI]	<i>Biomphalaria</i> AIC=258			MSE	<i>Bulinus</i> AIC=248			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-6.14 [-8.56, -3.73]	-2.67 [-8.90, 3.55]	-5.48 [-7.64, -3.31]	0.443	-7.26 [-12.5, -2.02]	-2.50 [-8.51, 3.51]	-3.39 [-8.61, 1.84]	0.0686
Samama	-5.19 [-8.73, -1.65]			0.340	-4.80 [-11.0, 1.44]			0.214
MOET	-5.14 [-8.80, -1.49]			0.253	-8.26 [-13.5, -3.01]			0.0688
Koche	-3.87 [-13.7, 5.94]			0.102	-9.97 [-15.5, -4.43]			0.0251
St Augustine 2					-5.73 [-11.3, -0.177]			1.08
Ndembo	-5.30 [-8.73, -1.87]			0.185	-4.20 [-12.1, 3.74]			0.623
Sungusya	-3.95 [-13.3, 5.39]			0.379	-7.42 [-12.8, -2.03]			1.13
St Martins	-3.81 [-14.8, 7.21]			0.420	-9.26 [-14.8, -3.70]			0.0570
Chikomwe	-5.33 [-8.71, -1.95]			0.487	-8.05 [-13.4, -2.69]			0.262
Chipeleka	-3.96 [-13.3, 5.36]			0.290	-6.73 [-12.1, -1.37]			0.390
Makumba	-5.57 [-8.63, -2.52]			0.673	-8.51 [-13.9, -3.11]			0.256
Mtengeza	-4.98 [-9.26, -0.714]			0.453	-6.53 [-11.9, -1.12]			0.885

Figure D.15: Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species

Single- β_s no space with $\alpha = 0.05$

Parameters/Schools [CI]	<i>Biomphalaria</i> AIC=250			MSE	<i>Bulinus</i> AIC=351			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-1.48 [-4.87, 1.92]	0.148 [-3.24, 3.54]	-2.61 [-2.85, -2.36]	0.613	-5.50 [-17.5, 6.49]	-4.63 [-16.8, 7.56]	-3.70 [-12.2, 4.85]	0.101
Samama				0.275				1.24
MOET				0.190				0.329
Koche				0.233				0.546
St Augustine 2								1.44
Ndembo				0.262				1.78
Sungusya				0.519				1.15
St Martins				0.372				0.540
Chikomwe				0.535				0.542
Chipeleka				0.476				0.389
Makumba				0.506				0.541
Mtengeza				0.326				0.926

Figure D.16: Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Parameter estimates for single- β spatial model for each species

Multi- β_s space with $\alpha = 0.10$

Parameters/Schools [CI]	<i>Biomphalaria</i> AIC=258			MSE	<i>Bulinus</i> AIC=250			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-2.94 [-5.26, -0.618]	-2.74 [-8.83, 3.34]	-5.56 [-7.62, -3.49]	0.443	-6.32 [-8.62, -4.02]	-2.80 [-8.51, 2.91]	-4.65 [-6.79, -2.51]	0.0676
Samama	-2.31 [-5.71, 1.09]			0.340	-4.53 [-7.80, -1.25]			0.248
MOET	-4.81 [-8.33, -1.30]			0.254	-7.32 [-9.63, -5.00]			0.0676
Koche	-3.14 [-12.5, 6.19]			0.102	-8.95 [-11.9, -5.98]			0.0260
St Augustine 2					-5.03 [-7.79, -2.27]			1.13
Ndembo	-1.08 [-4.38, 2.22]			0.185	-4.38 [-7.99, -0.775]			0.648
Sungusya	1.50 [-7.43, 10.4]			0.378	-6.32 [-8.91, -3.73]			1.15
St Martins	1.53 [-8.85, 11.9]			0.421	-8.26 [-11.2, -5.36]			0.0584
Chikomwe	-1.93 [-5.19, 1.33]			0.487	-7.55 [-10.1, -5.05]			0.263
Chipeleka	0.604 [-8.33, 9.53]			0.289	-6.15 [-8.62, -3.68]			0.394
Makumba	0.273 [-2.67, 3.21]	0.674	-7.45 [-10.0, -4.87]	0.257				
Mtengeza	0.748 [-3.35, 4.85]	0.454	-5.84 [-8.33, -3.36]	0.879				

Figure D.17: Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species

Single- β , space with $\alpha = 0.10$

Parameters/Schools [CI]	Biomphalaria AIC=275			MSE	Bulinus AIC=341			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-1.00 [-1.67, -0.328]	-4.96 [-5.49, -4.42]	-6.86 [-7.45, -6.28]	0.682	-5.00 [-13.8, 3.80]	0.500 [-13.9, 14.9]	-3.55 [-12.2, 5.14]	0.0866
Samama				0.364				1.16
MOET				0.308				0.309
Koche				0.106				0.460
St Augustine 2								1.49
Ndembo				0.124				1.63
Sungusya				0.509				1.11
St Martins				0.753				0.483
Chikomwe				0.476				0.860
Chipeleka				0.297				0.399
Makumba				0.995				0.456
Mtengeza				0.837				0.913

Figure D.18: Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Parameter estimates for single- β spatial model for each species

Multi- β_s no space with $\alpha = 0.10$

Parameters/Schools [CI]	Biomphalaria AIC=258			MSE	Bulinus AIC=248			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-6.22 [-8.55, -3.90]	-2.74 [-8.80, 3.32]	-5.56 [-7.63, -3.48]	0.443	-7.26 [-12.5, -1.98]	-2.50 [-8.56, 3.57]	-3.38 [-8.64, 1.88]	0.0690
Samama	-5.27 [-8.67, -1.87]			0.340	-4.80 [-11.0, 1.44]			0.214
MOET	-5.23 [-8.74, -1.71]			0.254	-8.26 [-13.5, -2.97]			0.0694
Koche	-3.96 [-13.3, 5.38]			0.102	-9.97 [-15.5, -4.40]			0.0252
St Augustine 2					-5.73 [-11.3, -0.144]			1.07
Ndembo	-5.38 [-8.69, -2.08]			0.185	-4.19 [-12.1, 3.73]			0.621
Sungusya	-4.04 [-13.0, 4.89]			0.378	-7.43 [-12.9, -1.94]			1.12
St Martins	-3.91 [-14.3, 6.49]			0.421	-9.26 [-14.9, -3.67]			0.0573
Chikomwe	-5.41 [-8.67, -2.15]			0.487	-8.05 [-13.5, -2.65]			0.262
Chipeleka	-4.05 [-12.9, 4.89]			0.289	-6.73 [-12.1, -1.33]			0.389
Makumba	-5.66 [-8.60, -2.72]			0.674	-8.51 [-13.9, -3.07]			0.257
Mtengeza	-5.07 [-9.17, -0.971]			0.454	-6.52 [-12.0, -1.09]			0.887

Figure D.19: Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species

Single- β_s no space with $\alpha = 0.10$

Parameters/Schools [CI]	<i>Biomphalaria</i> AIC=250			MSE	<i>Bulinus</i> AIC=NA			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-1.48	0.148	-2.61	0.614	NA	NA	NA	NA
	[-4.80, 1.85]	[-3.17, 3.47]	[-2.84, -2.37]					
Samama				0.275				
MOET				0.190				
Koche				0.233				
St Augustine 2								
Ndembo				0.262				
Sungusya				0.519				
St Martins				0.372				
Chikomwe				0.535				
Chipeleka				0.476				
Makumba				0.506				
Mtengeza				0.326				

Figure D.20: Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Parameter estimates for single- β spatial model for each species

Multi- β_s space with $\alpha = 0.20$

Parameters/Schools [CI]	<i>Biomphalaria</i> AIC=259			MSE	<i>Bulinus</i> AIC=250			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-3.09 [-5.31, -0.883]	-2.89 [-8.66, 2.88]	-5.71 [-7.67, -3.75]	0.445	-6.33 [-8.81, -3.85]	-2.80 [-9.26, 3.66]	-4.65 [-6.95, -2.35]	0.0668
Samama	-2.47 [-5.69, 0.747]			0.341	-4.54 [-8.08, -0.995]			0.242
MOET	-4.98 [-8.31, -1.64]			0.255	-7.32 [-9.78, -4.85]			0.0745
Koche	-3.31 [-12.1, 5.44]			0.101	-8.96 [-12.1, -5.77]			0.0282
St Augustine 2					-5.04 [-8.01, -2.07]			1.11
Ndembo	-1.24 [-4.38, 1.91]			0.185	-4.38 [-8.29, -0.467]			0.639
Sungusya	1.34 [-7.08, 9.75]			0.377	-6.40 [-9.34, -3.46]			1.11
St Martins	1.34 [-8.30, 11.0]			0.423	-8.26 [-11.3, -5.25]			0.0614
Chikomwe	-2.08 [-5.19, 1.03]			0.488	-7.57 [-10.3, -4.86]			0.265
Chipeleka	0.442 [-8.01, 8.89]			0.287	-6.16 [-8.81, -3.51]			0.391
Makumba	0.110 [-2.70, 2.92]			0.676	-7.45 [-10.2, -4.72]			0.261
Mtengeza	0.578 [-3.30, 4.46]			0.456	5.84 [-8.51, -3.18]			0.888

Figure D.21: Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species

Single- β , space with $\alpha = 0.20$

Parameters/Schools [CI]	Biomphalaria AIC=271			MSE	Bulinus AIC=341			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-1.00 [-1.75, -0.253]	-5.13 [-5.74, -4.52]	-7.00 [-7.68, -6.32]	0.673	-6.00 [-14.6, 2.61]	0.500 [-12.8, 11.8]	-4.53 [-13.1, 4.00]	0.0867
Samama				0.361				1.17
MOET				0.305				0.307
Koche				0.114				0.459
St Augustine 2								1.49
Ndembo				0.127				1.64
Sungusya				0.492				1.10
St Martins				0.718				0.481
Chikomwe				0.479				0.858
Chipeleka				0.293				0.401
Makumba				0.951				0.457
Mtengeza				0.783				0.917

Figure D.22: Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Parameter estimates for single- β spatial model for each species

Multi- β_s no space with $\sigma = 0.20$

Parameters/Schools [CI]	<i>Biomphalaria</i> AIC=259			MSE	<i>Bulinus</i> AIC=248			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-6.38 [-8.59, -4.18]	-2.88 [-8.66, 2.89]	-5.71 [-7.67 -3.75]	0.445	-7.26 [-12.4, -2.16]	-2.50 [-8.40, 3.40]	-3.38 [-8.44, 1.68]	0.0697
Samama	-5.44 [-8.64, -2.25]			0.341	-4.80 [-10.8, 1.21]			0.212
MOET	-5.40 [-8.71, -2.09]			0.255	-8.26 [-13.4, -3.16]			0.0707
Koche	-4.15 [-12.8, 4.46]			0.101	-9.97 [-15.4, -4.57]			0.0256
St Augustine 2					-5.73 [-11.1, -0.338]			1.07
Ndembo	-5.55 [-8.67, -2.42]			0.185	-4.19 [-11.9, 3.50]			0.618
Sungusya	-4.22 [-12.5, 4.07]			0.377	-7.45 [-12.9, -2.04]			1.11
St Martins	-4.11 [-13.6, 5.33]			0.423	-9.26 [-14.7, -3.85]			0.0579
Chikomwe	-5.57 [-8.66, -2.48]			0.488	-8.05 [-13.3, -2.82]			0.263
Chipeleka	-4.23 [-12.5, 4.09]			0.287	-6.73 [-11.9 -1.52]			0.388
Makumba	-5.83 [-8.62, -3.04]			0.676	-8.51 [-13.8, -3.26]			0.258
Mtengeza	-5.25 [-9.10, -1.40]			0.456	-6.52 [-11.8 -1.29]			0.890

Figure D.23: Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Parameter estimates for multi- β_s spatial model for each species

Single- β_s no space with $\alpha = 0.20$

Parameters/Schools [CI]	<i>Biomphalaria</i> AIC=256			MSE	<i>Bulinus</i> AIC=NA			MSE
	$\log \beta$	$\log \sigma$	$\log \gamma$		$\log \beta$	$\log \sigma$	$\log \gamma$	
Mchoka	-4.96	-3.34	-5.60	0.644	NA	NA	NA	NA
	[-10.1, 0.211]	[-8.53, 1.86]	[-8.33, -2.86]					
Samama				0.341				
MOET				0.253				
Koche				0.165				
St Augustine 2								
Ndembo				0.186				
Sungusya				0.394				
St Martins				0.463				
Chikomwe				0.493				
Chipeleka				0.334				
Makumba				0.682				
Mtengeza				0.467				

Figure D.24: Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Parameter estimates for single- β spatial model for each species

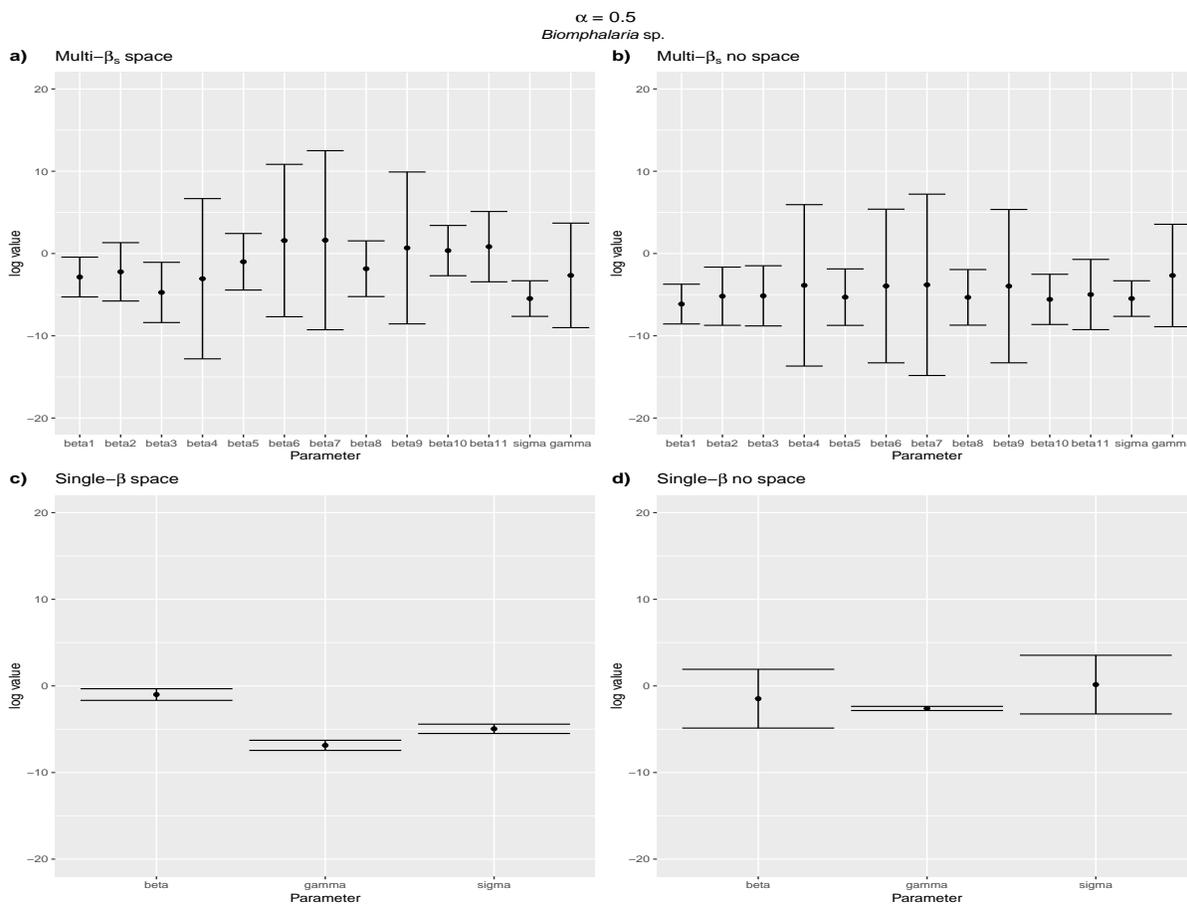


Figure D.25: Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Confidence intervals for parameter estimates for *Biomphalaria* sp. models with SAC prevalence at age 6 set as α set to 0.05 (5%) a) Multi- β_s space sp. b) Multi- β_s no space c) Single- β space d) Single- β no space

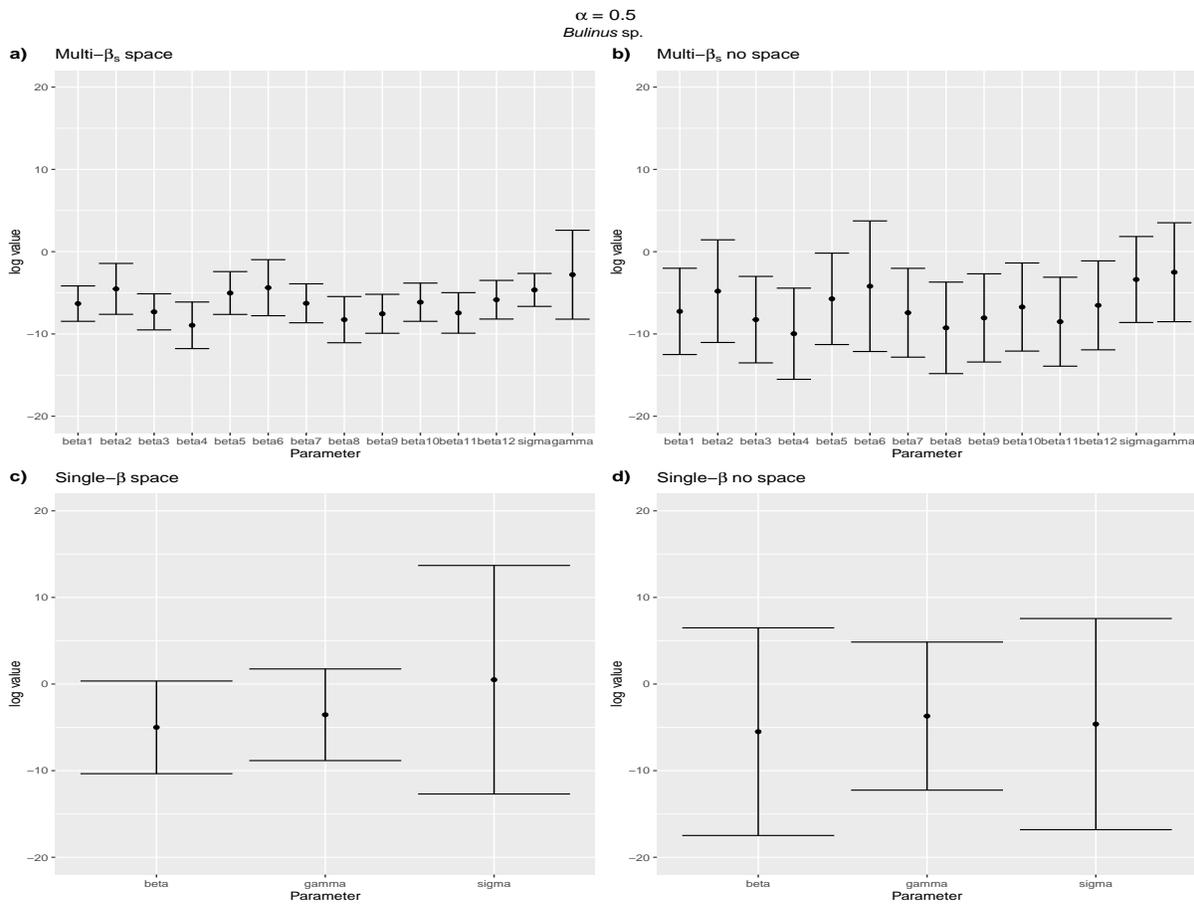


Figure D.26: Prior *Schistosoma* infection prevalence of $\alpha = 0.05$ for SAC age 6. Confidence intervals for parameter estimates *Bulinus* spp. models with SAC prevalence at age 6 set as α set to 0.05 (5%) a) Multi- β_s space sp. b) Multi- β_s no space c) Single- β space d) Single- β no space

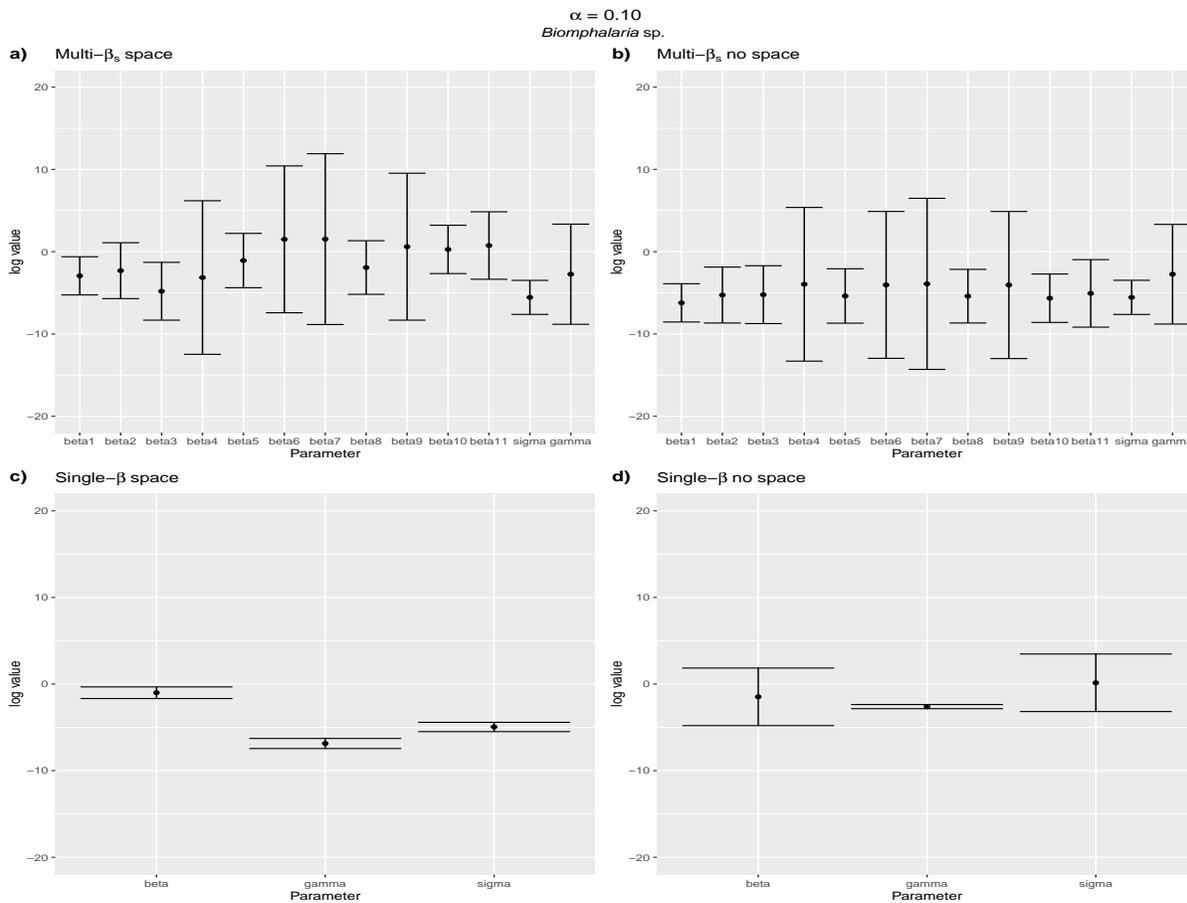


Figure D.27: Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Confidence intervals for parameter estimates for *Biomphalaria* sp. models with SAC prevalence at age 6 set as α set to 0.10 (10%); a) Multi- β_s space b) Multi- β_s no space c) Multi- β_s space d) Multi- β_s no space

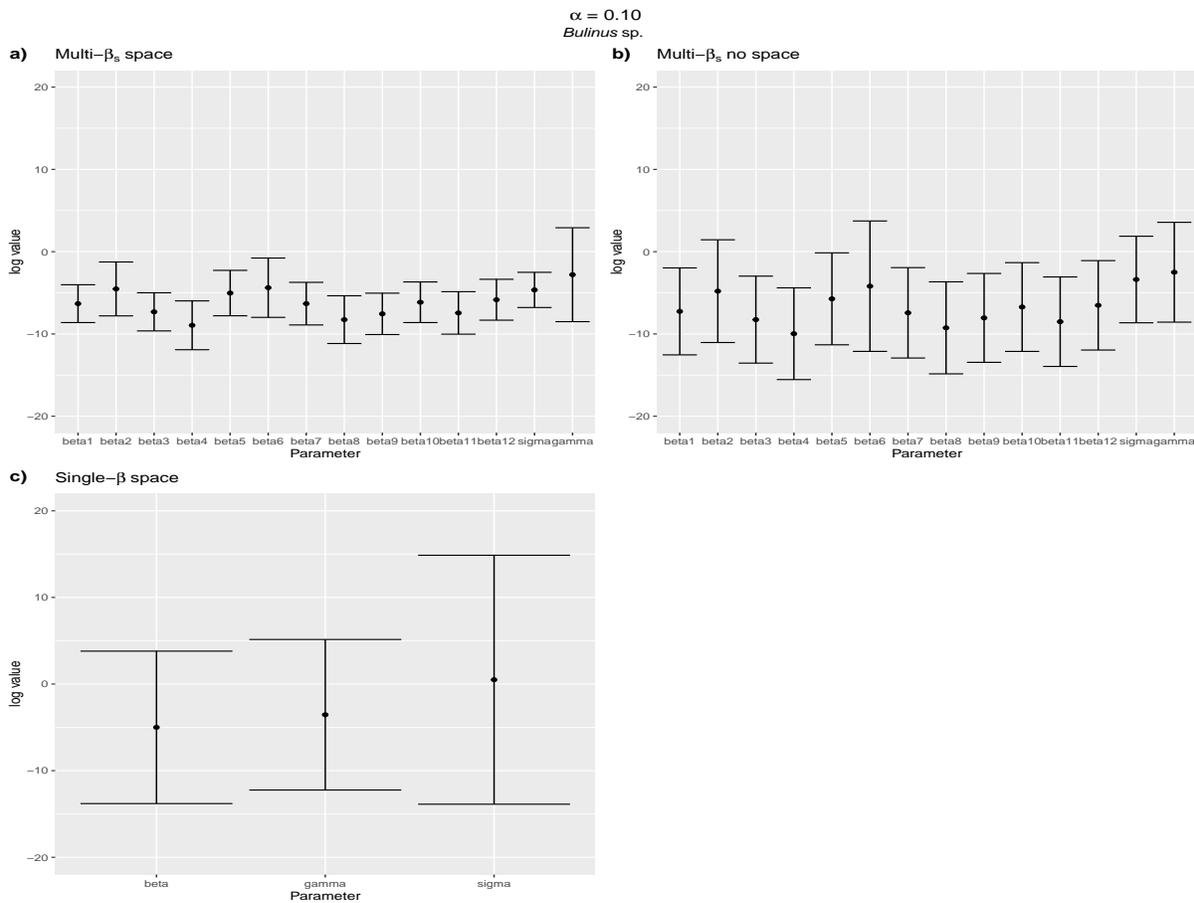


Figure D.28: Prior *Schistosoma* infection prevalence of $\alpha = 0.10$ for SAC age 6. Confidence intervals for parameter estimates for *Bulinus* spp. models with SAC prevalence at age 6 set as α set to 0.10 (10%); a) Multi- β_s space sp. b) Multi- β_s no space c) Single- β space d) Single- β no space

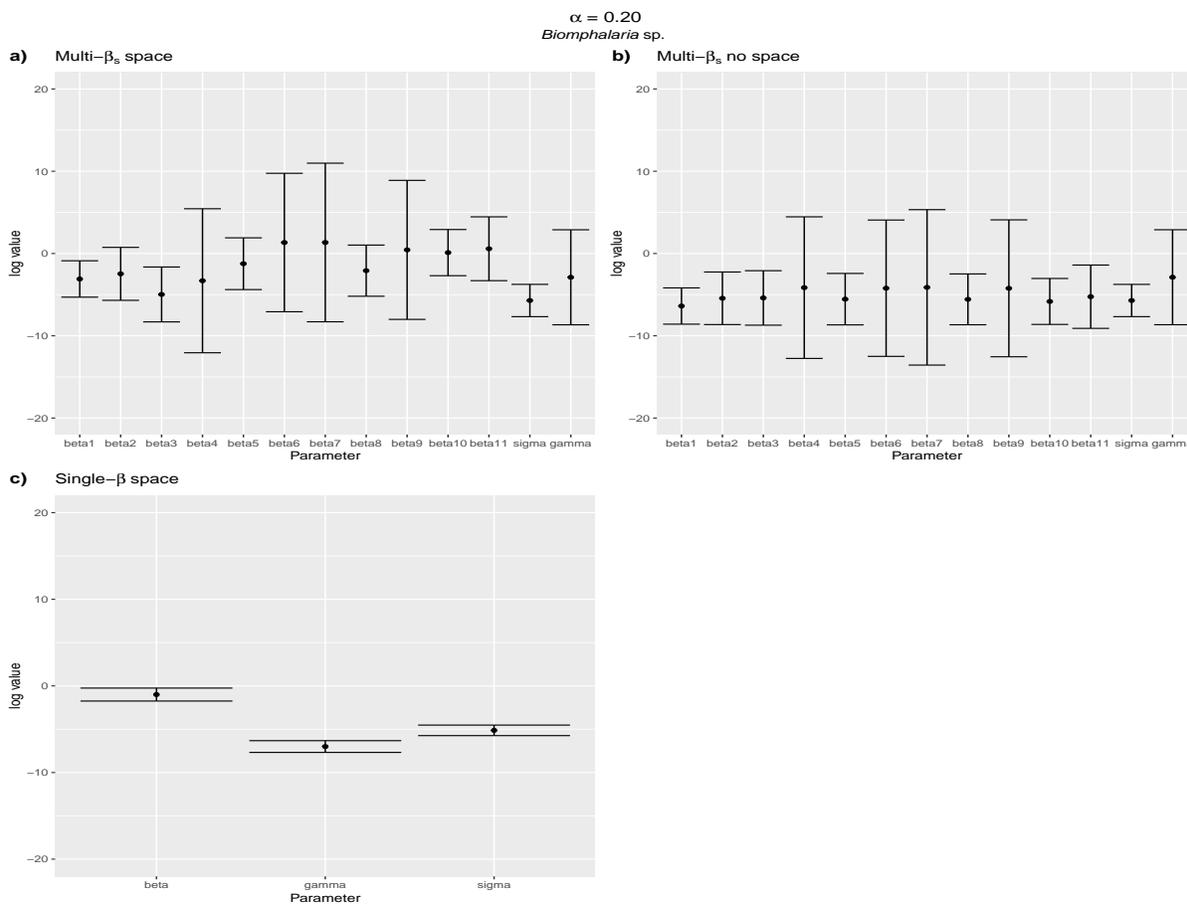


Figure D.29: Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Confidence intervals for parameter estimates for *Biomphalaria* sp. models with SAC prevalence at age 6 set as α set to 0.20 (20%): a) Multi- β_s space sp. b) Multi- β_s no space c) Single- β space d) Single- β no space

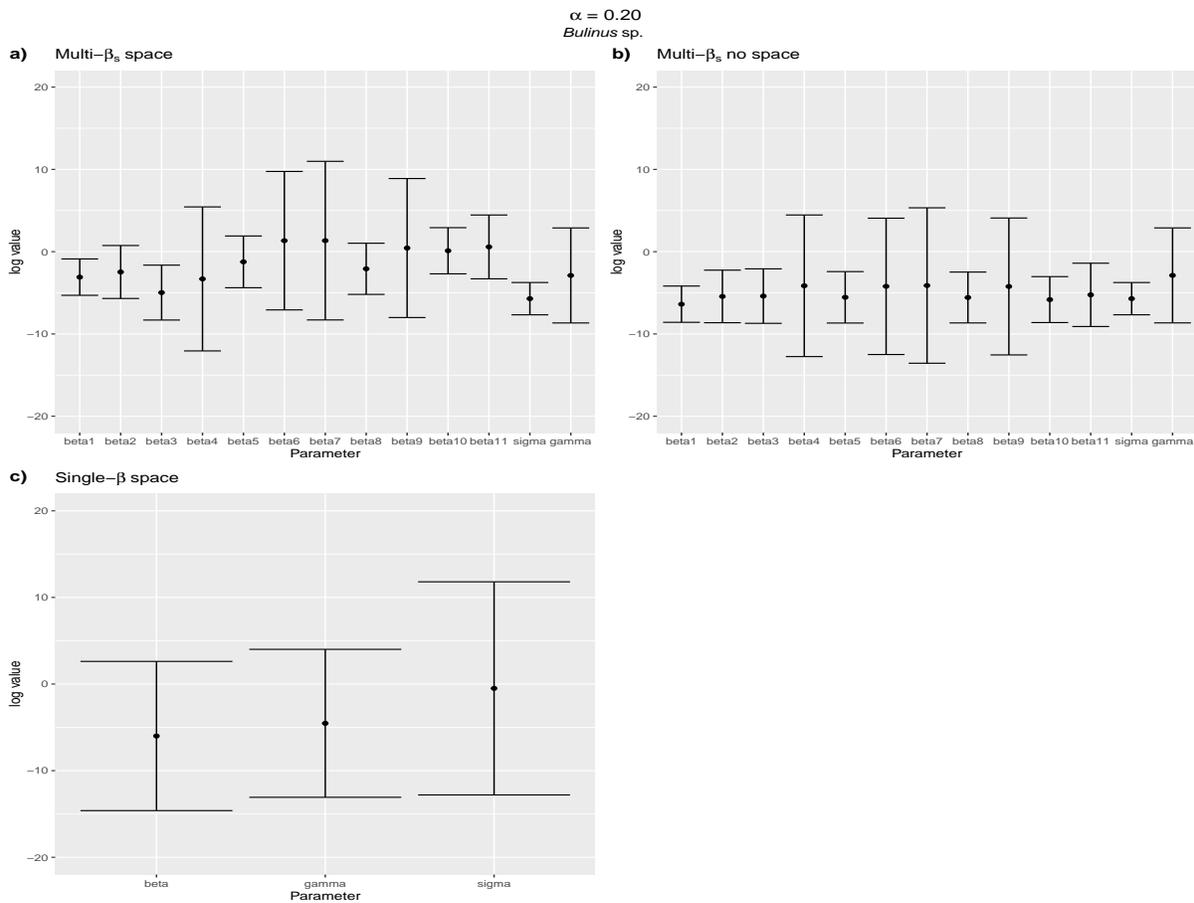


Figure D.30: Prior *Schistosoma* infection prevalence of $\alpha = 0.20$ for SAC age 6. Confidence intervals for parameter estimates for *Bulinus* spp. models with SAC prevalence at age 6 set as α set to 0.05 (5%): a) Multi- β_s space sp. b) Multi- β_s no space c) Single- β space d) Single- β no space

D.2 Log-likelihood profiles

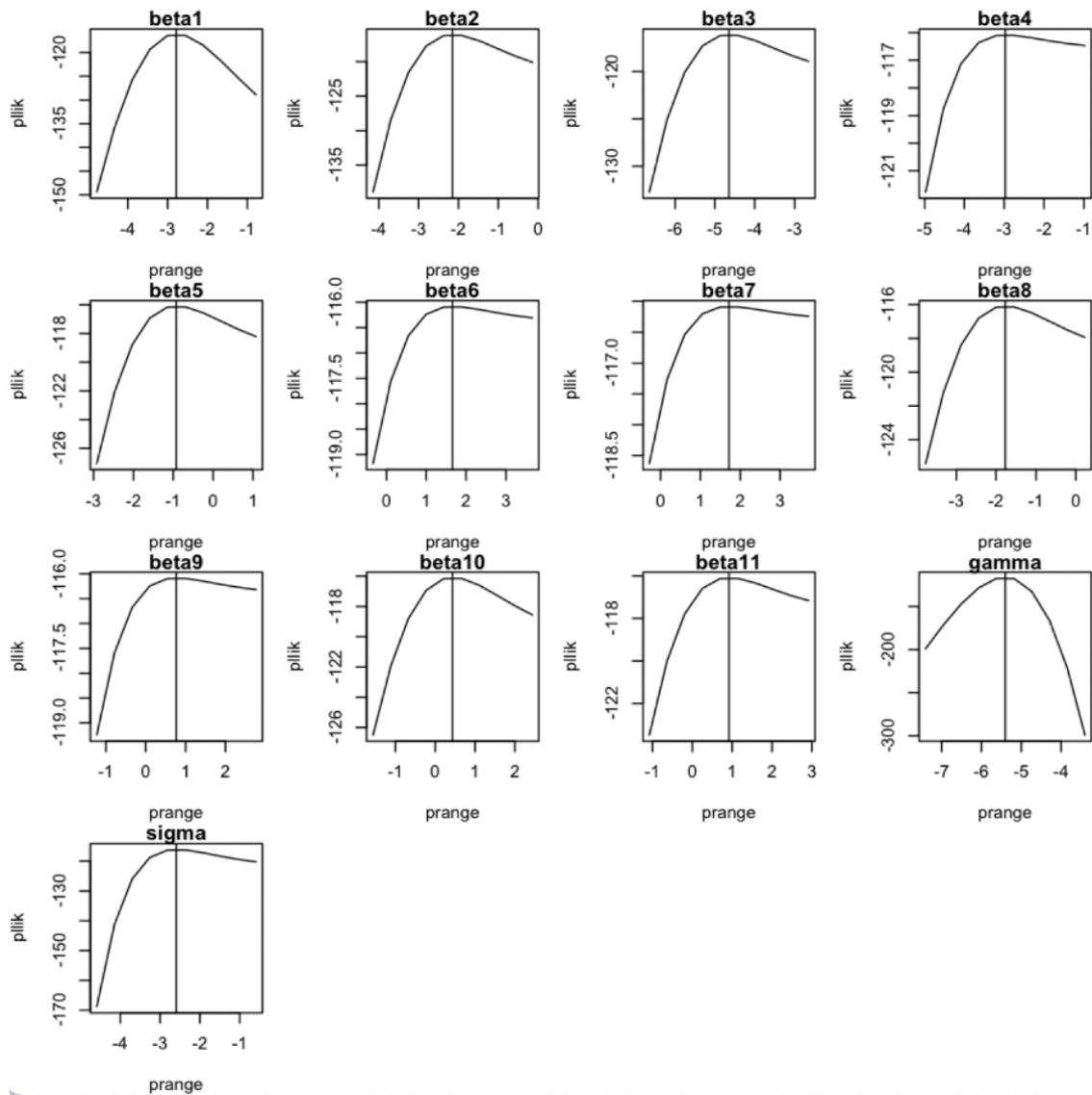


Figure D.31: *Biomphalaria* sp. Multi- β with space effect model profile likelihood against fit of model.

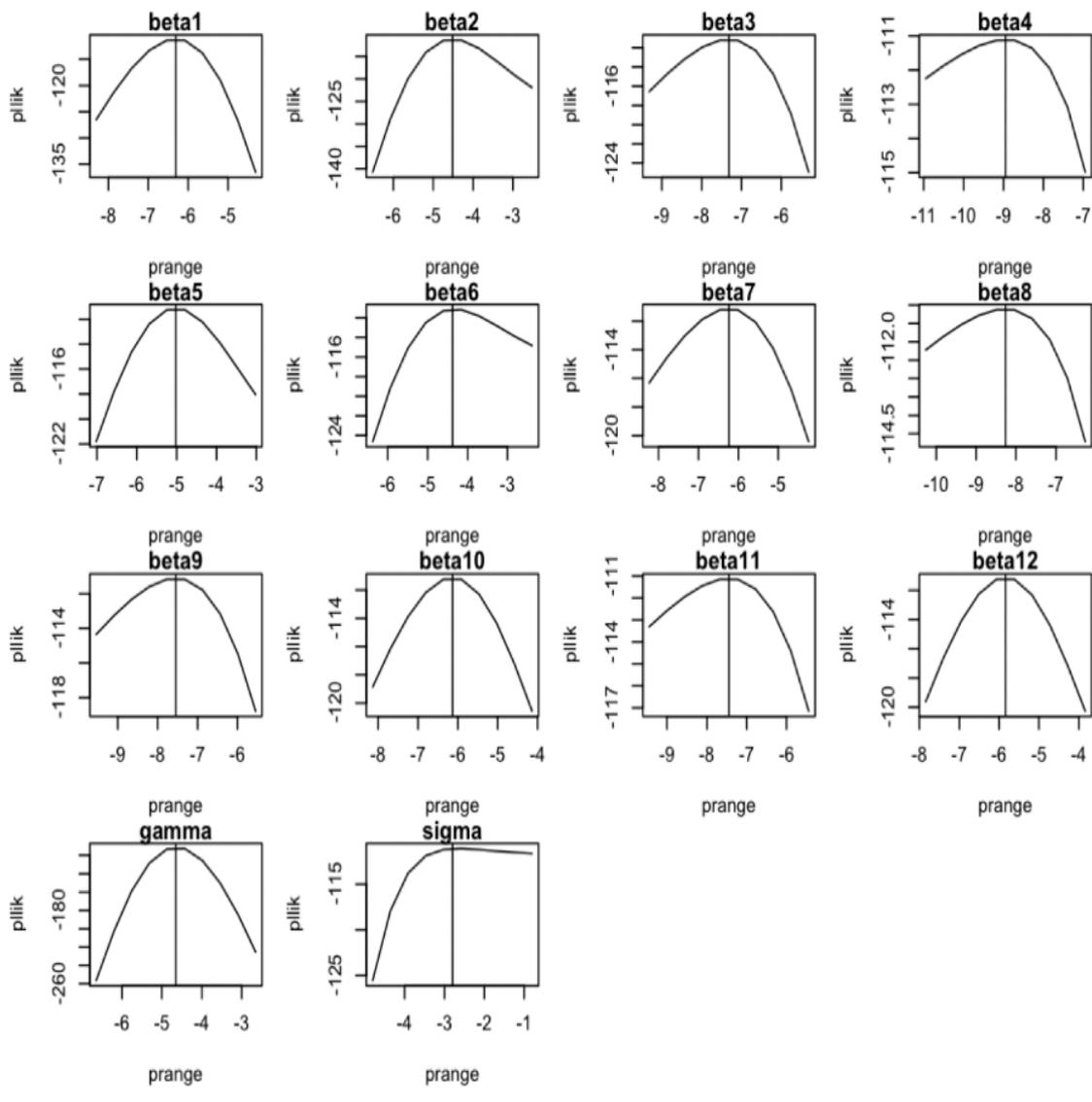


Figure D.32: *Bulinus* sp. Multi- β with space effect model profile likelihood against fit of model

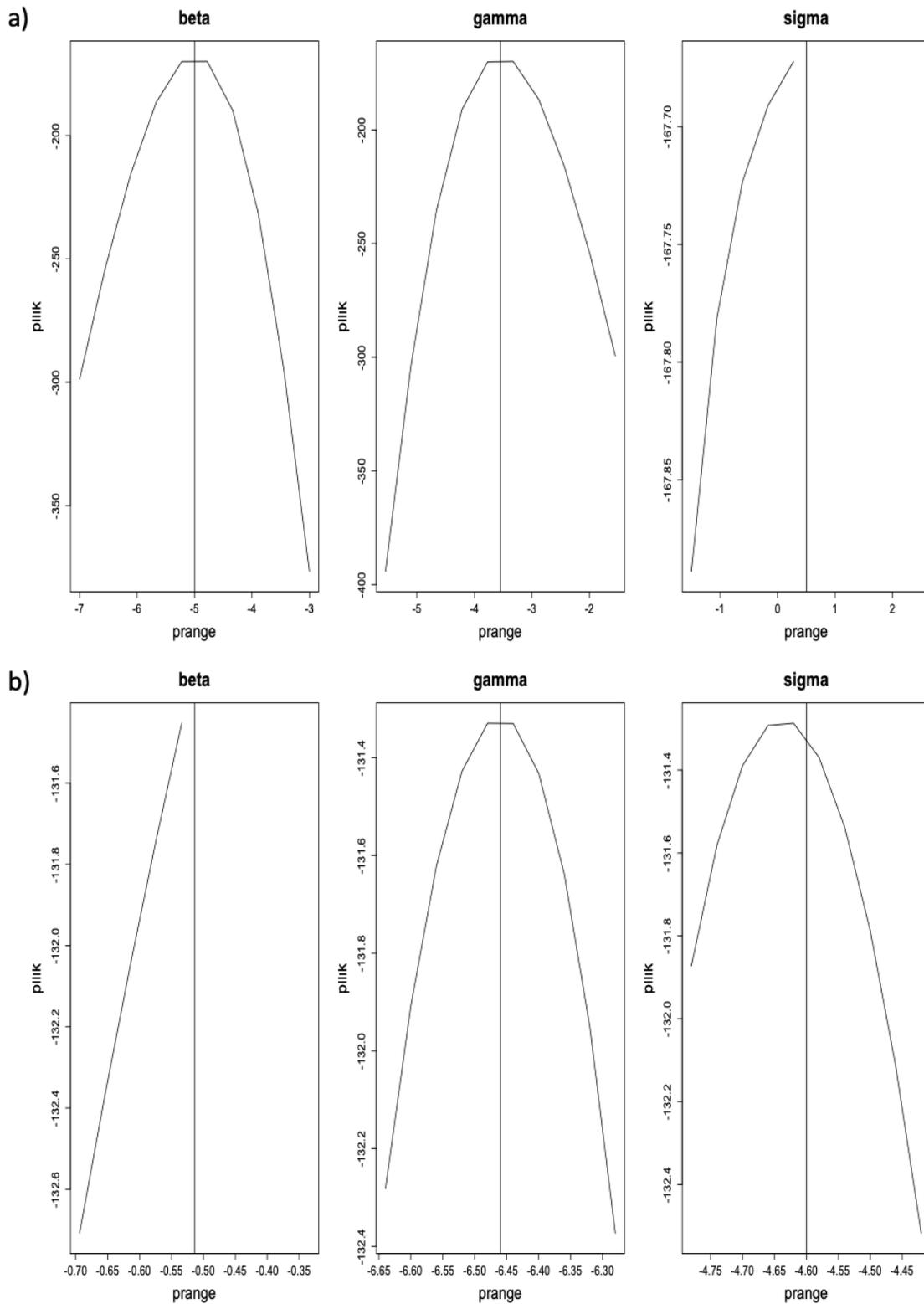


Figure D.33: Single- β with space effect model profile likelihood against fit of model a) *Biomphalaria* sp. b) *Bulinus* spp.

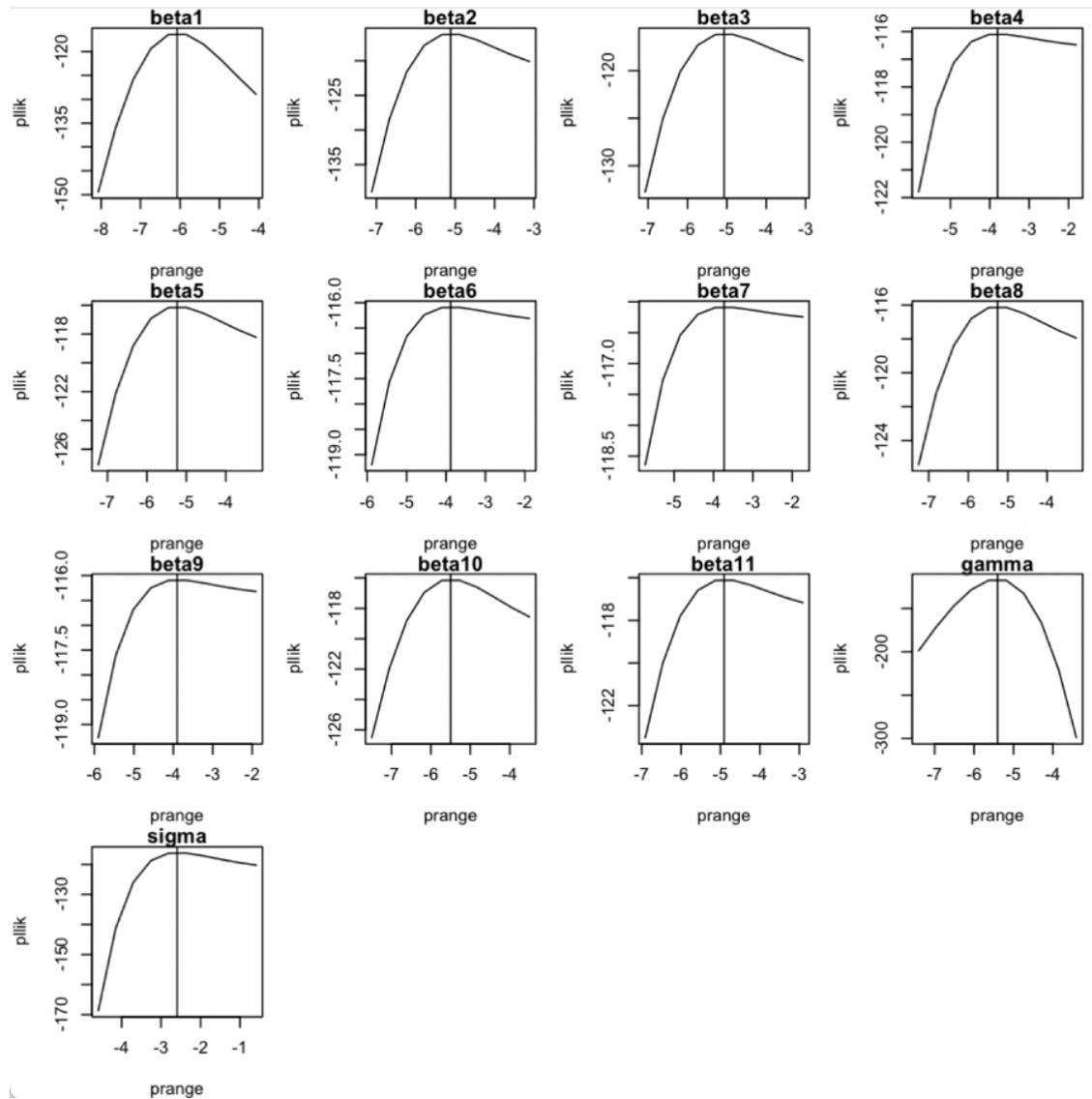


Figure D.34: *Biomphalaria* sp. Multi- β with no space effect model profile likelihood against fit of model.

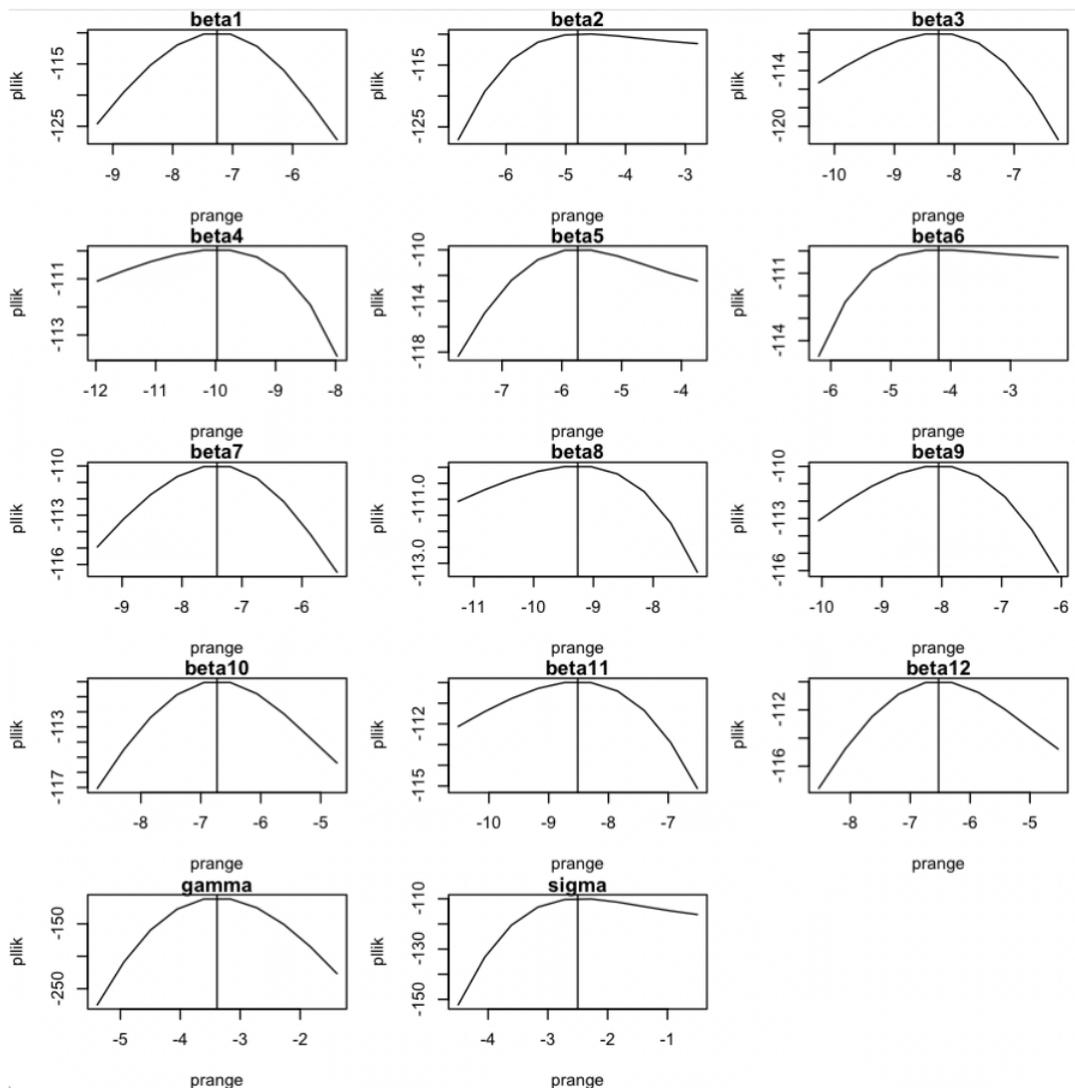


Figure D.35: *Bulinus* sp. Multi- β with no space effect model profile likelihood against fit of model

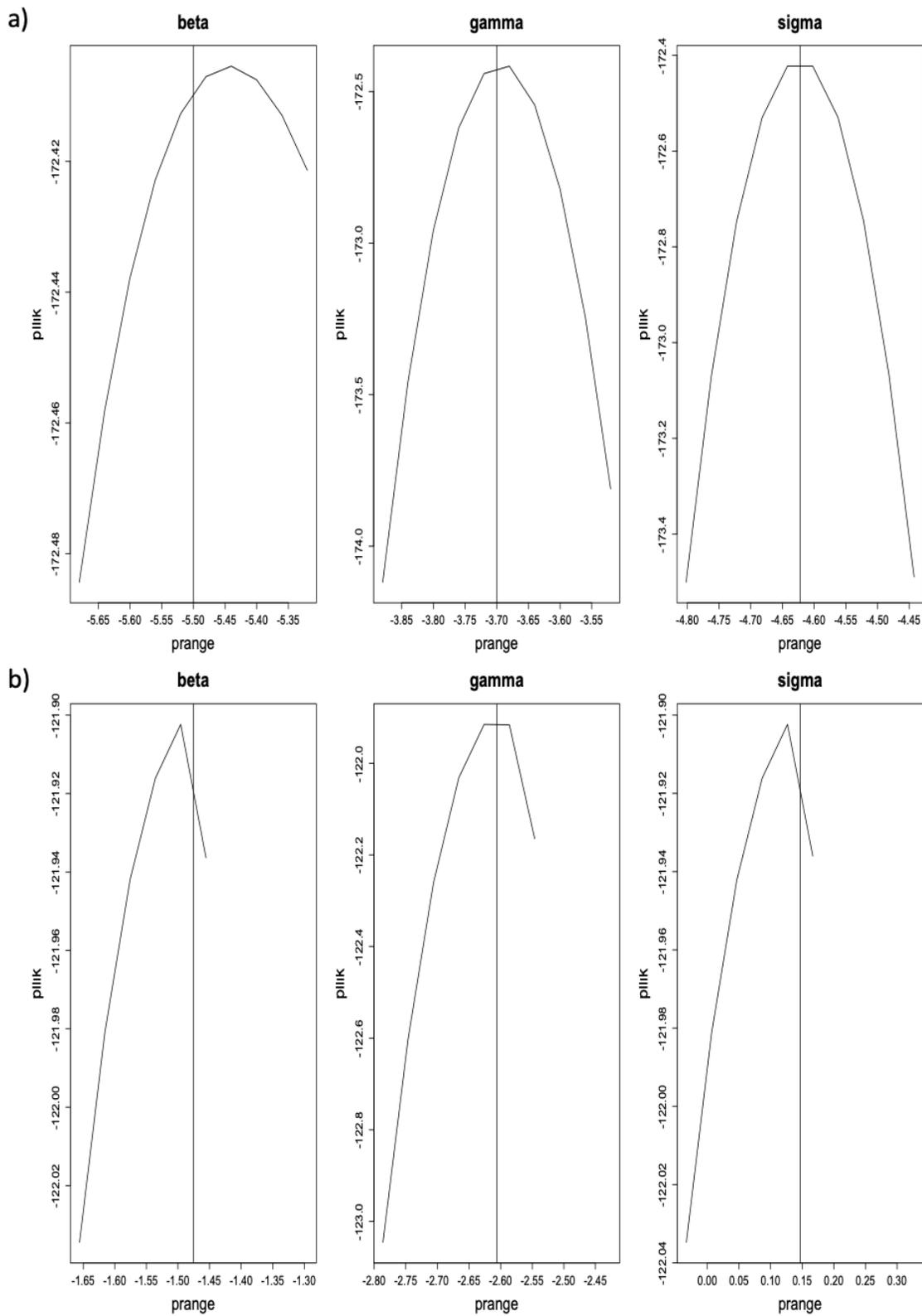


Figure D.36: Single- β with no space effect model profile likelihood against fit of model a) *Biomphalaria* sp. b) *Bulinus* spp.

Appendix E

Dissemination of research



Part A: Intestinal & urogenital schistosomiasis dual-infection focus:
Investigating age-infection relationships for school-aged children along shoreline of Lake Malawi

Amber Lydia Reed



Introduction

Schistosomiasis can cause lifelong morbidity with children most vulnerable to disease. Infection rates are often characterized by local heterogeneities in transmission, which leads to the importance of identifying high risk areas and ways to improve tailored control.

Along the shoreline of Lake Malawi, both intestinal (IS) and urogenital (UGS) schistosomiasis is now occurring from *Schistosoma mansoni* and *S. haematobium* infections respectively, despite annual praziquantel treatment in all schools.

In this poster, I present a secondary analysis of epidemiological data collected from school-aged children (SAC) in Mangochi District, Lake Malawi to assess age-infection profiles.



Model Formulation

Let Y_{ij} be a binary response for individual SAC i at a named school j . This follows Y_{ij} is either $Y_{ij}=1$ if SAC had an infection-positive result or $Y_{ij}=0$ if the SAC had an infection-negative result at named school j . The following equations can represent the distribution of the model as.

$$Y_{ij}|X_{ij} \sim \text{Bernoulli}(p_{ij}),$$

$$p_{ij} = E(Y_{ij} | x_{ij}).$$

These give a logistic regression with Bernoulli distribution and mean p_{ij} . The x_{ij} is a vector of the explanatory variables with the i th subject ($i=1, 2, \dots, n$) with j th school ($j=1, 2, \dots, k$), where n is the number of subjects and k is the sample of schools. We then represent the model as a generalised additive model (GAM),

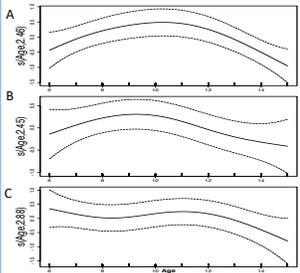
$$\text{logit}(p_{ij}) = \alpha + x_{ij}^T \beta + s(z_{ij}, \phi)$$

where s is the smooth function of z_{ij} given ϕ where z_{ij} denotes the age of the SAC i at school j . GAM allow us to adjust the smoothness of the predictor functions and make the assumption that the predictor relationship is smooth in nature when the true predictor relationship may be more noisy.

Results

GAM smooth term

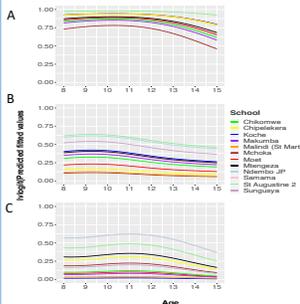
Inverse logit of smooth term for Schistosoma infection association with age of SAC adjusted for age and school



All schools dual-infection of A.S.m. [+], B.S.m. [-], C.S.h. [-]

GAM covariates

Inverse logit predicted values for Schistosoma infection association with age of SAC adjusted for age and school



Dual-infection of A.S.m. [+], B.S.m. [-], C.S.h. [-]

Table 2: IS and UGS GAM of the smooth term age adjusted for school

School	IS			UGS		
	Estimate	SE	CI	Estimate	SE	CI
AIC	88	50	45			
Smooth term for school	0.0022***	0.011	0.015			
Explanatory variables						
Sex						
Female	0.70*	[0.036, 1.51]	1.69**	0.718, 2.54	1.79**	[0.01, 2.47]
Male	0.76*	[-0.025, 1.62]	0.85*	[-0.162, 1.47]	0.94*	[-0.020, 0.198]
Koche	1.25*	[0.44, 2.01]	1.73**	[0.79, 2.68]	2.65*	[-0.48, -0.167]
St Augustine 2	2.47*	[0.374, 4.60]	2.64**	[1.35, 3.93]	1.9*	[0.271, 3.12]
Mchoka 2	0.62	[-0.465, 1.71]	2.23**	[1.07, 3.48]	1.56**	[0.462, 2.67]
Chipelakera	1.67*	[0.184, 2.73]	2.21**	[1.07, 3.48]	0.10	[-1.25, 0.972]
Malindi (St Martins)	1.67*	[0.184, 2.73]	0.10	[-1.25, 1.10]	4.02	[-3.09, 0.102]
Chitosema	0.68	[-0.262, 1.66]	1.2	[0.26, 2.12]	0.70	[-2.07, 0.728]
Chitosema	1.67*	[0.184, 2.73]	0.41	[-0.40, 1.27]	0.40	[-0.546, 1.44]
Mlangeni	0.68	[-0.262, 1.47]	1.57*	[0.48, 2.62]	1.8	[-2.23, 0.392]
Munguwa	0.68	[-0.169, 2.102]	1.67*	[0.580, 2.70]	0.65	[-0.314, 1.43]

Summary of Results

- Six schools (SA, Koche, St Augustine 2, Sungusya, Malindi (St Martins), Chipelakera) all had significant evidence to suggest as SAC aged the log odds of being positive for *S.m.* increased compared to MC school.
- Koche school had significant evidence to suggest as SAC the log odds of being positive *S.h.* decreases and St Augustine 2 and Ndembu increases compared to MC school.
- S.m.* infection [+T+] smooth term prediction goes from negative to positive versus age up to age 11 and before decreasing back to negative.
- For *S.m.*, St Augustine 2 had the highest and Sungusya school had the lowest prevalence rates.
- For *S.h.*, there was no clear pattern for prevalence versus age for SAC at all the schools.

Conclusions

There is an increasing prevalence of IS in SAC up to around 11 years before decreasing there afterwards. By contrast, no clear age-infection pattern for UGS was found. Peak of infection is expected around adolescence. This peak in prevalence at age 11 may be due acquired immunity which is known as 'peak shift' phenomenon or could be explained by factors for instance water exposure³. A further investigation (Part B) was carried out to study the role of IS and UGS co-infection impact on disease dynamics and whether the age-profiling of infection changes.

Acknowledgements

With thanks to Christopher Jewell, Russell Stothard and Michelle Stanton for her guidance with the project and R programming.

1. Kayuni, S.A., O'Ferrall, A.M., Baxter, H., Heskelh, J., Mainga, B., Lally, D., Al-Harbi, M.H., LaCourse, E.J., Juziwele, L., Musaya, J. and Makaula, P., 2020. An outbreak of intestinal schistosomiasis, alongside increasing urogenital schistosomiasis prevalence, in primary school children on the shoreline of Lake Malawi, Mangochi District. *Malawi Infectious Diseases of Parasites*, 9(1), pp.1-10.

2. Muhammad H Alharbi et al. "Biomphalaria pfeifferi Snails and Intestinal Schistosomiasis, Lake Malawi, Africa, 2017-2018". In: 25:3 (2019), pp. 613-615

3. M. EJ Woodhouse. "Patterns in Parasite Epidemiology: The Peak Shift". In: *Parasitology Today* 14:10 (1998), pp. 428-434.

Figure D.1: Poster Presented at British Parasitology Conference (BSP) conference 2021. Part A: Intestinal and urogenital schistosomiasis dual-infection focus: Investigating age-infection relationships for school-aged children along shoreline of Lake Malawi

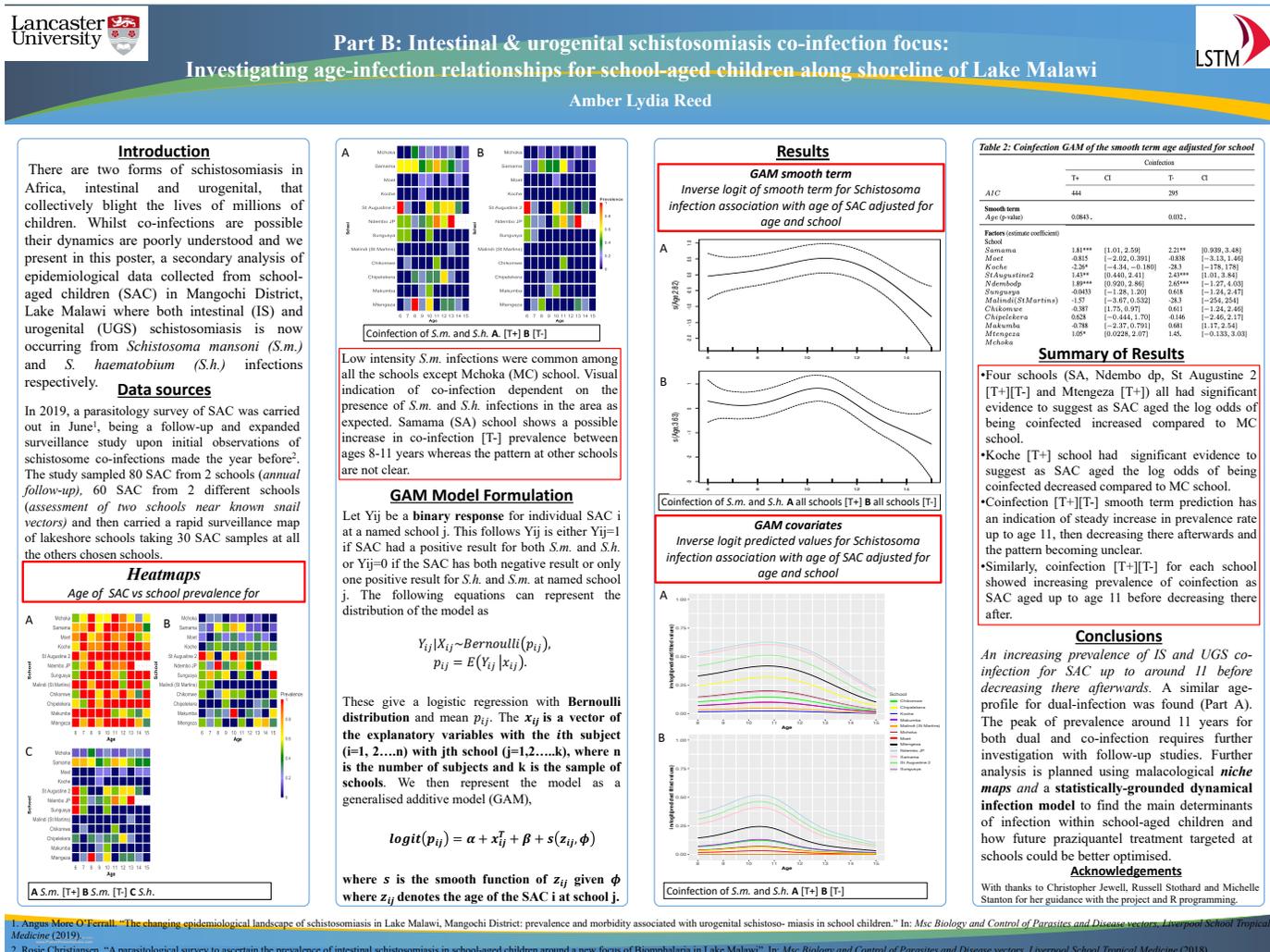


Figure D.2: Poster presented at BSP conference 2021. Part B: Intestinal and urogenital schistosomiasis co-infection focus: Investigating age-infection relationships for school-aged children along shoreline of Lake Malawi

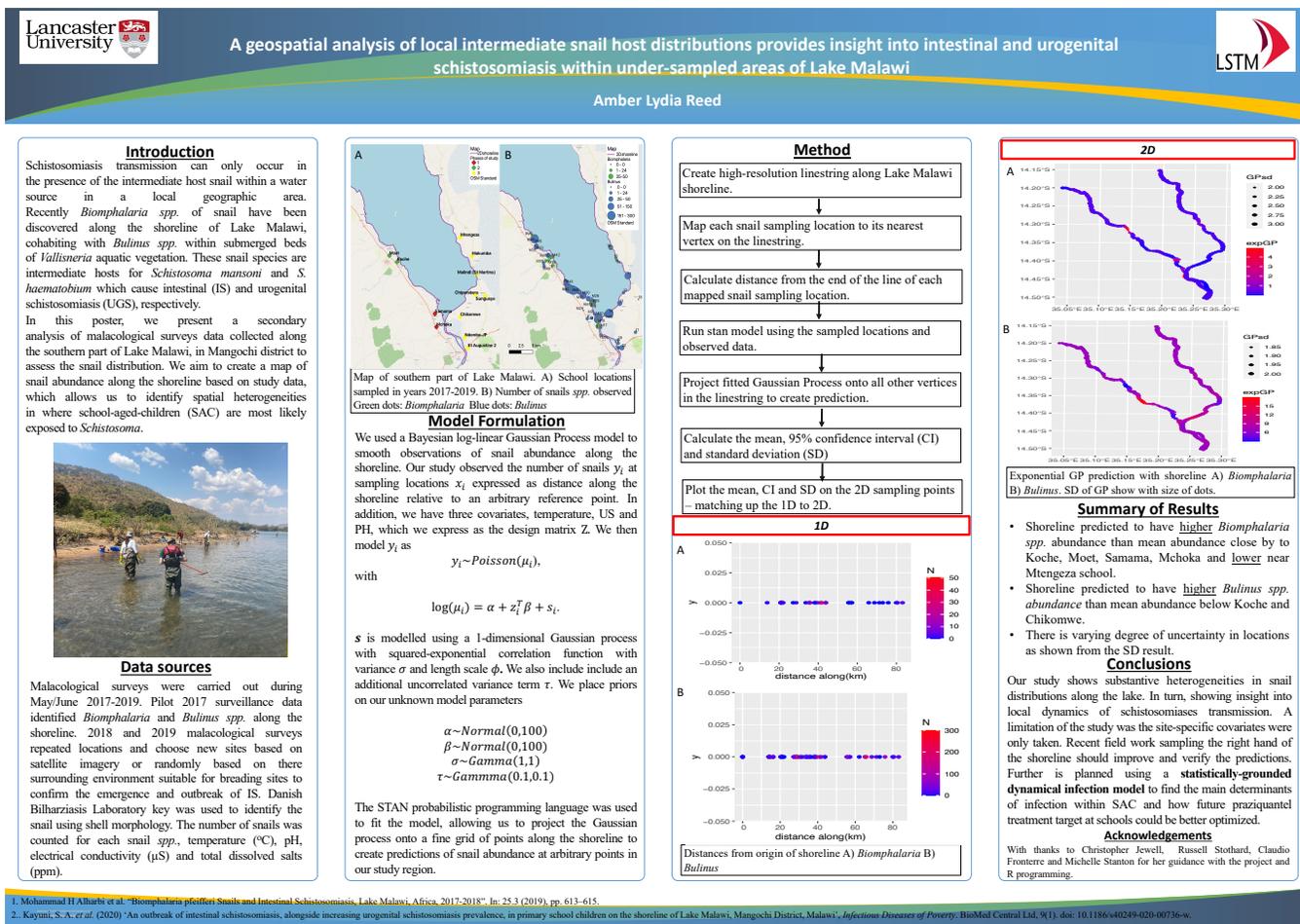


Figure D.3: Post presented at BSP conference 2022. A geospatial analysis of local intermediate snail host distributions provides insight into intestinal and urogenital schistosomiasis within under-sampled areas of Lake Malawi