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

Amber L. Reed[1*], Mohammad H. Al-Harbi[2,3], Peter Makaula[4], Charlotte Condemine[2], Josie Hesketh[2], John Archer[2], Sam Jones[2], Sekeleghe A. Kayuni[2,4], Janelisa Musaya[4], Michelle C. Stanton[5], J. Russell Stothard[2], Claudio Fronterre[1] and Christopher Jewell[6]

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). Analyses of all other environmental data were considered non-significant.

Conclusions The spatial range in which interpolation of snail distributions is possible appears < 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.

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

*Correspondence:
Amber L. Reed
amberreed21@hotmail.com
Full list of author information is available at the end of the article

Reed *et al. Parasites & Vectors*   (2024) 17:272

Page 2 of 13

Background

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* occur. While various species of *Bulinus* are present in Lake Malawi, with *Bulinus globosus* and *B. nyassanus* responsible for *Schistosoma haematobium* transmission, only in 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 [1–3].

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 in attempt to define the extent of schistosomiasis, particularly intestinal schistosomiasis (IS). These surveys demonstrated the need for further surveillance of freshwater snails, alongside emphasis upon updated and tailored interventions and policies for control of schistosomiasis in this lacustrine setting [1–3]. 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 [4]. 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 [5]. 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 [5–7]. Changes in climate and seasonal patterns are therefore likely to alter transmission of schistosomiasis over both space and time, increasing the need for identification of snail habitats to target appropriate control interventions [1]. However, although snail distribution within a geographical area can be measured through malacological surveillance, 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 [8] and is vital for those using it for irrigation, agriculture, water supply, fishing industries and tourism [9]. Due to the lack of adequate sanitation in Malawi, human urine and faecal materials continuously contaminate the shoreline facilitating the transmission of schistosomiasis, amongst other water-borne pathogens [10]. 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) [11, 12]. Lower temperatures and higher winds are reported on the eastern side [13], with low-lying areas such as the upper Shire River margins vulnerable to flooding [14]. More broadly, the climate of this southern part of the shoreline is affected by the migration of the Inter-Tropical Converge 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 [15, 16]. Rainfall is dependent on altitude and time of the year [17]. Lake water levels vary over time and are at their highest during wet season, which also affects evapotranspiration and outflows to the Shire River [2, 14]. Most important perhaps is an increasing human and livestock population which is leading to more frequent water contact, enhancing opportunities for transmission of schistosomiasis [2, 4].

The World Health Organisation (WHO) has supplied new guidelines to target elimination of schistosomiasis by reducing freshwater snail abundance, thus interrupting transmission [18]. Identifying locations where freshwater snails are most abundant therefore aids targeted control methods, preventing initial infection and re-infection and hence helping eliminate or reduce transmission [19–21].

Here, we undertook a secondary analysis of primary malacological data first reported by Al-Harbi et al. [1] and Kayuni et al. [2]. 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.

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.

Reed *et al. Parasites & Vectors*   (2024) 17:272

Page 3 of 13

Snail abundance

The primary dataset reported in Al-Harbi et al. [1] and Kayuni et al. [2], which this secondary analysis is based on, originally collected malacological surveys between 2017 and 2019 as shown in Fig. 1 and available at Additional file 1: Dataset. 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 1b shows a map of sampling sites, together with their relationship to primary schools in the region, demonstrating the importance of human proximity to the lake shore and hence potential for exposure to infected 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 values and assumed these values to be the closest lowest value, e.g. 300. The recorded sites considered in our study are shown in Fig. 1c and d.

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 Fig. 2. Rainfall (millimetres, mm) estimates were extracted between 1 November 2017 and 30 June 2019 from Tropical Applications of Meteorology using SATellite data and ground-based observations (TAMSAT) with a monthly frequency at 4-km resolution [22–24]. Land surface temperature (LST) (°C), evapotranspiration and Normalised Difference Vegetation Index (NDVI) raster data were obtained from Land Processes Distribution Active Archive Center (LPDAAC) [25–28]. LST data between 1 November 2017 to 30 June 2019 were
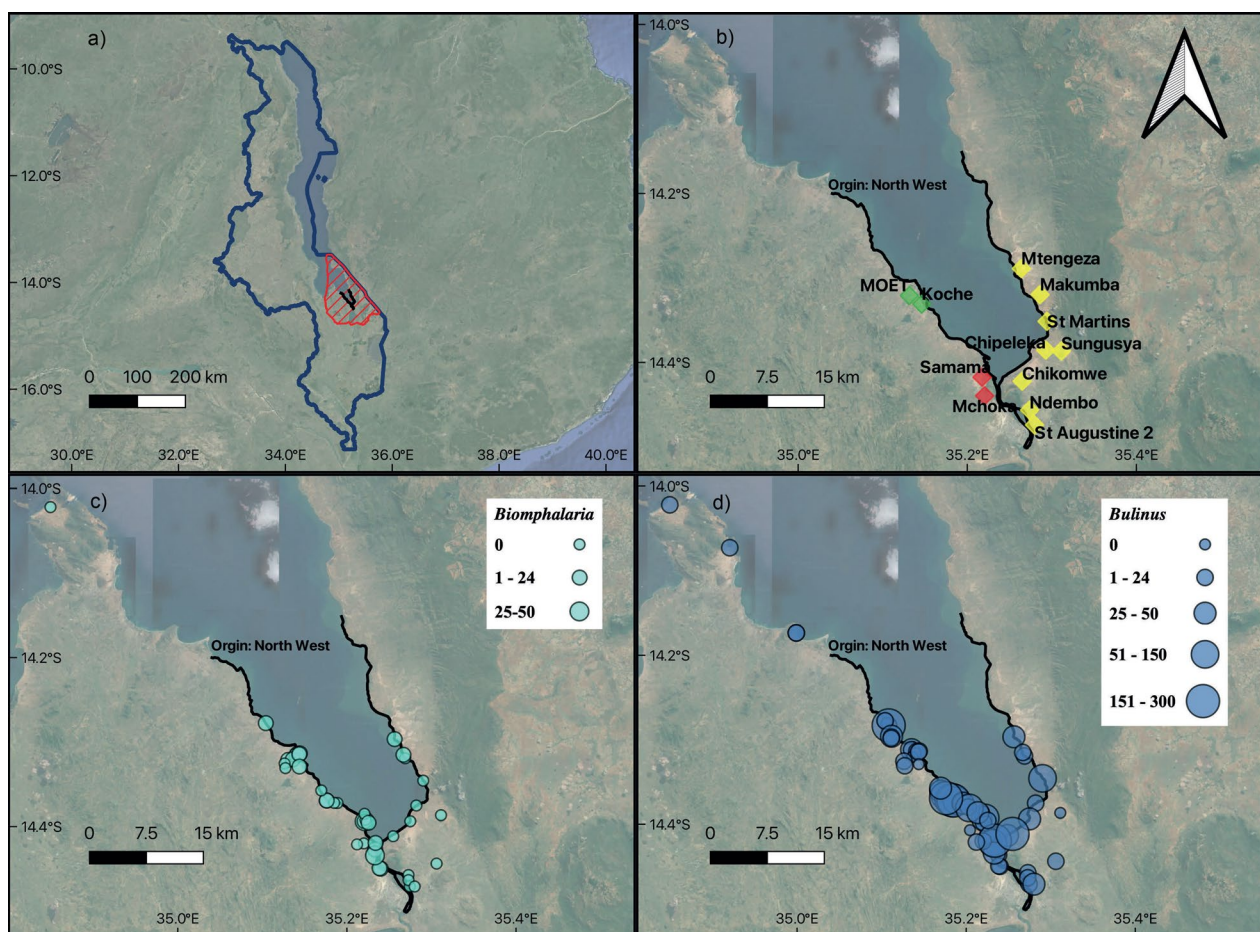

Fig. 1 Primary dataset collected data. **a** Map of Malawi in dark blue. Red crossed: study area; black line: prediction points. Parasitological surveys: **b** Primary school locations along the shoreline. Malacological surveys: **c** observed *Biomphalaria* sp. snails; **d** observed *Bulinus* spp. snails
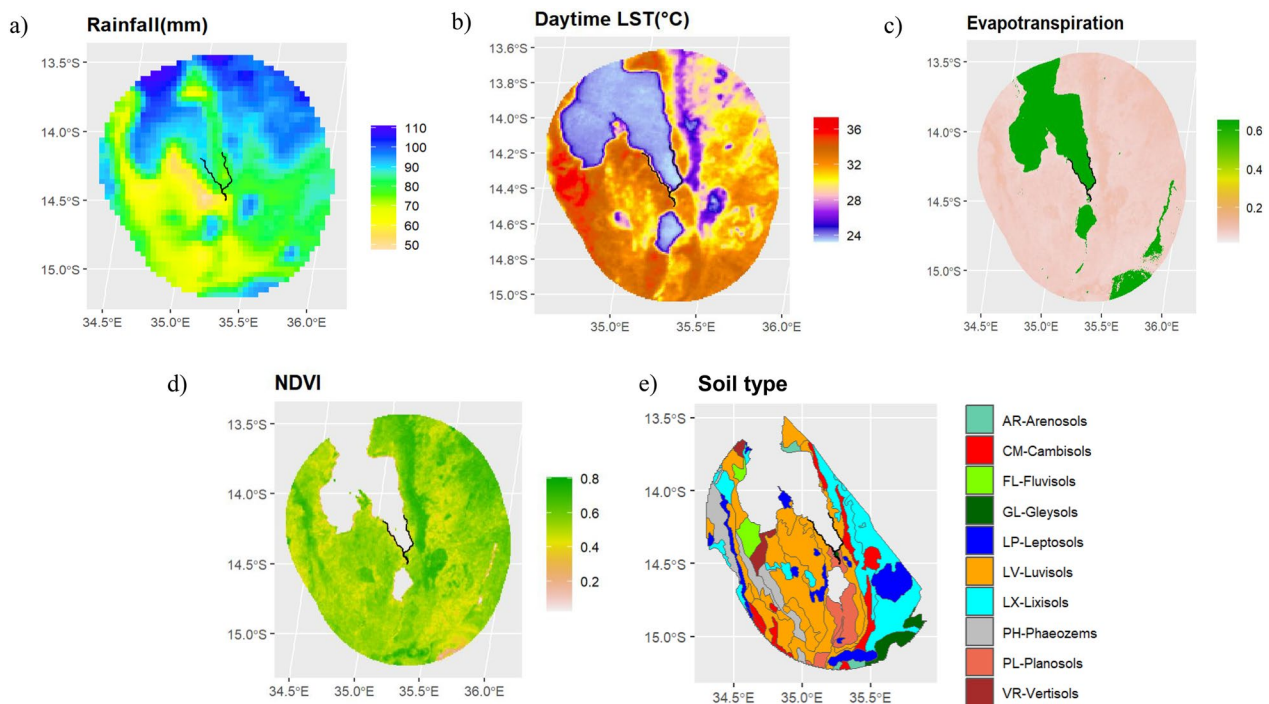
Reed *et al. Parasites & Vectors*   (2024) 17:272

Page 4 of 13

Fig. 2 Raster plot of extracted covariate data. **a** Rainfall (mm), **b** daytime land surface temperature (LST) (C), **c** evapotranspiration, **d** normalized difference vegetation index (NDVI). **e** Soil types for southern part of Lake Malawi shoreline, adapted from Dijkshoorn et al. [29]. Black line shows the shoreline template from which covariate values were extracted

extracted from Moderate Resolution Imaging Spectro-radiometer (MODIS)/Terra LST/Emissivity 8-Day L3 Global 1 km SIN Grid raster (MOD11A2v061) [25]. 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 500 m SIN Grid raster (MOD16A3GFv061) [26]. NDVI data between 1 November 2017 to 30 June 2019 were extracted from Modis/Terra vegetation indices 16-Day L3 Global 1 km SIN Grid raster (MOD13A2v0061) [27]. 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 1:1 m [29]. After extracting the values, the temporal covariates were aggregated by taking the mean of the values over the time frame.

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 (Fig. 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 (Additional file 2 Fig. S1 and Fig. S2). Additionally, we sub-sampled the 4000 vertices at equal intervals to a set of 200 prediction points.

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 1-km 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

Reed *et al. Parasites & Vectors*   (2024) 17:272

Page 5 of 13

as the first sampling point, the next collected value was taken.

One-dimensional Poisson latent Gaussian process regression

Bayesian multilevel model

A Bayesian Poisson multilevel model (BMLM) with a Gaussian latent process (GP) was developed using STAN programming language, 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 assumed 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 at the sum of spatially correlated variance (using quadratic, exponential, or Matérn) κ uncorrelated (or nugget) variance [30]. 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 (Crl)) 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 supplementary information (Additional file 3: Model Formulation) for a mathematical explanation of the model.

Results

Observed data

After cross-checking the observation data, as shown in Fig. 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) °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) °C. Additional file 3 shows the observed data for 1D (Additional file 4: Fig. S1) and 2D (Additional file 4: Fig. S2 and Fig. S3). A histogram of the centred and scaled covariates is shown in Additional file 5: Fig. S1.
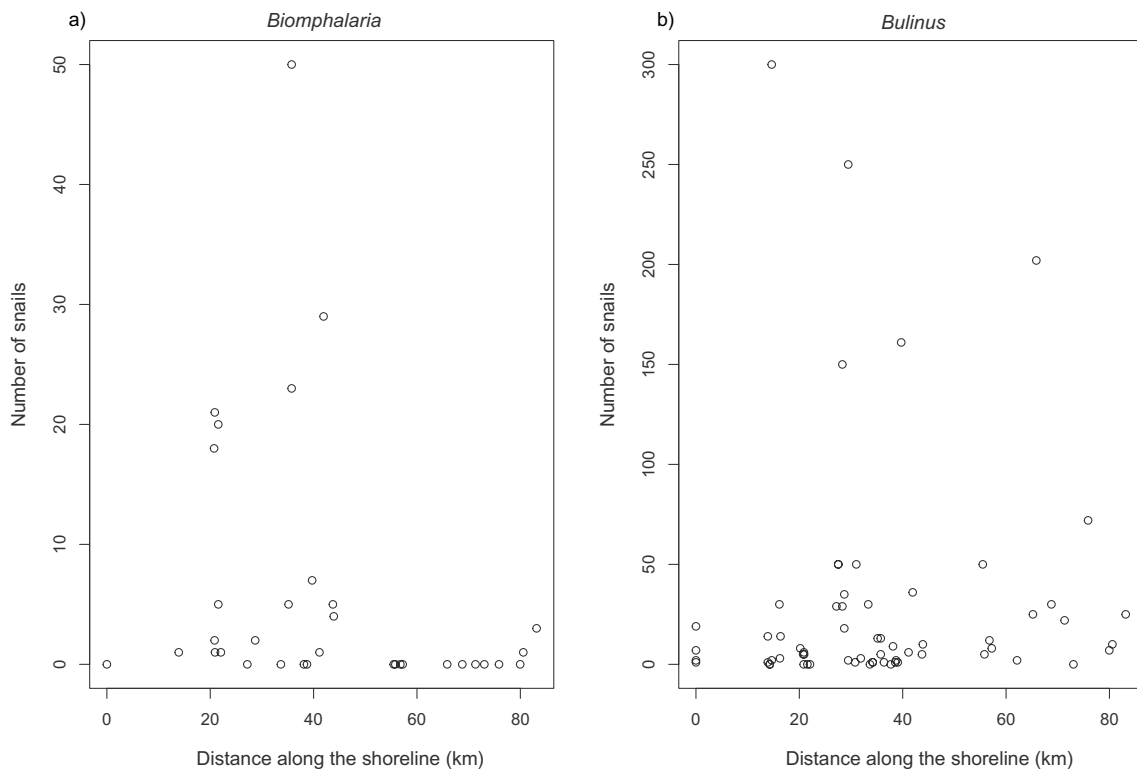

Fig. 3 Scatter plot of absolute snail numbers observed at sampling points versus distance along the shoreline in km. **a** *Biomphalaria* sp.; **b** *Bulinus* spp

Reed *et al. Parasites & Vectors*   (2024) 17:272

Page 6 of 13

Environmental data prediction points

The extracted environmental data prediction point ranges were: rainfall (59.82–90.37mm), LST [24.68–32.46 (°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 the eastern shoreline having the most rainfall and lowest LST (°C) compared to the western shoreline. Luvisolic (LV) soil type was absent around the River Shire compared to Gleysolic (GL) soil type; Planosolic (PL) soil type was present at the entrance to the River Shire and south of the River Shire compared to GL soil type. The distributions of the values of the environmental covariates in Fig. 4 can be viewed in Additional file 5: Fig. S2.

Covariance function comparison

As shown in Additional file 6: Fig. S1, 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 the shoreline for either *Biomphalaria* sp. or *Bulinus* spp. as shown in Additional file 6: Fig. S2. This suggests that the effect of

the covariates (environmental data) is more prominent than in the Gaussian process.

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 in Additional file 7: Fig. S1 and Additional file 8: Fig. S1.

Covariate effects

Figure 5 shows the posterior distributions for the environmental covariate effects (on the log scale) for each species of snail, with mean snail abundance at location i on the x axis, with 95% CrI filled. As shown in Fig. 5 and Table 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 μ_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 μ_i mean *Biomphalaria* sp. abundance changes by -1.42 (CrI -3.09, 0.10) (reduction). For a 1-SD increase in rainfall, the log μ_i mean *Bulinus* spp. abundance changes by -0.88 (CrI -2.15, 0.33) (reduction). For a 1-SD increase in LST, the log

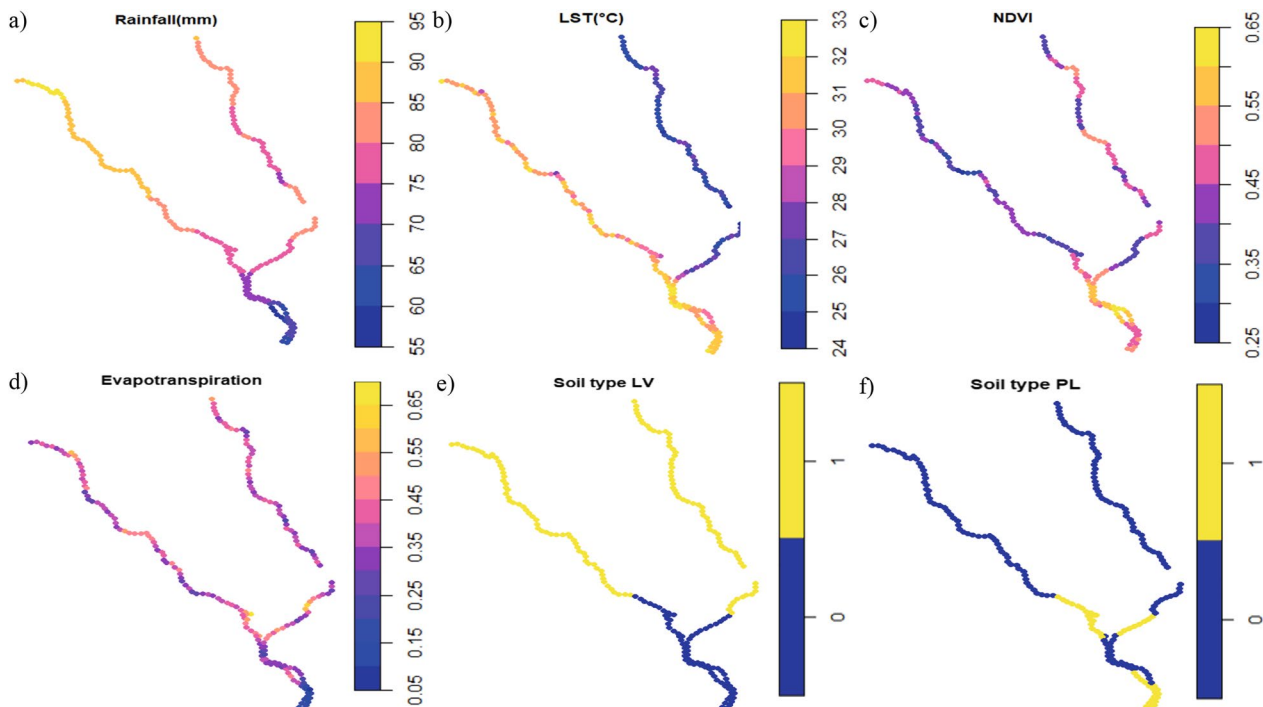**Fig. 4** Environmental data values extracted for each prediction point. **a** Daytime land surface temperature (LST) (°C), **b** rainfall (mm), **c** evapotranspiration, **d** Normalized Difference Vegetation Index (NDVI), **e** Luvisolic (LV) soil type, **f** Planosolic (PL) soil type. **e** and **f** are compared with Gleysolic (GL) soil type. Gap in shoreline is due to the removal of CM soil type
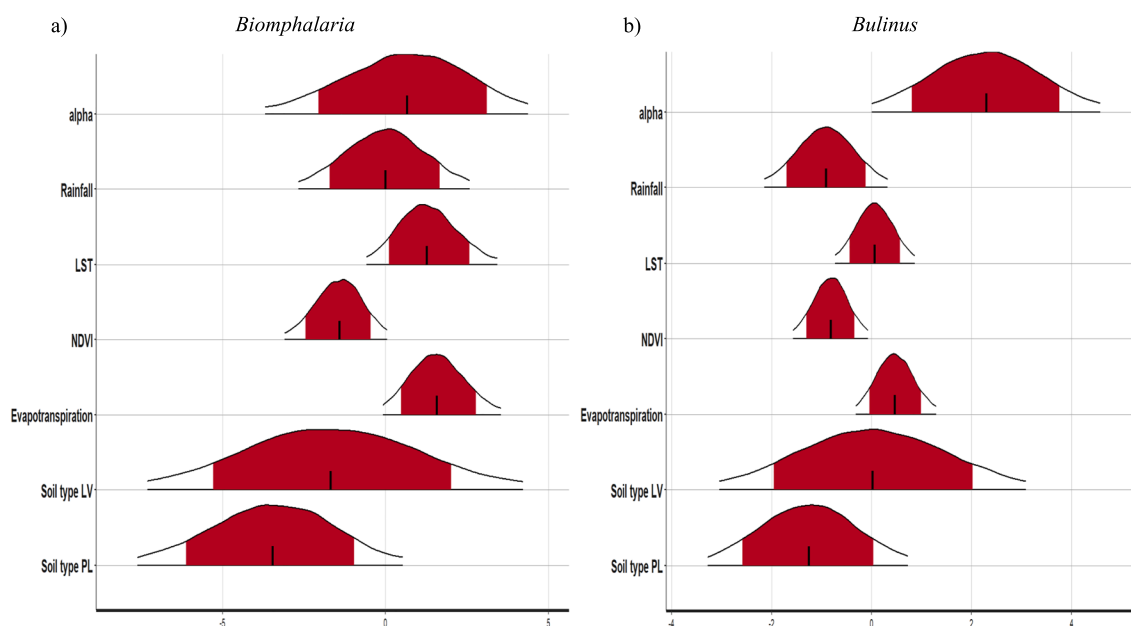
Fig. 5 Posterior plot for each species **a** *Biomphalaria* sp.; **b** *Bulinus* spp. Red shaded area represents the 80% credible intervals (CrI) and the extent of the curve is the 95% CrI

Table 1 Estimated parameter values, mean and CrI

Parameter	*Biomphalaria*		*Bulinus*	
	Mean	95% CrI	Mean	95% CrI
α	0.46	[− 3.98,4.40]	2.32	[$3.69\,e^{-03}$,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				

CrI credible interval, *LST* land surface temperature, *NDVI* Normalized Difference Vegetation Index, *LV* Luvisols, *PL* Planosols, *GL* Gleysols

μ_i mean *Biomphalaria* sp. abundance changes by 1.30 (CrI − 0.56, 3.4) (increased). For a 1-SD increase in evapotranspiration, the log μ_i mean *Biomphalaria* sp. abundance changes by 1.61 (CrI − 0.04, 3.56); similarly, the log μ_i mean *Bulinus* spp. abundance changes by 0.46 (CrI − 0.34, 1.28) (increased). For an increase of PL soil type compared to the baseline GL soil type, the log μ_i mean *Biomphalaria* sp. abundance changes by − 3.48 (CrI − 7.13, 0.52) (reduction); similarly, the log μ_i mean *Bulinus* spp. abundance changes by − 1.29 (CrI − 3.29, 0.71) (reduction). No association could be found for LV soil compared to GL soil type.

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, Fig. 6; 1D version, Additional file 9: Fig. S1).
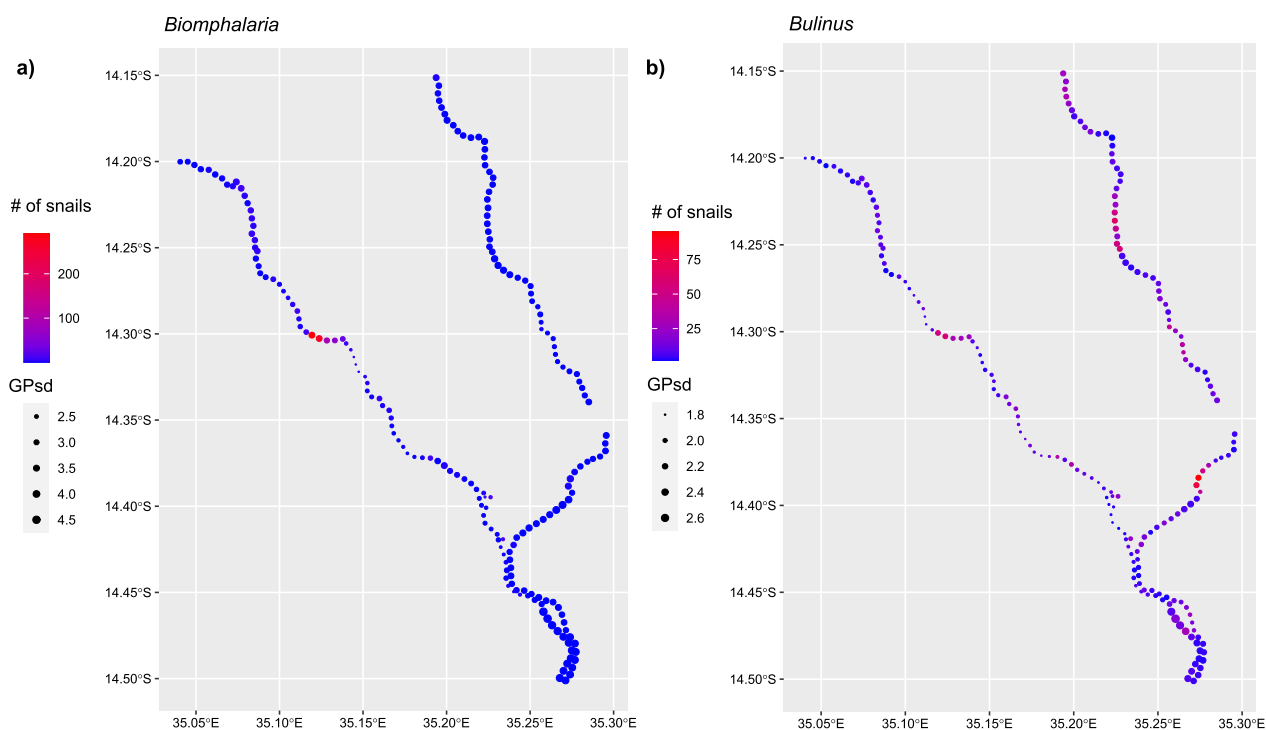
Reed *et al. Parasites & Vectors*   (2024) 17:272

Page 8 of 13

Fig. 6 Two-dimensional mean Gaussian process (GP) prediction of number of snails log($\hat{u}_i$) 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

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 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 (Fig. 7) 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 [31–33]. 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 [34, 35].

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, disrupting freshwater snail habitats [32]. Second, an increase in

Reed *et al. Parasites & Vectors*   (2024) 17:272

Page 9 of 13

Fig. 7 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 the southern part of the Lake Malawi shoreline (unpublished). Pictures taken by Alexandra Juhasz

rainfall has been reported to increase turbidity of water and in turn decrease the presence of snails (through disrupting their habitat) [32, 33]. 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 snails, causing an overall reduction in snail abundance [35].

In contrast to our result, some cases studies have found increase in snail abundance during increase in rainfall. For instance, when excess rainfall, known as 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 [5]. Consequently, flooding can change the human-snail contact interplay, through an indirect effect on human behaviour, and thus the associated risk of schistosomiasis transmission [36, 37]. 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 [31, 36]. 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 [2]. This could impact the snail abundance and its presence spatially and temporally and indirectly affect human behaviour as mentioned before [2, 36, 37], for example, if the lake levels are regulated by needs for hydroelectricity or because many individuals prefer to make contact with shallower and more safe areas of the lake [38, 39].

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 determine the optimal temperature for snail survival. For *Biomphalaria* sp. snails the optimum temperature has been found to between 15 and 30 ℃, where there is a decrease in snail abundance above 30 and 35°*C*, and no snails survive above 35 ℃ [37, 38].For our prediction points along the shoreline, the LST ranged between 25 and 32 ℃, which suggests *Biomphalaria* sp. snail abundance still increases above 30 ℃. This difference could be due to it being in a natural environment where snails are able to adapt to climate change [40]. It has also been reported that freshwater snails move further into the lake when temperatures increase, which we did not consider

Reed *et al. Parasites & Vectors*   (2024) 17:272

Page 10 of 13

in our model because it was constrained to the shoreline and buffer area [41].

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 been investigated in other studies [5].

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 are muddy when rainfall occurs (become water-logged) [42]. A previous study by Koch et al. [7] found the opposite results 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 [7]. The difference between PL and GL soils is that GL is known for its iron reduction [43]. Kulina et al. [44] reported an increase in transmission of snail risk in groundwater with higher iron concentration [44]. 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 [29]. 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 [45].

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 10 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 [41, 46]. 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 [35].

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 [35, 36, 47]. Furthermore, other factors such as food source, pollution (e.g. discarded plastics), parasitism and even the competition for snail habit with other organisms within an area were not considered in our model [47]. 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 [48].

One limitation of our study is the restricted study period (November 2017 to June 2019, except for the 5-year evapotranspiration time frame) as well as taking the mean values for each prediction location. Rabone et al. [41] 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 [47]. 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 Additional file 4: Fig. S1, Fig. S2 and Fig. S3; 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 [41, 49]. 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 [49]. However, it was excluded from the study because of missing River Shire values as shown in Additional file 10: Fig. S1, Fig S2 and Fig S3. 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 [1–3].

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.

Conclusions

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.

Abbreviations

BMLM	Bayesian multilevel model
CM	Cambisols
°C	Degree Celsius
GP	Gaussian process
GL	Gleysols
GLOBathy	GLObalBathymetric
IS	Intestinal schistosomiasis
ISRIC	International Soil Reference and Information Centre
ITCZ	Inter-Tropical Converge Zone
LPDAAC	Land Processes Distribution Active Archive Center
LST	Land surface temperature
LV	Luvisols
MCMC	Markov chain Monte Carlo
Millimetre	Mm
MODIS	Moderate-resolution imaging spectroradiometer
NDVI	Normalised Difference Vegetation Index
1D	One-dimensional
PL	Planosols
SAC	School-aged children
SOTER	Soil Terrain Database for Malawi
TAMSAT	Tropical Applications of Meteorology using SATellitedata and ground-based observations
2D	Two-dimensional
UGS	Urogenital schistosomiasis

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13071-024-06353-y.

Additional file 1. Dataset.

Additional file 2. Figure S1, Figure S2.

Additional file 3. Model formulation.

Additional file 4. Figure S1, Figure S2 and Figure S3.

Additional file 5. Figure S1 and Figure S2.

Additional file 6. Figure S1 and Figure S2.

Additional file 7. Figure S1.

Additional file 8. Figure S1.

Additional file 9. Figure S1.

Additional file 10. Figure S1, Figure S2 and Figure S3.

Author contributions
MHA, PM, CC, JH, JA, SJ, SAK and JM collected the primary data used in this secondary analysis. ALR developed the analysis, and wrote the paper with data analysis and interpretation of the results with guidance from MCS, JRS, CF and CJ. All authors read and approved the final manuscript.

Availability of data and materials
All code for this publication is accessible on Zenodo https://doi.org/10.5281/zenodo.10410622. The primary data are provided in Additional file 1: Dataset.

Declarations

Competing interests
The authors declare that they have no competing interests.

Ethics approval and consent to participate
Research approvals were granted in Malawi by the National Health Sciences Research Committee (1805), Mangochi District Health Office Research Committe (26.04.2019) and in the UK by LSTM Research Ethics Committee (30.04.2019). Written informed guardian consent was obtained for each school child before participation in surveys.

Consent for publication
Not applicable.

Author details
[1]Lancaster Medical School, Lancaster University, Bailrigg House, Bailrigg, Lancaster LA1 4YE, UK. [2]Tropical Disease Biology, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK. [3]Ministry of Health, 52367 Buraydah, 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, UK. [6]Mathematics and Statistics, Lancaster University, Bailrigg House, Bailrigg, Lancaster LA1 4YE, UK.

Reed *et al. Parasites & Vectors*   (2024) 17:272

Page 12 of 13

References

1. Alharbi MH, Condemine C, Christiansen R, LaCourse EJ, Makaula P, Stanton MC, et al. *Biomphalaria pfeifferi* snails and intestinal schistosomiasis, Lake Malawi, Africa, 2017–2018. Emerg Infect Dis. 2019;25:613–5. https://doi.org/10.3201/eid2503.181601.

2. Kayuni SA, O'Ferrall AM, Baxter H, Hesketh J, Mainga B, Lally D, 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. *Infect Dis Poverty*. 2020. https://doi.org/10.1186/s40249-020-00736-w.

3. Alharbi MH, Iravoga C, Kayuni SA, Cunningham L, LaCourse EJ, Makaula P, et al. First molecular identification of *Bulinus africanus* in Lake Malawi implicated in transmitting Schistosoma parasites. *Trop Med Infect Dis*. 2022;7(8), 195. https://doi.org/10.3390/TROPICALMED7080195.

4. Madsen H, Stauffer JR. Schistosomiasis control under changing ecological settings in lake Malawi. EcoHealth. 2022;19:320–3. https://doi.org/10.1007/S10393-022-01606-7/METRICS.

5. Adekiya TA, Aruleba RT, Oyinloye BE, Okosun KO, Kappo AP. The effect of climate change and the snail-schistosome cycle in transmission and bio-control of schistosomiasis in sub-saharan africa. *Int J Environ Res Public Health*. 2020;17(1), 181. https://doi.org/10.3390/ijerph17010181.

6. Walz Y, Wegmann M, Dech S, Raso G, Utzinger J. Risk profiling of schistosomiasis using remote sensing: approaches, challenges and outlook. *Parasites Vectors*. 2015. https://doi.org/10.1186/s13071-015-0732-6.

7. De Kock KN, Wolmarans CT, Bornman M. Distribution and habitats of *Biomphalaria pfeifferi*, snail intermediate host of *Schistosoma mansoni*, South Africa. Water SA. 2004;30:29–36. https://doi.org/10.4314/wsa.v30i1.5023.

8. Madsen H, Bloch P, Makaula P, Phiri H, Furu P, Stauffer JR. Schistosomiasis in Lake Malaŵi villages. EcoHealth. 2011;8:163–76. https://doi.org/10.1007/s10393-011-0687-9.

9. Ogutu-Ohwayo R, Balirwa JS. Management challenges of freshwater fisheries in Africa. Lakes Reserv Res Manag. 2006;11:215–26. https://doi.org/10.1111/J.1440-1770.2006.00312.X.

10. Stauffer JR, Madsen H. A one health approach to reducing schistosomiasis transmission in Lake Malawi. Prev Med Commun Heal. 2018;1:1–4. https://doi.org/10.15761/PMCH.1000115.

11. Mapoma HWT, Xie X. Basement and alluvial aquifers of Malawi: an overview of groundwater quality and policies. Afr J Environ Sci Technol. 2014;8:190–202. https://doi.org/10.4314/ajest.v8i3.

12. Mkandawire PP. Groundwater resources of Malawi, isarm.org, 2004. https://isarm.org/sites/default/files/resources/files/Managing-Shared-Aquifer-Resources-in-Africa.pdf#page=91. Accessed 8 Apr 2023.

13. Eccles DH. An outline of the physical limnology of Lake Malawi. Limnol Oceanogr. 1974. https://doi.org/10.4319/lo.1974.19.5.0730.

14. Bhave AG, Bulcock L, Dessai S, Conway D, Jewitt G, Dougill AJ, et al. Lake Malawi's threshold behaviour: a stakeholder-informed model to simulate sensitivity to climate change. J Hydrol. 2020;584:124671. https://doi.org/10.1016/J.JHYDROL.2020.124671.

15. Statistical Office N, Macro I. Malawi 2010 Demographic and Health Survey, NSO ICF, 2011. http://www.nsomalawi.mw/images/stories/data_on_line/demography/MDHS2010/MDHS2010-report.pdf. Accessed 18 Apr 2023.

16. Kayuni SA, Corstjens PLAM, LaCourse EJ, Bartlett KE, Fawcett J, Shaw A, 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. Parasitology. 2019;146:1785–95. https://doi.org/10.1017/S0031182019000969.

17. 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., Malawi's Updated Nationally determined contributions 105 UNFCC. 2023. https://unfccc.int/documents/268340. Accessed 19 Apr 2023.

18. WHO. WHO GUIDELINE on control and elimination of human schistosomiasis. 2022. https://www.who.int/publications/i/item/9789240041608. Accessed 1 Dec 2022.

19. WHO. Ending the neglect to attain the sustainable development goals: a road map for neglected tropical diseases 2021–2030. Geneva: Word Health Organisation; 2020. p. 196.

20. Allan F, Ame SM, Tian-Bi YNT, Hofkin BV, Webster BL, Diakité NR, et al. Snail-related contributions from the schistosomiasis consortium for operational research and evaluation program including xenomonitoring, focal mollusciciding, biological control, and modeling. Am J Trop Med Hyg. 2020;103:66–79. https://doi.org/10.4269/AJTMH.19-0831.

21. King CH, Binder S, Shen Y, Whalen CC, Campbell CH, Wiegand RE, et al. SCORE studies on the impact of drug treatment on morbidity due to *Schistosoma mansoni* and *Schistosoma haematobium* infection. Am J Trop Med Hyg. 2020;103:30–5. https://doi.org/10.4269/AJTMH.19-0830.

22. Maidment RI, Grimes D, Allan RP, Tarnavsky E, Marcstringer M, Hewison T, et al. The 30 year TAMSAT African rainfall climatology and time series (TARCAT) data set. J Geophys Res Atmos. 2014;119:10619–44. https://doi.org/10.1002/2014JD021927.

23. Maidment RI, Grimes D, Black E, Tarnavsky E, Young M, Greatrex H, et al. A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa. Sci Data. 2017;4:170063. https://doi.org/10.1038/SDATA.2017.63.

24. Tarnavsky E, Grimes D, Maidment R, Black E, Allan RP, Stringer M, et al. Extension of the TAMSAT satellite-based rainfall monitoring over Africa and from 1983 to present. J Appl Meteorol Climatol. 2014;53:2805–22. https://doi.org/10.1175/JAMC-D-14-0016.1.

25. Wan Z, Hook S, Hulley G. MODIS/Terra land surface temperature/emissivity 8-day L3 global 1km global SIN grid V061 , NASA EOSDIS Land Processess DAA. 2023. https://lpdaac.usgs.gov/products/mod11a2v061/#citation. Accessed 19 Apr 2023.

26. Running S, Mu Q, Zhao M, Morena A. MODIS/Terra net evapotranspiration gap-filled yearly L4 global 500m SIN grid V061 , NASA EOSDIS Land Processes DAAC. 2023. https://lpdaac.usgs.gov/products/mod16a3gfv061/. Accessed 19 Apr 2023.

27. Didan K. MODIS/TERRA Vegetation Indices 16-Day L3 Global 1km SIN Grid V061, NASA, E0SDIS Land Processes DAAC. 2023. https://lpdaac.usgs.gov/products/mod13a2v061/#citation. Accessed 19 Apr 2023.

28. Golon D. The Land Processes Distributed Active Archive Center (LP DAAC) | U.S. Geological Survey, Earth Resources Observation and Science (EROS) center. 2023. https://www.usgs.gov/publications/land-processes-distributed-active-archive-center-lp-daac. Accessed 19 Apr 2023.

29. Dijkshoorn J., Leenaars J, Huting J, Kempen B. ISRIC Report 2016/01 Soil and Terrain database of the Republic of Malawi | ISRIC, ISRIC - World Soil Information. 2023. https://www.isric.org/documents/document-type/isric-report-201601-soil-and-terrain-database-republic-malawi. Accessed 18 Apr 2023.

30. 6.13 Gaussian Process Covariance Functions—Stan Functions reference.2023. https://mc-stan.org/docs/2_29/functions-reference/gaussian-process-covariance-functions.html. Accessed 5 May 2023.

31. Gong Y, Tong Y, Jiang H, Xu N, Yin J, Wang J, et al. Three Gorges Dam: Differential determinants and spatial-temporal effects of the change of snail density. Res Sq. 2023. https://doi.org/10.21203/RS.3.RS-2867328/V1.

32. Zheng Y, Shi Y, Li S, Deng W, Zhou J, Hu B, et al. Environmental determinants for snail density in Dongting lake region: an ecological study incorporating spatial regression. Am J Trop Med Hyg. 2022;107:1178–84. https://doi.org/10.4269/AJTMH.22-0238.

33. Nwoko OE, Manyangadze T, Chimbari MJ. Spatial distribution, abundance, and infection rates of human schistosome-transmitting snails and related physicochemical parameters in KwaZulu-Natal (KZN) province, South Africa. Heliyon. 2023;9:e12463. https://doi.org/10.1016/J.HELIYON.2022.E12463.

34. Wood S. Package "mgcv" Mixed GAM computation vehicle with automatic smoothness estimation, version. 1.8–31, 2019. https://cran.r-project.org/web/packages/mgcv/mgcv.pdf. Accessed 22 Jun 2020.

35. Odero S, Ogonda L, Sang D, Munde E, Shilulil C, Chweya P. Distribution of *Biomphalaria* snails in associated vegetations and schistosome infection prevalence along the shores of Lake Victoria in Mbita, Kenya: a cross-sectional study. East African Heal Res J. 2019;3:172–7. https://doi.org/10.24248/EAHRJ-D-19-00013.

36. Woodhall DM, Wiegand RE, Wellman M, Matey E, Abudho B, Karanja DMS, et al. Use of geospatial modeling to predict *Schistosoma mansoni* prevalence in Nyanza Province. Kenya, PLoS One. 2013;8:e71635. https://doi.org/10.1371/JOURNAL.PONE.0071635.
37. Alharbi MH, Condemine C, Hesketh J, Kayuni SA, Arme TM, Archer J, et al. *Biomphalaria pfeifferi* (Gastropoda: Planorbidae) Lake Malawi and Upper Shire River, Mangochi District, Malawi: distribution, genetic Diversity and pre-Patent schistosome infections. Trop Med Infect Dis. 2023;8:126. https://doi.org/10.3390/TROPICALMED8020126.
38. Chasimpha S, McLean E, Chihana M, Kachiwanda L, Koole O, Tafatatha T, 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. BMC Public Health. 2015;15:1–9. https://doi.org/10.1186/S12889-015-2323-Z/TABLES/4.
39. Tyler MD, Richards DB, Reske-Nielsen C, Saghafi O, Morse EA, Carey R, et al. The epidemiology of drowning in low- and middle-income countries: a systematic review. BMC Public Health. 2017;17:1–7. https://doi.org/10.1186/S12889-017-4239-2/FIGURES/3.
40. Stensgaard AS, Vounatsou P, Sengupta ME, Utzinger J. Schistosomes, snails and climate change: current trends and future expectations. Acta Trop. 2019;190:257–68. https://doi.org/10.1016/j.actatropica.2018.09.013.
41. Rabone M, Wiethase JH, Allan F, Gouvras AN, Pennance T, Hamidou AA, 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. Parasites Vectors. 2019;12:1–20. https://doi.org/10.1186/s13071-019-3745-8.
42. Tiner RW. Ecology of wetlands: classification systems. In: Tiner RW, editor. Encyclopedia of Inland Waters. Cambridge: Academic Press; 2009. p. 516–25. https://doi.org/10.1016/B978-012370626-3.00057-0.
43. Britannica. Gleysol, Encyclopedia Britannica. 2011. https://www.britannica.com/science/Gleysol. Accessed 10 May 2023.
44. Kulinkina AV, Walz Y, Koch M, Biritwum NK, Utzinger J, Naumova EN. Improving spatial prediction of *Schistosoma haematobium* prevalence in southern Ghana through new remote sensors and local water access profiles. PLoS Negl Trop Dis. 2018. https://doi.org/10.1371/journal.pntd.0006517.
45. Poggio L, De Sousa LM, Batjes NH, Heuvelink GBM, Kempen B, Ribeiro E, et al. SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. Soil. 2021;7:217–40. https://doi.org/10.5194/soil-7-217-2021.
46. Ernould JC, Kaman AK, Labbo R, Couret D, Chippaux JP. Recent urban growth and urinary schistosomiasis in Niamey, Niger. Trop Med Int Heal. 2000;5:431–7. https://doi.org/10.1046/j.1365-3156.2000.00577.x.
47. Appleton CC, Madsen H. Human schistosomiasis in wetlands in Southern Africa. Wetl Ecol Manag. 2012;20:253–69. https://doi.org/10.1007/S11273-012-9266-2/FIGURES/4.
48. Opisa S, Odiere MR, Jura WGZO, Karanja DMS, Mwinzi PNM. Malacological survey and geographical distribution of vector snails for schistosomiasis within informal settlements of Kisumu City, western Kenya. Parasit Vectors. 2011;4:226. https://doi.org/10.1186/1756-3305-4-226.
49. Khazaei B, Read LK, Casali M, Sampson KM, Yates DN. GLOBathy, the global lakes bathymetry dataset. Sci Data. 2022;9:1–10. https://doi.org/10.1038/s41597-022-.

Publisher's Note