Space and time predictions of schistosomiasis snail host population dynamics across hydrologic regimes in Burkina Faso

Javier Perez-Saez, Theophile Mande, Andrea Rinaldo

Laboratory of Ecohydrology, Federal Polytechnic School of Lausanne, Lausanne, Switzerland

Abstract

The ecology of the aquatic snails that serve as obligatory intermediate hosts of human schistosomiasis is driven by climatic and hydrological factors which result in specific spatial patterns of occurrence and abundance. These patterns in turn affect, jointly with other determinants, the geography of the disease and the timing of transmission windows, with direct implications for the success of control and elimination programmes in the endemic countries. We address the spatial distribution of the intermediate hosts and their seasonal population dynamics within a predictive ecohydrological framework developed at the national scale for Burkina Faso, West Africa. The approach blends river network-wide information on hydrological ephemeral nature which conditions snail habitat suitability together with ensembles of time ecological models forced by remotely sensed estimates of temperature and precipitation. The models were validated against up to four years of monthly snail abundance data. Simulations of model ensembles accounting for the uncertainty in remotely sensed products adequately reproduce observed snail demographic fluctuations observed in the field across habitat types, and produce national scale predictions by accounting for spatial patterns of hydrological conditions in the country. Geospatial estimates of seasonal snail abundance underpin large-scale, spatially explicit predictions of schistosomiasis incidence. This work can therefore contribute to the development of disease control and elimination programmes.

Introduction

Spatio-temporal modelling of the ecology of intermediate hosts or vectors of water-associated infectious diseases has been applied to malaria and other mosquito-borne diseases, in particular in terms of temperature suitability for Anopheles spp. (Weiss et al., 2014, 2015), and on the effects of rainfall on seasonal variations of mosquito densities (Bombly et al., 2008; Bombly and Eltahir, 2009; Tompkins and Ermert, 2013; Rumisha et al., 2014; Diboulo et al., 2015). Much less work has focused on the aquatic snails that serve as intermediate hosts of Schistosoma spp. which cause human schistosomiasis for which more than 220 million people are in need of preventive chemotherapy, most of which in sub-Saharan Africa (Garba Djimay and Montresor, 2017). The availability of information on the spatio-temporal patterns of snail abundance could be leveraged with spatially-explicit models of disease transmission to provide insight into the timing of transmission windows and their epidemiological effects, with implications for the design and deployment of national surveillance-response strategies for disease elimination (Bergquist et al., 2015; Perez-Saez et al., 2015).

Spatial studies on the intermediate hosts of the two main Schistosoma species in sub-Saharan Africa, Schistosoma mansoni and Schistosoma haematobium causing the intestinal and urogenital forms of the disease, respectively, have principally focused on the determination of their ecological ranges using historical presence/absence in species distribution models (SDMs) (Stensgaard et al., 2013, 2016; Pedersen et al., 2014). Following similar lines but with recent malacological surveys at multiple times during the year, Manyangadze et al. (2016) took into account seasonal variation in snail occurrence in South Africa by partitioning SDMs by season (cold/dry, hot/dry, rainy, post-rainy), thus producing sea-
son-dependent suitability maps. The inclusion of seasonality in spatial estimations of snail occurrence was an innovative step; however, their approach did not account for variations in snail abundance – only presence/absence – and did not incorporate temporal auto-correlation in snail abundance from one season to the other, while temporal ecology modelling has focused only on local snail population dynamics in specific habitats (Woolhouse and Chandiwana, 1990; Perez-Saez et al., 2016; Gurarie et al., 2017).

Here, we propose a methodology aimed at bridging the gap in spatio-temporal predictions of snail population dynamics for use in schistosomiasis risk mapping and incorporation into spatially-explicit models of disease transmission. We extend previous work concerning three experimental sites across the climatic regions in Burkina Faso (Perez-Saez et al., 2016, Perez-Saez, Mande, Zongo, Rinaldo, Comparative analysis of time-based and quadrat sampling in seasonal population dynamics of intermediate hosts of schistosomiasis, unpublished data) by building monthly discrete-time models of snail demography forced by remotely sensed estimates of key environmental covariates. The models are calibrated to reproduce multi-season ecological data, and are properly regionised by using spatial estimates of hydrologic ephemerality.

Materials and Methods

Snail abundance data and ecological models

The field sites and protocols for the collection of field ecological and environmental data have been described elsewhere (Perez-Saez et al., 2016, Perez-Saez et al., unpublished data)

Briefly, snail sampling was carried out in three villages in Burkina Faso covering the country’s climatic regions from May 2014 to December 2017. We focused on two sites in which natural habitats were studied, including a temporary pond and an ephemeral stream in the Centre-East region of the country (Lioulgou), and a permanent stream in the South-West (Panamasso) (Figure 1). The Sudanian climate in Panamasso is characterised by annual rainfall of around 1100mm with a longer rainy season stretching from June to October, and the Sudan-Sahelian climate experienced by Lioulgou characterised by annual rainfall less than 900mm per year and a shorter rainy season from July to September (Sivakumar and Gnoumou, 1987). Two types of sampling schemes were performed, one time-based using a weekly frequency and one quadrat-based using a monthly frequency. In the perspective of providing spatially consistent predictions of snail abundance, the use of time-based sampling has the disadvantage of being subject to observer-dependent frailty and can only yield relative abundance counts which cannot be compared across sites (Hairston et al., 1958). Quadrat sampling can provide estimates of absolute snail densities, subject to the suitability of the type of habitat to this method. However, it is possible to establish the correspondence between the two methods which then enables the transformation of one type of counts into the other (Perez-Saez et al., unpublished data).

Following accepted methods (Perez-Saez et al., 2016), the Malthusian, Ricker and Gompertz discrete time models are here used to describe snail population dynamics, allowing for piecewise linear effects of the covariates on the natural rate of increase (Toms and Lesperance, 2003). The use of piecewise linear models was motivated by the need to account for sharp changes in population growth and collapse that characterise snail ecological dynamics in the absence of environmental data gathered at high temporal resolutions as it has been the case in other studies as in Perez-Saez et al. (2016), in which data on air temperature, humidity, water temperature, level and conductivity had been collected at a 5min temporal resolution throughout the study period. The regionalisation of the predictions to the rest of the country needs to rely on remotely sensed estimates of environmental variables (section 1 of the Appendix). Specifically, the change of the relative snail abundance \( N \), between time points \( t \) and \( t + 1 \), with a Ricker-type density feedback, was modelled as (Eq. 1):

\[
\log\left(\frac{N_t}{N_{t-1}}\right) = a + b \Gamma_t + \sum_{i=1}^{m} \sum_{j=1}^{n} c_i \cdot d_{ij} \cdot \left(I_{y \geq 0} \cdot f(y)\right) \cdot \left(\frac{X_{t-1}^j}{T_{x,j}} > T_{x,j}\right)
\]

Eq. 1

where: \( b \leq 0 \) is the negative density-feedback parameter, and \( c_i \cdot d_{ij} \) denotes the coefficient of environmental covariate feature \( X_{t-1}^j \) multiplied by the indication function \( I(\cdot) \) of covariate \( X_{t-1}^j \) being larger than the covariate- and lag-specific threshold \( T_{x,j} \), taken at non-negative integer time lags \( r_x, r_z \leq T \).

Given the monthly time step, only lags of 1 month were considered. Lagged covariate features were taken in all possible combinations up to a maximum of \( m + n = 5 \), yielding a total of \( 500,000 \) models to test for each habitat/species combination. The thresholds \( T_{x,j} \) were fit for each covariate combination using the Nelder-Mead algorithm (Nelder and Mead, 1965). Product interactions of the type \( \eta_i \cdot X_{t-1}^j \) were not considered. As proposed by Perez-Saez et al. (2016), the Gompertz-type density feedback, where \( N \) is substituted by \( \log(N) \) on the right hand side of Eq. 1, and the Malthusian model \( b=0 \) were also tested. The time-series of monthly mean time-based counts was used for model calibration. This choice was due to the fact that quadrat-based counts presented gaps due to logistical constraints during the summer of 2016 in Panamasso during which logistical constraints prevented the collection of quadrat data during two months. Time. Time-based counts were collected throughout the study period. The functional relationships between time- and quadrat-based counts identified by Perez-Saez et al., unpublished data, where then used to transform model outputs back to absolute densities with the latter as basis for the analysis of spatio-temporal patterns of snail abundance.

Remote sensing

Remotely sensed data were used for the regionalisation of predictions with respect to snail intermediate host population dynamics throughout the country. The choice of which covariates to include in the spatial models of snail ecology took previous identification results into consideration (Perez-Saez et al., 2016) as well as the availability of relevant remotely sensed products providing estimates at suitable spatial and temporal resolutions. Among the six environmental covariates measured in situ (air temperature, water temperature, conductivity, level, precipitation amounts and number of intense rainfall events), only precipitation
and air temperature were retained due to their systematic appearance in the model identification process and the availability of corresponding remotely sensed products. For both covariates, the latter were selected among possible candidates, corrected against ground measurements and their uncertainty characterised by fitting appropriate error models. According to common approaches in hydrology (Hong et al., 2006; Moradkhani and Sorooshian, 2009), such fitting is a necessary step as it accounts for the propagation of uncertainty in estimated environmental covariates into the outputs of snail demography models. A detailed account of the selection of remotely sensed products, contrasting against ground measurements, de-biasing (for precipitation), estimation (for temperature), and the quantification of associated errors are given in the Appendix. Briefly, for precipitation we retained Rainfall Estimate, v. 2.0 (RFEV2) of the Climate Prediction Centre’s decadal estimate at ≈10 km spatial resolution (Xie and Arkin, 1997) which we contrasted to monthly precipitation measurements from the Burkina Faso meteorological agency (Direction Générale de la Météo, Ouagadougou, Burkina Faso) for the period 2006-2016 and quantified uncertainty multiplicative error models according to McMillan et al. (2011) and Tian et al. (2013). For temperature, we estimated monthly mean maximum and minimum daily temperature with the MODIS Aqua/Terra temperature products (MYD11A2) 8-day composites at 1-km spatial resolution (Wan et al., 2015) using ground measurement from the meteorological agency of Burkina Faso following the multivariate regression methodology proposed by Garske et al. (2013) for temperature-dependent modelling of malaria transmission.

Model implementation and regionalisation

The ecological model fitting procedure was analogous to the one presented previously by Perez-Saez et al. (2016). The model fitting procedure consisted in computing the residual sum of squares for each combination of covariates and model type (considering density feedbacks or lack thereof). Model fitting for the ephemeral habitats in Lioulgou was done for the data up to December 2016 when sampling was interrupted due to logistical constraints, and up to December 2017 in the permanent habitat in Panamasso. The ecological models were calibrated using the mean estimated covariates from remote sensing. Ecological models were ranked using the compensated Akaike information criterion (AICc; Burnham and Anderson, 2002) and combined into an ensemble model using a weighted sum of model outcomes, where the weight \( w_i \) assigned to model \( i \) is a function of the difference between its AICc score and the smallest one obtained in the model set, \( \Delta \text{AIC}_i \), as expressed in the equation by Burnham and Anderson (2002) (Eq. 2):

\[
\text{Eq. 2: } w_i = \frac{e^{-\Delta \text{AIC}_i}}{\sum_i e^{-\Delta \text{AIC}_i}}
\]

Model ensemble predictions were realised by performing 1000 simulations of model outcomes using perturbed remotely sensed estimates of rainfall and temperature drawn from their respective error models, which depended on climatic region and month (section 2 of the Appendix). Model simulations were run for the entire available observation window for each habitat. Initial conditions for the simulation were drawn from a uniform distribution whose bounds correspond to the minimum and maximal sampled snail density in the first sampling month in each species/habitat dataset.

After comparison with observed ecological time series in the experimental sites, the simulation methodology was applied to the pixel-based prediction of snail abundance at the national scale. The resolution of the pixels in the simulations was of 10 km². All covariates were resampled to the simulation resolution using a bilinear interpolation. For the intermediate host of S. mansoni in Burkina Faso, Biomphalaria pfeifferi, the species probability of presence, based on hydrological ephemeral regions as previously delineated by Perez-Saez et al. (2017), were used as a mask for snail presence. The intermediate hosts of S. haematobium, from the genus Bulinus, are known to be present in both permanent and ephemeral rivers (Poda et al., 2004). Each type of habitat has their corresponding ecological model ensemble (i.e. the models fits to the ephemeral stream in Lioulgou, and the permanent stream in Panamasso). The two model ensembles were each applied to their respective hydrological ephemeral region as defined in Perez-Saez et al. (2017) (Figure 1). As done in the epidemiological analysis by Perez-Saez et al. (2017), we assumed that at least one Bulinus species can be present in all hydrological ephemeral conditions in Burkina Faso (Poda et al., 2004): Simulations for the permanent and ephemeral river habitats were extracted for each river stretch of the national river network. Accurate maps of the location of small temporary ponds that sustain Bulinus spp. are not available despite advances in remotely sensed approaches to small water bodies (Amitrano et al., 2017). Simulations for snail ecology in temporary ponds were therefore presented by pixel for the whole country.

Results

The simulated population dynamic models forced by remotely sensed covariates captured observed seasonal fluctuations of snail abundance well in all three habitats in the two sampling villages (Figure 2). Overall, 85%-100% of the 95% confidence intervals (CIs) of the mean time-based snail densities were within the 95% model simulation envelopes. Moreover, 79%-100% of the quadrat-based CIs of mean snail abundance where within the transformation of model simulations into absolute densities. We recall that the models were fitted against the mean monthly time-based relative abundance counts, and simulations were transformed to equivalent, absolute-density counts for comparison with quadrat-based snail density measurements (Perez-Saez et al., unpublished data). The simulations reproduced the seasonal dynamics of Bulinus spp. well in the temporary ponds and ephemeral streams, including their population bursts during the rainy season. The variability in the predicted onset of snail population is well captured by the temperature and precipitation error models, with some models greatly overshooting and some under-estimating the initial population growth in June after the first rains in the temporary ponds. Although the median simulation tends to under-estimate initial growth, the latter falls within the 95% simulation envelopes (Figure 2). The dry-season peak in the permanent river in Panamasso was also well followed by the median simulations. The marked wet-season dips in the abundance of Bulinus spp. were particularly well simulated with little sensitivity to uncertainty with regard to precipitation and temperature (average width of the 95% simulation envelope of 1.78 of the simulation median in the months of October and November). Models of Bi. pfeifferi are more sensible to covariate variations, in particular in the March-April population increase (average width of the 95%
Figure 1. Situation map of snail research in Burkina Faso. A) The capital of Burkina Faso (white dot) and two study sites (red dots) located in both permanent (all year-round river flow, blue lines) and ephemeral (more than six months of dry bed conditions, orange lines) hydrological conditions. B) Temporary pond sampled in Lioulgou. C) Ephemeral river sampled in Lioulgou. D) Permanent stream sampled in Panamasso. E) Time series of monthly snail abundance estimates with the time-based and quadrat protocols (error bars: 95% confidence interval).

Figure 2. Ecological predictions using remote-sensing covariates. Comparison between observed (points) and simulated (median: black line, shading: 95% simulation envelopes (SE) of 1000 simulations) monthly population dynamics using remotely sensed covariates in terms of whether 95% confidence interval (CI) of the monthly mean snail densities intersects 95% SE (purple points) or not (green points), with the fraction of data points that do so given for each snail species-habitat combination.
The simulation envelope of 2.77 of the simulation median, although the median simulation tracks the observed population fluctuations well. Overall the median simulations were in good agreement with the observed monthly abundance counts, in particular when considering simulation envelopes in addition to the median simulation.

Simulating habitat and species-specific ecological models using the remotely sensed covariates enabled the exploration of spatio-temporal patterns of snail abundance at the national scale (Figure 3). Overall *Bulinus* spp. presented the largest abundances in the permanent river network in the south-western part of the country (mean seasonal densities: 35 snails/m$^2$), followed by populations in temporary ponds (peak mean seasonal densities: $\approx 20$ snails/m$^2$), with simulated densities in the ephemeral streams being the lowest (peak mean seasonal densities: $\approx 10$ snails/m$^2$). *Bi. pfeifferi* presented similar simulated abundances than *Bulinus* spp. in permanent streams (peak mean seasonal densities: $\approx 30$ snails/m$^2$). Seasonal mean abundances highlighted the differences, both in timing and magnitude of snail abundance peaks among the habitat types. *Bulinus* spp. showed similar patterns with peak densities during the rainy season (July-September), with declining populations in the winter season (October-January) and close to zero in the dry season (February-June). Interestingly, the populations in the temporary ponds were not predicted to reach such low values during the dry season in very southern and south-western regions of the country, possibly due to the more abundant rainfall in the Sudanian climate in Panamasso with respect to the Sudano-
Sahelian climate experienced by Lioulgou. The simulated population dynamics of *Bulinus* spp. in the permanent stretches of the river network suggest that local environmental conditions yield different density patterns depending on location. The mean snail abundances in the winter period are high in all permanent river stretches, however the patterns differ for the dry and wet periods of the year. Strong densities were predicted in the rainy season in the West, with higher densities during the dry season in the most south-western parts of the country. Simulated spatio-temporal patterns of *Bi. pfeifferi* were more homogeneous across suitable habitats, with maximum densities during the dry season and low densities during the rainy and winter seasons.

The extraction of the timing of peak densities for each species and across habitats provided an additional perspective on the differences between snail population dynamics in the country (Figure 4). In the ephemeral aquatic habitats, the simulations indicated that the population peaks of *Bulinus* spp. in the temporary ponds are between August and September throughout the country. On the other hand, the peaks appeared earlier, around July, in the ephemeral streams in most regions, with some parts in the centre of the country and the most northern regions exhibiting a later peak at the same time as the temporary ponds (September). *Bulinus* spp. in permanent streams presented a much more heterogeneous timing with respect to peak density which was already visible from the seasonal means, with an overall East-to-West gradient of timing in the dry season extending into the early rainy season with peak densities occurring from January to June. In rare cases in the South-West, the timing of peak densities occurred in September. The timing of *Bi. pfeifferi* peak densities was homogeneous and occurred during the dry season (around March).

**Discussion and Conclusions**

Together, the outputs of the simulations highlight the species and habitat-specific heterogeneity in the population dynamics of the snail intermediate hosts at the national scale. The proposed methodology supports an objective evaluation of spatio-temporal fluctuations in the abundance of the schistosomiasis snail host across the highly seasonal climate of Burkina Faso. Model simulations forced by remotely sensed estimates of temperature and precipitation were, in fact, capable to reproduce observed snail popu-
lution dynamics across species, habitat types and hydrological conditions. The regionalisation of the results at the national scale enables the visualisation of the heterogeneity of the magnitude snail abundance seasonal means, as well as the timing of the peaks of the strongly seasonal population dynamics of these snails of medical importance.

Our methodology builds on a novel combination of ecological models fitted to in situ snail abundance time series, together with a national-scale categorisation of hydrological conditions to determine the spatial occurrence of distinct snail species as a function of their requirements in terms of aquatic habitat characteristics. It is important to note that the results presented in Figure 2 are full simulations in the sense that they are run for the entire simulation window fixing only the initial conditions, as opposed to 1-step ahead predictions (Perez-Saez et al., 2015).

Possible extensions of the modelling framework could consist, on the one hand of incorporating discharge estimates in the ecological models for rivers, and on the other of extending the pond models to account for snail ecology in man-made reservoirs and lakes. Regarding rivers, incorporating predicted monthly discharge into the ecological models, instead of rainfall alone, would allow the differentiation of snail population dynamics between tributaries and the main branches of river networks. National-scale predictions of discharge at the river-stretch level poses non-trivial challenges due to the issue of data scarcity and the time series heterogeneity in length, accuracy, and reliability (Perez-Saez et al., 2017). However, a hydrologic modelling approach based on the Budyko framework has shown good results for monthly discharge predictions (Zhang et al., 2008) providing an avenue for discharge prediction in data-scarce regions.

Regarding the extension of the ecological modelling framework to man-made reservoirs and lakes, the main factor which needs to be taken into account is the duration of surface water presence, which can support snail populations that typically extends beyond that of temporary ponds in small topographic sinks. This is well illustrated by the historical data in a small reservoir in the centre of Burkina Faso (Poda et al., 1996). Estimation of the retention capacity of these small reservoirs using remote sensing could provide the necessary information for generalising our approach to these types of habitats (Amiratino et al., 2016, 2017; Peckel et al., 2016). This water-availability extension of the modelling framework would bring it closer to the resource-based models (Gurarie et al., 2017), though allowing for density feedbacks and temperature effects.

The main limitation of the current simulation of snail population dynamics at the national scale is the application of the models fitted at only two locations (Lioulgou and Panamasso) to the rest of the territory with no contemporary validation points. However, this does not affect the methodological framework, but model-guided validation of the simulations presented in this work could involve the confirmation of the different timing of peak abundances of Bulinus spp. in permanent rivers in the south-western part of the country. Moreover, the abundance and seasonality of Bulinus spp. should be validated in the Sahelian climatic zone in the North of the country, and area for which we have no available data, and where transmission of urogenital schistosomiasis is the most intense (Perez-Saez et al., 2017). In addition, the realisation of validation sampling campaigns is hindered by important logistical difficulties and financial costs of long-term ecological monitoring efforts, such as the one that yielded the data for this analysis. In this perspective, the guidance provided by the model may greatly help in minimising the amount of resources needed to validate the modelling results we present in this work. Another limitation of the study is that interactions between snail genera where ignored in the ecological models, in particular in the permanent rivers, such as at the site of Panamasso where both Bulinus globosus and Bulinus pfeifferi are known to co-occur (Poda et al., 2004). To our knowledge there is no evidence for direct competition between the genera of the two species, despite anecdotal accounts of its role for the substitution of S. haematobium by S. mansoni in South-Western Burkina Faso (Colette et al., 1982). However, this has not been tested directly and this hypothesis could be the matter for future work using the available time series of abundances in Panamasso.

Leveraging remotely sensed data and hydrological insight we here provide a first example of national-scale predictions of the abundance fluctuations of the snail intermediate hosts of schistosomiasis in strongly seasonal climates. The spatio-temporal simulations of snail population dynamics presented in this work are represent a unique opportunity for the development and deployment of spatially explicit models of schistosomiasis transmission (Perez-Saez et al., 2015). As such, the ability to simulate snail demography at any specific location with reasonable accuracy paves the way for the implementation of national-scale model-based surveillance response systems to support schistosomiasis elimination activities in endemic countries.

References
System (HDSS) data, Burkina Faso. Parasite Vector 8:118.


