



A spatiotemporal recommendation engine for malaria control

QIAN GUAN*, BRIAN J. REICH, ERIC B. LABER

Department of Statistics, North Carolina State University, 2311 Stinson Dr. Raleigh, NC 27695-8203, USA

qguan2@ncsu.edu

SUMMARY

Malaria is an infectious disease affecting a large population across the world, and interventions need to be efficiently applied to reduce the burden of malaria. We develop a framework to help policy-makers decide how to allocate limited resources in realtime for malaria control. We formalize a policy for the resource allocation as a sequence of decisions, one per intervention decision, that map up-to-date disease related information to a resource allocation. An optimal policy must control the spread of the disease while being interpretable and viewed as equitable to stakeholders. We construct an interpretable class of resource allocation policies that can accommodate allocation of resources residing in a continuous domain and combine a hierarchical Bayesian spatiotemporal model for disease transmission with a policy-search algorithm to estimate an optimal policy for resource allocation within the pre-specified class. The estimated optimal policy under the proposed framework improves the cumulative long-term outcome compared with naive approaches in both simulation experiments and application to malaria interventions in the Democratic Republic of the Congo.

Keywords: Infectious disease; Malaria; Resource allocation; Sequential optimization; Satiotemporal model.

1. INTRODUCTION

Malaria is a vector-borne infectious disease affecting a large population worldwide, especially in tropical and subtropical regions. It is estimated that there were 216 million cases of malaria resulting in 445 000 deaths globally in 2016 (WHO, 2017), and thus it remains a major public health problem. Effective malaria interventions have been developed, including insecticide-treated mosquito nets (ITNs) (Lengeler, 1998), indoor residual spraying (IRS) (Pluess and others, 2010), and Artemisinin-based combination therapy (ACT) (Eastman and Fidock, 2009). However, these interventions are too costly to be given to everyone in need. The malaria research community has made great strides in mapping disease prevalence (Hay and others, 2009; Kang and others, 2018), modeling its transmission (Bhadra and others, 2011; Griffin and others, 2014, 2016), developing and testing effective interventions (Okell and others, 2014; Stuckey and others, 2014; Walker and others, 2016), and implementing these interventions in practice. We proposed to build on this work to develop a real-time recommendation engine (RE) for precision interventions to help policy-makers decide how to best allocate limited resources.

*To whom correspondence should be addressed.

Efforts have been made to recommend optimal interventions to different areas to reduce malaria incidence. *Okell and others* (2014) and *Stuckey and others* (2014) conducted cost-effectiveness analysis of different combinations of interventions for malaria control and evaluated how they vary depending on malaria transmission-related factors. However, they only evaluated the cost-effectiveness of several pre-designed intervention combinations. Also, they did not provide individualized intervention recommendations that are tailored to small areas or allow these recommendations change over time based on accumulative information. *Walker and others* (2016) used a dynamic mathematical model to capture the effect of interventions under different malaria transmission settings and recommended intervention choices at national, provincial, and pixel level. They only recommended what interventions should be implemented but did not specify the optimal amount of each intervention for each pixel. Also, they did not allow for recommendations to update over time.

Our proposed RE can be formalized as a sequence of decision rules, one per stage of intervention, that map current information to a resource allocation vector specifying the resources allocated to each unit (such as a health zone). The optimal RE should optimize the expectation of some long-term cumulative outcome, e.g., the average malaria prevalence over a given time horizon, and also account for the need for interpretability and fairness. It is related to the idea of dynamic treatment regimes (DTR), which are sequential decision rules of giving treatment recommendation to individual patient at each stage based on up-to-date patient information (*Murphy, 2003; Robins, 2004; Chakraborty and Moodie, 2013; Schulte and others, 2014*). However, there are unique challenges in the spatiotemporal resource allocation problem that distinguish it from typical DTR problems and prevent the direct application of existing DTR methods. First, methods for DTR typically assume that individuals to receive treatment are independent and a large number of independent trajectories of the individuals are available in the observed data. In our allocation problem, the districts or zones to receive resources are spatially dependent so that we have to treat them as a single allocation objective over the spatial domain at each time point. As a result, there is only single observation available at each time point without independent replications. Second, most of the existing methods for DTR can only solve problems with discrete treatment space including only a small number of treatment options. However, in our problem the allocation to each region is the proportion of residents to be given a bednet, and there are many regions that we need to determine the allocation, so the action space is continuous and high dimensional. *Chen and others* (2016), *Laber and Zhao* (2015), and *Rich and others* (2016) dealt with continuous-valued treatments, but the samples in their methods are assumed independent.

An optimal policy for spatiotemporal resource allocation (e.g., as found using dynamic programming) is a massively complex mathematical function of many geographic and epidemiologic inputs, requiring extensive computation due to the large number of spatial locations under consideration, and is essentially a black box that generates no new knowledge. In contrast, we propose an approach that estimates the optimal RE within an interpretable class of regimes and permits straightforward computation and is amenable to visualization, scrutiny, and stakeholder input. We build a spatiotemporal model for the progression and transmission of the disease and then draw posterior samples from the fitted system dynamic model to simulate future prevalence for any RE. The mean outcome for each regime within a pre-specified class of interpretable policies is estimated, and the optimal RE is maximizer over the class. The idea is related to the policy-search method for DTR which models the mean outcome as a function of each regime and chooses the maximizer as the optimal regime (*Robins and others, 2008; Orellana and others, 2010; Zhang and others, 2012; Zhao and others, 2012; Zhang and others, 2013; Zhao and others, 2015; Zhang and others, 2018*).

However, due to the challenges mentioned above, existing methods cannot be implemented directly to the continuous resource allocation problem. *Guan and others* (2020) used policy-search with g-computation to estimate optimal personalized recommendations for recall intervals for patients with periodontal disease. They built a nonparametric Bayesian model for disease progression process, specified

a class of policies on a clinically interpretable risk score, and estimated the optimal policy within this class. However, their observed data included a large number of patients with independent longitudinal profiles, and their decision space was discrete involving only two treatment options. [Laber and others \(2018\)](#) developed a spatiotemporal treatment allocation strategy to slow the spread of white nose syndrome in bats. They used a statistical spatial gravity model to forecast the spread of the disease and developed an algorithm to identify a subset of locations to receive treatment at each time point aiming to optimize long-run control of the disease. While they solved a similar problem, their settings are simpler than what we consider here. Their spatial gravity model is only used to predict the locations that might be affected by the disease at future time points, whereas we need to predict the disease prevalence at each location in the future. Also, they only decided whether or not to apply treatments at each location at each time point so that the decision for each location is binary, whereas we aim to decide the amount of resources to be assigned to each location so that the action space for each location is continuous. Because of these differences, we build a more comprehensive spatiotemporal model to predict disease prevalence and define a new class of policies that can accommodate a continuous action space while remaining interpretable.

Even though our proposed resource-allocation framework is sufficiently flexible to handle a continuous and high-dimensional decision space, the methodology is designed specifically for the setting where the data and interventions are defined annually at the district level. For applications with individual-level data or finer-scale interventions, the proposed methodology cannot be applied directly. In these cases, the spatiotemporal model and the class of policies would need to be redefined.

2. PROBLEM STATEMENT

Assume the spatial domain is divided into n health zones. For health zone $l \in \mathcal{L} = \{1, \dots, n\}$, the environmental covariates $(X_{l1}, \dots, X_{lp})^T$ are observed at baseline and remain constant over time. At each time point $t = 1, \dots, T$, in health zone l , the malaria disease rate, Z_{lt} , is estimated by the sample proportion of individuals with malaria in the malaria field survey conducted at time t . Following [Bhatt and others \(2015\)](#), we use $Y_{lt} = \text{logit}(Z_{lt})$ as the response. For settings with many health zone disease rates near zero, a binomial, Poisson, or negative binomial model might be preferable. The intervention resource allocated to health zone l at time point t is denoted $A_{lt} \in [0, 1]$. In our analysis, A_{lt} is the bed-net coverage, i.e., the proportion of individuals sleeping under a bednet. The data available after time t are $\mathbf{H}_t = \{\mathbf{X}, \mathbf{Y}_0, \mathbf{A}_1, \mathbf{Y}_1, \dots, \mathbf{A}_t, \mathbf{Y}_t\}$, where $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_p)$, $\mathbf{X}_k = (X_{1k}, \dots, X_{nk})^T$ for $k = 1, \dots, p$, $\mathbf{Y}_t = (Y_{1t}, \dots, Y_{nt})^T$ and $\mathbf{A}_t = (A_{1t}, \dots, A_{nt})^T$ represent the collection of covariates, disease rate, and resource allocation, respectively for all health zones.

Let $\mathbf{S}_t = \phi_t(\mathbf{H}_t) \in \mathbb{R}^{q_0}$ be a summary of information collected at time t , a resource allocation policy π be a function from $\mathcal{S} = \text{supp } \mathbf{S}_t$ to the action space $\mathcal{A} = [0, 1]^n$, and Π be the class of policies under consideration which we assume are parameterized by $\boldsymbol{\alpha} \in \mathbb{R}^K$. Hence, the policy is determined by the parameter vector $\boldsymbol{\alpha}$ which maps the currently available information to a resource allocation recommendation $\pi(\mathbf{S}_t; \boldsymbol{\alpha})$. Define $\Psi(\mathbf{s}_t) = \Psi(\phi_t(\mathbf{h}_t))$ to be the set of possible resource allocation actions with realized history information $\mathbf{h}_t$ so that $\pi(\mathbf{s}_t; \boldsymbol{\alpha}) \in \Psi(\mathbf{s}_t)$.

We formalize the optimal resource allocation policy using potential outcomes. Define $\mathbf{Z}_t^*(\bar{\mathbf{a}}_t)$ to be the potential disease rate at time t under the sequence of allocations $\bar{\mathbf{a}}_t = \{\mathbf{a}_1, \dots, \mathbf{a}_t\}$ up to time t . Define $\mathbf{Z}_t^*(\pi)$ to be the potential outcome at time t if the resources were allocated under the resource allocation policy π up to time t . The loss value associated with a policy is defined as the expectation of a loss function of the potential outcomes under the policy, such as the mean malaria prevalence over the next 5 years (assuming resources are allocated once per year), $L(\pi) = \mathbb{E}\left\{\left(\sum_{l=1}^n \sum_{t=T+1}^{T+5} Z_{lt}^*(\pi)\right)/5n\right\}$. The optimal policy within the pre-specified class is then defined as $\pi_{\text{opt}} = \arg \min_{\pi \in \Pi} L(\pi)$. In order to estimate the optimal policy using the observed data, we make the following assumptions ([Robins, 2004](#); [Schulte and others, 2014](#)):

(i) sequential ignorability, $\{\mathbf{Z}_k^*(\bar{\mathbf{a}}_k) : \text{for all } \bar{\mathbf{a}}_k \in \overline{\mathcal{A}}_k\}_{k \geq 1} \perp\!\!\!\perp \mathbf{A}_t | \mathbf{H}_t$ for $t = 1, \dots$, where $\overline{\mathcal{A}}_k = \mathcal{A}^k$ is the set of all possible resource allocation actions up to time k ; (ii) consistency, $\mathbf{Z}_t = \mathbf{Z}^*(\bar{\mathbf{A}}_t)$, where $\bar{\mathbf{A}}_t$ is the sequence of observed resource allocation actions up to visit t ; (iii) positivity, let $g(\mathbf{a}_t | \mathbf{h}_t)$ denote the conditional treatment density given realized history information $\mathbf{h}_t$, and then there exists $\epsilon > 0$ so that $g(\mathbf{a}_t | \mathbf{h}_t) \geq \epsilon$ for all $\mathbf{a}_t \in \Psi(\phi_t(\mathbf{h}_t))$ and $t = 1, 2, \dots$. With these assumptions, the loss value function to be optimized can be expressed using g-computation and the data-generating model, which is the malaria transmission model given in Section 3.

3. BAYESIAN SPATIOTEMPORAL MODEL

We extend the hierarchical Bayesian spatiotemporal model proposed in [Mugglin and others \(2002\)](#) to model the spread of the disease. We assume that

$$Y_{lt} = \eta_{lt} + v_{lt}, \quad (3.1)$$

where $v_{lt} \stackrel{iid}{\sim} \text{Normal}(0, \sigma_e^2)$ is measurement error and $\boldsymbol{\eta}_t = (\eta_{1t}, \dots, \eta_{nt})'$ is a spatiotemporal process. Let m_l be the number of neighbors of zone l and $\mathcal{I}_l$ be the index set of its neighbors. We assume at each time t and each health zone l that

$$\begin{aligned} \eta_{lt} - \eta_{lt-1} = & c_0 + b_0 A_{lt} + (c_1 + b_1 A_{lt}) \eta_{lt-1} + (c_2 + b_2 A_{lt}) \left(\frac{1}{m_l} \sum_{j \in \mathcal{I}_l} \eta_{jt-1} \right) \\ & + \sum_{k=1}^p \beta_{1k} X_{kl} + \sum_{k=1}^p \beta_{2k} X_{kl} A_{lt} + \epsilon_{lt}, \end{aligned} \quad (3.2)$$

so that the progression of the disease rate depends on its previous disease rate, its neighbors' previous disease rate, environmental covariates, and the intervention level. The coefficients $c_0, b_0, c_1, b_1, c_2, b_2$ can be interpreted as the intercept, the main effect of resource allocation, the main effect of the previous disease status, the interaction effect of resource allocation and the previous disease status, the main effect of previous neighborhood disease status, the interaction effect of resource allocation, and previous neighborhood disease status, respectively. The innovation process $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \dots, \epsilon_{nt})^T$ is modeled as a Gaussian Markov random field that is independent in time and spatially correlated. Specifically,

$$\boldsymbol{\epsilon}_t \sim MVN\{0, \sigma_s^2 (M - \rho G)^{-1}\}, \quad (3.3)$$

where M is diagonal with diagonal elements $m_1, \dots, m_n$; G is the adjacency matrix with $g_{ll} = 0, g_{lk} = 1$ if zone k is a neighbor of zone l , and $g_{lk} = 0$ otherwise; σ_s^2 is the variance parameter, and $\rho \in (0, 1)$ controls spatial dependence. An alternative to this model is the intrinsic model with $\rho = 1$ which would provide some computational savings at the expense of model flexibility.

Equivalently, the spatiotemporal process model can be written in matrix form as the Gaussian first-order autoregressive process:

$$\boldsymbol{\eta}_t = \mathbf{W}_t \boldsymbol{\eta}_{t-1} + c_0 \mathbf{1} + b_0 \mathbf{A}_t + \sum_{k=1}^p \beta_{1k} \mathbf{X}_k + \sum_{k=1}^p \beta_{2k} (\mathbf{X}_k \circ \mathbf{A}_t) + \boldsymbol{\epsilon}_t.$$

Spatiotemporal dependence is determined by the $n \times n$ propagator matrix $\mathbf{W}_t$ that depends on $\mathbf{A}_t$ and is parameterized by c_1 , b_1 , c_2 , and b_2 such that its (i, j) element is

$$w_{ij} = \begin{cases} (1 + c_1) + b_1 A_{it} & \text{if } j = i; \\ (c_2 + b_2 A_{it})/m_i & \text{if } j \in \mathcal{I}_i, \text{ i.e., zone } j \text{ is the neighbor of zone } i; \\ 0 & \text{otherwise.} \end{cases}$$

The term $c_0 + b_0 \mathbf{A}_t + \sum_{k=1}^p \beta_{1k} \mathbf{X}_k + \sum_{k=1}^p \beta_{2k} (\mathbf{X}_k \circ \mathbf{A}_t)$ varies depending on the resource allocation $\mathbf{A}_t$ and environmental covariates $\mathbf{X}$ to control the latent process η_t to switch among growth, recession, or stable phases. The latent process is able to incorporate the interaction between the spatial and temporal correlation through the propagator matrix $\mathbf{W}_t$. Different levels of interventions will affect the environmental effects on the disease rate and also how the disease spreads in space and time.

Here, we assume a first-order autoregressive model in time. For the motivating malaria analysis, this should be sufficient because the time from exposure to manifestation is much less than the yearly time step of the data. In other settings, higher-order autoregressive models can be fit to account for incubation periods or other delayed effects.

4. RECOMMENDATION ENGINE

We use policy search to estimate an optimal resource allocation policy within a class of policies that are interpretable to domain experts. We use a global utility function to parameterize the class of policies under consideration. The global utility function $U_G(\mathbf{a}_t, \mathbf{s}_t; \boldsymbol{\alpha})$ summarizes the current available information $\mathbf{s}_t$ via the parameter $\boldsymbol{\alpha}$ and is a function of the allocation $\mathbf{a}_t$. We consider the class of policies of the form $\pi(\mathbf{s}_t; \boldsymbol{\alpha}) = \arg \max_{\mathbf{a}_t} U_G(\mathbf{a}_t, \mathbf{s}_t; \boldsymbol{\alpha})$ subject to some cost constraints, where $\mathcal{U} = \{U_G(\mathbf{a}_t, \mathbf{s}_t; \boldsymbol{\alpha})\}_{\boldsymbol{\alpha} \in \mathbb{R}^K}$ is a class of global utility functions s.t. $U_G(\mathbf{a}_t, \mathbf{s}_t; \boldsymbol{\alpha})$ measures the “goodness” of allocation $\mathbf{a}_t$ in state $\mathbf{s}_t$ under parameter $\boldsymbol{\alpha}$. The class $\mathcal{U}$ is chosen to capture salient features of the decision problem, for example, cost, fairness, spread dynamics, and logistics. Thus, the resultant class of policies is interpretable and the estimated optimal policy, say indexed by $\hat{\boldsymbol{\alpha}}_{\text{opt}}$, can be plugged into this utility function to identify features driving optimal malaria control. In practice, the total available resources are limited by annual logistical or budgetary constraints. For example, constraints are placed on the bednet distribution in each region, $0 \leq a_{lt} \leq 1$, and on the total number of bednets distributed so that $\sum_{l=1}^n a_{lt} N_{lt} \leq \mathcal{C}(\sum_{l=1}^n N_{lt})$, where N_{lt} is the population in health zone l at time t .

To construct the global utility function for each health zone $l = 1, \dots, n$, we first define a priority score that represents its priority level of receiving resources so that the zone with higher priority score should be allocated more resources. The priority score is constructed based on all available data before resource allocation at time t and depends on the policy parameter $\boldsymbol{\alpha}$, that is, $p_{lt}(\mathbf{s}_t, \boldsymbol{\alpha})$. Then, we define a local utility $U(a_{lt}, p_{lt}(\mathbf{s}_t, \boldsymbol{\alpha}))$ as a function of a_{lt} that calculates the local utility for health zone l with priority $p_{lt}(\mathbf{s}_t, \boldsymbol{\alpha})$ can gain if a_{lt} resources are allocated to the zone. The utility function should be monotonically increasing with the resource allocation since additional resources will bring additional benefits. The global utility function is constructed using the total local utility aggregated over the whole spatial domain.

4.1. Priority score

The priority score for each zone l and time t is defined to depend on q user-specified risk factors $\{f_{1lt}, \dots, f_{qlt}\}$ constructed from currently available data $\mathbf{S}_t$. There is great flexibility in constructing the risk factors. They can include the standardized environmental covariates $f_{klt} = X_{jt}$, such as temperature and precipitation, or the zone’s current disease status $f_{klt} = Y_{lt}$, the gradient of disease status $Y_{lt} - Y_{lt-1}$, or even nonlinear

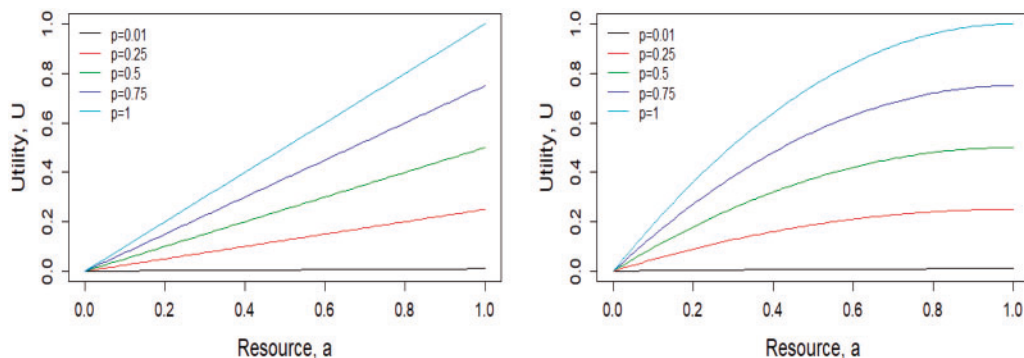


Fig. 1. Linear (left) and quadratic (right) utility function $U(a, p)$ as a function of resource allocation a for different values of the priority scores, p . The linear utility function is $U(a, p) = ap$ and the quadratic utility function is $U(a, p) = -p(a - 1)^2 + p$ for $0 \leq a \leq 1$.

summaries of the characteristics of the zone. The priority score is then defined as $p_{it}(\mathbf{s}_t, \boldsymbol{\alpha}) = 1 / \{1 + \exp[-(f_{1it}\alpha_1 + \dots + f_{qit}\alpha_q)]\} \in [0, 1]$, where $(\alpha_1, \dots, \alpha_q)$ determine the weights given to each risk factor in the priority score.

4.2. Local utility function

Intuitively, a reasonable local utility function $U(a, p)$ that depends on both the resource allocation and the priority score should satisfy the following assumptions: (i) $U(0, p) = 0$ for all p , that is, the utility is 0 if no resource is assigned; (ii) $U(a, p)$ is increasing in a for all p , that is, more resources will bring more benefits; (iii) $U(a, p)$ is increasing in p for $0 \leq a \leq 1$, that is, the same level of resources assigned to the health zone with higher priority score will bring more benefits. This property assures the health zone with higher priority score tend to receive more resources in order to maximize the overall utility.

A natural and simple utility function satisfying the above properties is the linear utility function $U_{lin}(a, p) = ap$. This utility function assumes that the individual utility increases linearly with resource allocation, that is, an additional unit of resource will bring constant marginal utility that is set to be the priority of the individual zone.

Equity or fairness is an important factor to consider when deciding the resource allocation in healthcare (Daniels and others, 2000; Gibson and others, 2004). Fairness concerns dictate that everyone should have the same right to the resources and the priority should be given to the zones with more severe situation (Nord and others, 1999; Nord, 2015). Similar to the idea proposed in Nord and others (1999), a convex utility function can accommodate some fairness concerns. One additional unit of resources will bring more utility if the zone has been assigned less resources. In other words, the marginal utility value should decrease with the resources allocation. So, the utility function that also satisfy the following additional property can account for fairness: (iv) $\frac{\partial^2}{\partial a^2} U(a, p) \leq 0$ for $0 \leq a \leq 1$ and all p , that is, the marginal utility value is decreasing when allocated resource is increasing. An example of the utility function that satisfies all above properties is the quadratic utility function $U_{quad}(a, p) = -p(a - 1)^2 + p$. Figure 1 illustrates the two utility functions with different values of priority scores. Both utility functions increase with a , but the gradient of the linear function in terms of a is constant while the gradient of the quadratic function is decreasing.

4.3. Global utility function

The global utility is defined as the summation of all local utilities but also with a penalty on the differences of resource allocations among neighborhood health zones:

$$U_G(\mathbf{a}_t, \mathbf{s}_t; \boldsymbol{\alpha}) = \sum_l U(a_{lt}, p_{lt}(\mathbf{s}_t, \boldsymbol{\alpha})) - \alpha_0 \sum_{i \sim j} (a_{it} - a_{jt})^2,$$

where $\alpha_0 \geq 0$, $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_q)$, and $i \sim j$ indicates that zone i and zone j are neighbors. The penalty term with a positive weight can smooth the resource allocation and account for fairness. A large penalty α_0 encourages spatial clusters of zone to be given intense treatments which could be more effective than scattered sites with intensive treatments. As the weight of the penalty term is to be optimized, this additional term allows for a more flexible class of policies to be considered.

The RE suggests allocations that maximize the global utility subject to the resource constraints on the total number of bednets distributed to the whole region:

$$\pi(\mathbf{s}_t; \boldsymbol{\alpha}) = \arg \max_{\mathbf{a}_t} U_G(\mathbf{a}_t, \mathbf{s}_t; \boldsymbol{\alpha}) \quad \text{such that} \quad \sum_l a_{lt} N_{lt} / (\sum_l N_{lt}) \leq C \text{ and } 0 \leq a_{lt} \leq 1. \quad (4.4)$$

Other nuances in the resource allocation decision making can be easily taken into consideration by adding more constraints to (4.4). For example, if it is agreed that allocating bednets to zones with malaria prevalence $< 1\%$ is ill-advised, we can add a constraint that $a_{lt} = 0$ if $z_{lt} < 0.01$.

5. POLICY SEARCH

The optimal priority score weights $\boldsymbol{\alpha}$ minimizes an optimality criterion $L(\boldsymbol{\alpha})$, such as the estimated expected cumulative (over space and time) malaria prevalence over the next 5 years. Therefore, although the policy only gives the resource allocation of one time point, the policy optimizes long-term outcomes. Given the posterior samples of the parameters in the Bayesian spatiotemporal model, the future malaria prevalence can be simulated to construct an estimator of the expected cumulative prevalence over space and time $\tilde{L}(\boldsymbol{\alpha})$. The plug-in estimator of the optimal weights is $\tilde{\boldsymbol{\alpha}}_{\text{opt}} = \arg \min_{\boldsymbol{\alpha}} \tilde{L}(\boldsymbol{\alpha})$. For numerical stability, we

replace $\tilde{L}(\boldsymbol{\alpha})$ with $\hat{L}(\boldsymbol{\alpha}) = \tilde{L}(\boldsymbol{\alpha}) + 0.0001 \sum_{i=0}^q \alpha_i^2$. Then $\hat{\boldsymbol{\alpha}}_{\text{opt}}$ is defined as the minimizer of $\hat{L}(\boldsymbol{\alpha})$.

Similar to the optimal allocation strategy estimation method in [Laber and others \(2018\)](#), we also draw samples from the posterior distribution over the postulated system dynamic model and use simulation to estimate the allocation strategy that optimizes long-term outcome. But different from the setting in [Laber and others \(2018\)](#) where the decision space is a binary vector, the action space in our setting is a high dimensional continuous space. We use a Kriging-based optimization method ([Picheny and others, 2013](#)), which achieves balance between exploration and exploitation, along with the simulation to estimate $\boldsymbol{\alpha}$ which minimizes $\hat{L}(\boldsymbol{\alpha})$. At the first step, we evaluate the value for the initial 100 points of $\boldsymbol{\alpha}$ from a Latin hypercube design generated using “optimumLHS” function in the R package “LHS.” We do a linear transformation of the initial design by setting the range of $\alpha_1, \dots, \alpha_q \in [-5, 5]$ and $\alpha_0 \in [0, 1]$. The value $\hat{L}_k$ corresponding to each point $\boldsymbol{\alpha}_k$ in the design is estimated using Monte Carlo simulation given the posterior samples of the parameters. By using a different posterior sample of the model parameters for each simulated trajectory in the Monte Carlo simulation, trajectories are samples from the full posterior predictive distribution of future prevalence and thus our policy accounts for both parametric and aleatoric uncertainty. Using the initial training data $\{\boldsymbol{\alpha}_k, \hat{L}_k\}$, a Gaussian process regression model is fit to predict the value corresponding to a new $\boldsymbol{\alpha}$ using kriging. The algorithm sequentially selects next weight $\boldsymbol{\alpha}$ to visit that optimizes the expected improvement (EI) defined in [Jones and others \(1998\)](#) and updates the

Gaussian process model parameters at each iteration. We use the the R package “DiceOptim” (Roustant and others, 2012) to implement the optimization procedure.

At each of the future years, if new data are available, the Bayesian spatiotemporal model can be refitted and the posterior distribution of the parameters can be updated. The resource allocation decision for the future years can be made by reoptimizing the priority score weights based on the updated posterior samples.

Each posterior sample of the parameters and predicted malaria prevalence in the Markov chain Monte Carlo (MCMC) procedure corresponds to an optimal α . We can quantify the uncertainty of $\hat{\alpha}_{\text{opt}}$ by applying the optimization procedure for each posterior draw to get the posterior distribution for α_{opt} .

6. SIMULATION

6.1. Generative model

For the simulation, we assume that there are $n = 100$ health zones that are arranged as a 10×10 square grid with grid spacing 1 between adjacent sites and include only one environmental covariate X_{l1} simulated from the Gaussian process with mean zero and variance one, and correlation $\text{Cor}(X_{l1}, X_{j1}) = \exp(-d_{lj}/2)$, where d_{lj} is the distance between the centroids of zone l and zone j . We simulate the baseline latent process for the health zones from a Gaussian process such that $\eta_0 = (\eta_{10}, \dots, \eta_{n0})^T \sim \text{MVN}(0, 0.5^2(M - 0.9G)^{-1})$ and the corresponding logit transformation of the disease rates are simulated from $Y_{l0} \sim N(\eta_{l0}, 0.01^2)$ for $l = 1, \dots, n$. We simulate the disease spread for $T = 5$ years following the generative model such that for $l = 1, \dots, 100, t = 1, \dots, 5, \eta_{lt} = (0.9 - 0.1A_{lt})\eta_{l,t-1} + (0.1 - 0.1A_{lt})/m_l \sum_{j \in N_l} \eta_{jt-1} + 0.2 - 0.7A_{lt} + 0.12X_{l1} - 0.1X_{l1}A_{lt} + \epsilon_{lt}$, where $\epsilon_t \sim \text{MVN}(0, 0.1^2(M - 0.9G)^{-1})$, and resource allocation is assumed increasing over time (to mimic the real malaria data) such that A_{lt} are simulated from $N(0.1 * t, 0.05^2)$ and are truncated at 0 and 1. The disease rates for the 5 years are simulated from $Z_{lt} \sim N(Y_{lt}, 0.01^2)$ for $l = 1, \dots, n$ and $t = 1, \dots, 5$. We simulate 100 data sets using this generative model.

6.2. Policy estimation

We consider three risk factors in the policy: the environmental covariate $f_{1lt} = X_{l1}$, logit of disease rate at previous time point $f_{2lt} = Y_{l,t-1}$, mean logit of disease rates of neighborhood zones at previous time point $f_{3lt} = \sum_{j \sim l} Y_{jt-1}/m_l$. The priority score of zone l at time t is then $1/\{1 + \exp[-(\alpha_1 X_{l1} + \alpha_2 Y_{l,t-1} + \alpha_3 \sum_{j \sim l} Y_{jt-1}/m_l)]\}$. We assume that the number of individuals in different health zones are the same and give the constraint that $\frac{1}{n} \sum_l A_{lt} \leq C$. We consider three scenarios with different values of resource constraint level $C = 0.2, 0.5$, and 0.8 .

We consider the following resource allocation policies for comparison:

1. Linear utility (Linear): our proposed policy with linear local utility function and a spatial penalty term to smooth the resource allocation.
2. Quadratic utility (Quad): our proposed policy with quadratic local utility function and a spatial penalty term to smooth the resource allocation.
3. Highest rate (Highest_rate): assign a bednet to each individual in the nC zones with highest disease rates and no bednets to the remaining zones.
4. Even: assign the same percentage of bednets C to each healthzone.

For each simulated data set, we use all information simulated up to year $T = 5$ as the training data to fit our proposed Bayesian spatiotemporal model using MCMC sampling with 5000 iterations. For the first two policies, the optimality criterion $L(\alpha)$ for each α is estimated using Monte Carlo simulations

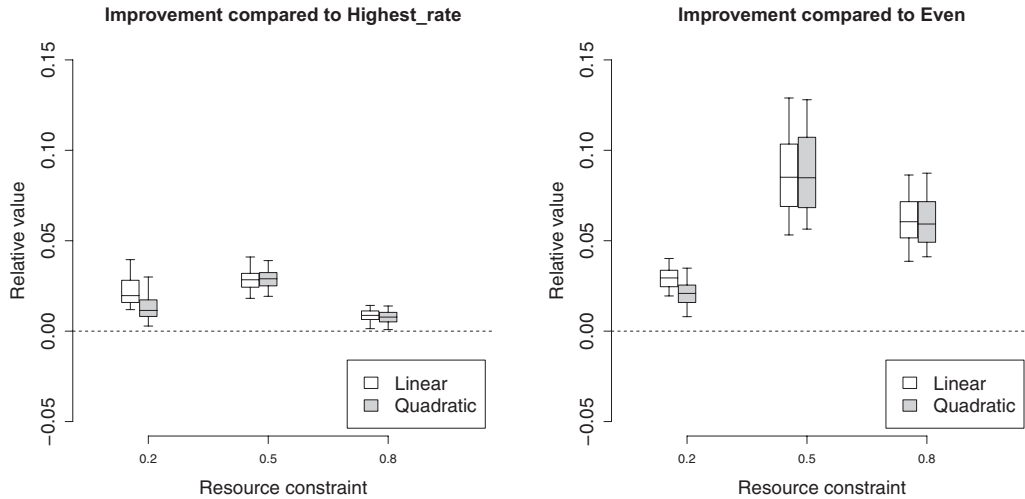


Fig. 2. The improvement of the proposed policies with linear utility function or quadratic utility function compared with “Highest_rate” policy (left) and “Even” policy (right) with different resource constraints $C = 0.2, 0.5$, and 0.8 when the model is correctly specified.

given the posterior samples. Sequential optimization is used to estimate the optimal policy that minimizes the expected mean malaria prevalence in the future 5 years within the pre-specified class. The loss value associated with estimated policy are approximated by 1000 Monte Carlo simulations given the true generative model. For the last two policies, there is no need to fit the model and no parameters to estimate. We just use 1000 Monte Carlo simulations given the true generative model to approximate the expected mean malaria prevalence in the future 5 years under the policy. The approximated loss values associated with the four policies for i th simulated data set are denoted as $L_{l}^i, L_{q}^i, L_{hr}^i, L_{ev}^i$, respectively.

We use the two naive policies “Highest_rate” and “Even” as two baseline policies and show the improvement of our proposed policies in terms of loss value compared with the baseline policy. For each of the simulated data sets, we compute the improvement as $(L_{hr}^i - L_{l}^i)/L_{hr}^i$, $(L_{ev}^i - L_{l}^i)/L_{ev}^i$ and $(L_{hr}^i - L_{q}^i)/L_{hr}^i$, $(L_{ev}^i - L_{q}^i)/L_{ev}^i$. Figure 2 plots the sampling distribution of the improvement of our proposed policies with different utility functions. Under this simulation setting, “Highest_rate” policy is preferred over “Even” policy. But we can see our proposed policies have significant improvement compared with either of the naive policies, especially when there are moderate level of total resources ($C = 0.5$). The policy with the linear utility function works slightly better than the policy with quadratic utility function when the total resource level is low ($C = 0.2$). This indicates more extreme resource allocation might improve the overall benefits under this specific simulation setting.

Supplemental materials available at *Biostatistics* online include an additional simulation study assuming the disease transmission model is misspecified to check the robustness of our method. From the simulation results, we can see under different simulation settings, our proposed policies are significantly better than naive policies and also consider fairness by allocating the resources more smoothly.

7. APPLICATION TO THE DRC DATA

We illustrate our method using data in the Democratic Republic of the Congo (DRC) primarily based on Demographic Health Surveys (DHS). DHS are cross-sectional, population-based cluster household surveys. In each survey, clusters are randomly chosen to be representative of the national population.

Table 1. *The posterior mean and 95% credible interval for parameters. The posterior mean with “★” represents the corresponding 95% credible intervals that excludes zero*

Response	Mean	95% CI		Mean	95% CI
Intercept (c_0)	-0.131★	(-0.192, -0.070)	σ_e^2	0.011★	(0.010, 0.012)
ITN (b_0)	-0.302★	(-0.366, -0.234)	σ_2^2	0.142★	(0.139, 0.146)
Temperature (β_{11})	0.033★	(0.020, 0.045)	ρ	0.999★	(0.998, 0.999)
Precipitation (β_{12})	0.003	(-0.012, 0.018)			
Temperature*ITN (β_{21})	-0.097★	(-0.125, -0.069)			
Precipitation*ITN (β_{22})	-0.053★	(-0.088, -0.019)			
Previous ($1 + c_1$)	0.917★	(0.902, 0.932)			
Previous*ITN (b_1)	0.096★	(0.062, 0.129)			
Prev_neighbor (c_2)	0.040★	(0.016, 0.065)			
Prev_neighbor*ITN (b_2)	-0.092★	(-0.145, -0.039)			

Within each cluster, households are randomly selected to participate in the survey. Two DHS program surveys—one in 2007 and another in 2014—were conducted to study malaria prevalence and treatment allocations in DRC. In the survey, structured questionnaires are administered to selected households to collect malaria-related information, such as their treatment status including bednet use. Also, dried blood spots were collected to test the malaria status.

Bhatt *and others* (2015) made use of the data from the DHS program surveys and five additional non-DHS program surveys and built a Bayesian hierarchical model to construct a malaria endemicity map across Africa from 2000 to 2015 in terms of *Plasmodium falciparum* parasite rate (PfPR). Interventions coverage levels from 2000 to 2015, including ITN, IRS, and ACT, are also estimated. We download the surface data of PfPR and ITN for DRC from <https://map.ox.ac.uk/country-profiles/#!/COD>. In the surface data, PfPR and ITN rate are estimated at a 5 km by 5 km resolution. Using the smoothed surfaces as data may bias our parameter estimates, but they greatly expand the spatiotemporal coverage of our data which is needed to build a disease progression model.

As malaria intervention resources are allocated in health zone level in DRC, we map the surface data of PfPR (Y_{it}) and ITN coverage rate (A_{it}) to each of the 515 health zones in DRC by taking the average values of the small cells lying in each health zone as the value of PfPR or ITN coverage rate in the corresponding health zone. We make the assumption that the populations at each 5 km by 5 km cell within a health zone are the same so that the mean rate of the cells can be used as the rate of the health zone. As there are almost no ITN usage before 2007, we only use the mapped data from 2007 to 2015 to train the Bayesian spatiotemporal model. In the model, we include two environmental covariates ($\mathbf{X}_1$ and $\mathbf{X}_2$): annual average temperature and annual average precipitation. As the annual average temperature and annual average precipitation are relatively stable in a certain area, we assume the annual average temperature and annual average precipitation are constant over time in each health zone. We download the monthly worldwide average temperature and precipitation data for 1970–2000 from <http://worldclim.org/version2> and map them to health zone level. We use the standardized mean annual average temperature and precipitation in 1970–2000 as the constant environment covariate values used in the model training and prediction.

The collected and processed data are fitted to the model in (3.1), (3.2), and (3.3). We use 5000 iterations in Gibbs sampling and discard first burn-in 2000 samples to obtain 3000 posterior samples. Table 1 summarizes the posterior mean and 95% credible interval of the parameters in the model. Temperature, previous disease status, and previous neighborhood disease status all have significant effects on the disease spread and the intervention ITN can significantly decrease the disease rate. Environmental covariates and previous disease status also have interaction effects with ITN on the disease progression.

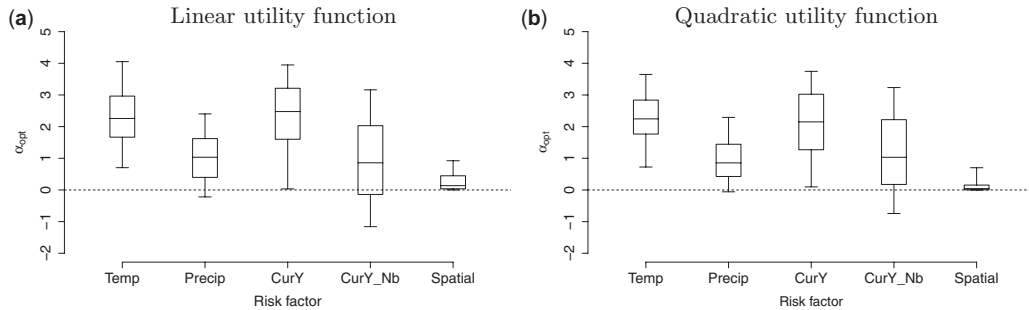


Fig. 3. Posterior distribution (5th, 25th, 50th, 75th, and 95th percentiles) of the weights α_{opt} for risk factors and the spatial penalty term when using linear utility function (left) or quadratic utility function (right). The risk factors include temperature (standardized), precipitation (standardized), current disease rate (logit), and current neighborhood disease rate (logit).

We consider four risk factors in the priority score in the policy: the standardized annual average temperature ($f_{1lt} = X_{1t}$) and the standardized annual average precipitation ($f_{2lt} = X_{2t}$), logit of disease rate at previous time point $f_{3lt} = Y_{lt-1}$, and mean logit of disease rate of neighborhood zones at previous time point $f_{4lt} = (1/m_l) \sum_{j \sim l} Y_{jlt-1}$. We evaluate the policies with either the linear local utility function or the quadratic local utility function and use the resource constraint level $C = 0.5$.

We randomly draw 100 posterior samples of the parameters and estimate the optimal policy in terms of α_{opt} corresponding to each posterior draw to get the posterior distribution of α_{opt} . Unlike typical MCMC sampling which suffers from autocorrelation, the samples of α_{opt} should be independent draws from the posterior and so fewer samples are needed for these parameters than are needed in typical MCMC sampling. Figure 3 plots the posterior distribution of α_{opt} when using either the linear utility function or the quadratic utility function. The posterior mean weights for all risk factors are positive, which suggests the priority of being allocated resources for each health zone tends to be positively correlated with temperature, precipitation, current disease status, and current neighborhood disease status. For both utility functions, temperature and current disease rate of the health zone seem to be most important factors in determining the risk score and priority. The posterior distribution of weight for the spatial penalty term concentrates more towards zero when using the quadratic utility function compared with using the linear utility function. This suggests that the spatial penalty term does help to better allocate the resources in terms of smoothness and efficiency when the linear utility function is used but does little to help when the quadratic utility function is used as the quadratic utility function is able to smooth the resource allocation inherently.

We also estimate one optimal policy averaging over the uncertainty of the parameters. The estimated optimal policy with the linear local utility function is with the priority score

$$P_{lt} = 1 / \{1 + \exp[-(2.1X_{1t} + 1.3X_{2t} + 3.1Y_{lt-1} + 0.77 \sum_{j \sim l} Y_{jlt-1}/m_l)]\}$$

and the weight for the spatial penalty $\alpha_0 = 0.06$. The estimated optimal policy with the quadratic local utility function is with the priority score

$$P_{lt} = 1 / \{1 + \exp[-(3.5X_{1t} + 1.1X_{2t} + 3.3Y_{lt-1} + 0.23 \sum_{j \sim l} Y_{jlt-1}/m_l)]\}$$

and the weight for the spatial penalty $\alpha_0 = 0.03$. The estimated optimal weights suggest that health zones with higher temperature, more precipitation, higher disease rate in the zone, and neighborhood zones tend

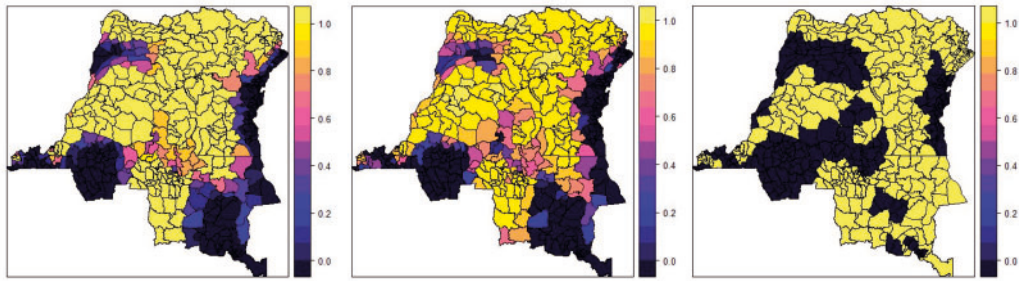


Fig. 4. The resource allocation next year using the estimated optimal policies with either the linear local utility function (left) or the quadratic local utility function (middle), or using “Highest_rate” policy (right).

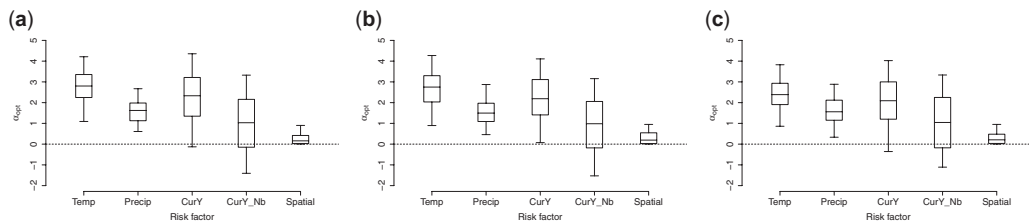


Fig. 5. Posterior distribution (5th, 25th, 50th, 75th, and 95th percentiles) of the weights α_{opt} for risk factors and the spatial penalty term when using linear utility function and data in year 2007–2012 (left), year 2007–2013 (middle), and year 2007–2014 (right) as training data, respectively. The risk factors include temperature (standardized), precipitation (standardized), current disease rate (logit), and current neighborhood disease rate (logit). (a) Training data: year 2007–2012. (b) Training data: year 2007–2013. (c) Training data: year 2007–2014.

to have higher priority to be allocated more resources. The loss value corresponding to the two optimal policies are 0.135 and 0.136, respectively and the loss values corresponding to “Highest_rate” policy and “Even” policy are 0.140 and 0.149, all with standard error 0.0005. The proposed policies improve the value by about 3% and 9% compared to the two naive policies. This is a substantial improvement when considering the number of disease cases that can be eliminated.

Figure 4 illustrates the resource allocation next year using the two estimated optimal policies or using “Highest_rate” rate policy. The estimated optimal policy using the quadratic utility function allocates the resources most smoothly while the “Highest_rate” policy only give extreme resource allocation (0 or 1).

We refit the model using data in years 2007–2012, 2007–2013, and 2007–2014 as training data and estimate the risk factor weights α_{opt} with linear utility function for resource allocation recommendation in year 2013, 2014, and 2015, respectively. The posterior distribution of optimal weights are given in Figure 5. They show that the updated estimated optimal policy is stable across years suggesting that a dynamic model would not lead to dramatic improvements for this analysis. Of course, for cases where treatment is applied annually there is ample time to update the policy, and so updating the policy using all available data would be advisable.

8. DISCUSSION

We develop a recommender system for spatiotemporal resource allocation to maximize the efficacy of malaria control efforts. Our proposed statistical framework deals with the challenges of spatial dependence and continuous action space. We used a hierarchical Bayesian spatiotemporal model to approximate

the system dynamics of the disease transmission involving the effect of environmental covariates and allocated resources, and construct a flexible and interpretable class of allocation policies that is also computationally feasible for searching the optimal resource allocation policy with the continuous action space. The simulation experiments suggest the proposed method performs well, and it is shown to be able to improve the resource allocation efficacy compared with naive policies in both simulation studies and the application to DRC data.

There are some limitations in our studies which provide possible directions for the future work. Our proposed RE relies on the postulated spatiotemporal model for the malaria transmission. A more flexible semi- or nonparametric model can be constructed to improve the robustness of the method to model misspecification. In the current framework, we only consider one intervention (ITN) at a time when optimizing the resource allocation. Our method can be extended to consider several interventions simultaneously and recommend the optimum allocation policy for all the related resources. The current framework assumes the resources are yearly allocated, and we do not consider seasonality in the model. If a monthly or even more frequent allocation is necessary, the seasonality can be incorporated as a covariate to the model, and the strategy can be defined for more frequent resource allocation with the constraint that an annual budget to be allocated across in space and time. As malaria intervention resources are allocated at health zone level in DRC, we build the spatiotemporal model based on the use of data aggregated at the district level and define RE to decide how to allocate resources to a finite number of regions. If applying treatment to points rather than a finite number of regions is more of interest and data are available at point locations, then a geostatistical model would be preferable to avoid bias in estimating covariate effects and a completely new framework is required to define the point level treatment allocation policy, and it is a topic for future work.

When we apply our method to the DRC data, the data we use to fit the model is the generated PfPR and ITN surface data estimated in [Bhatt and others \(2015\)](#) using the Bayesian hierarchical model based on very sparse survey data instead of the real yearly collected health zone level data. Also, we make the assumption that the populations at each 5 km by 5 km cell within a health zone are the same so that the mean rate of the cells can be used as the rate of the health zone. As a result, we only use the data as the illustration purpose instead of accurately reflecting the situation in DRC.

SOFTWARE

The R code that implements the proposed method for the simulation study and DRC data analysis is available at <https://github.com/qianguan/SpatiotemporalTreatmentAllocation>. All the data used for DRC data analysis is available at <https://figshare.com/account/home#/projects/99599>.

SUPPLEMENTARY MATERIAL

Supplementary material is available online at <http://biostatistics.oxfordjournals.org>.

ACKNOWLEDGMENTS

The authors thank the Bill and Melinda Gates Foundation (OPP1161913) and the National Institutes of Health (R01ES031651-01) for supporting this research and Dr. Mark Janko of Duke University for assistance with the DRC data.

Conflict of Interest: None declared.

FUNDING

Bill and Melinda Gates Foundation (OPP1161913) and the National Institutes of Health (R01ES031651-01).

REFERENCES

- BHADRA, A., IONIDES, E. L., LANERI, K., PASCUAL, M., BOUMA, M. AND DHIMAN, R. C. (2011). Malaria in Northwest India: data analysis via partially observed stochastic differential equation models driven by Lévy noise. *Journal of the American Statistical Association* **106**, 440–451.
- BHATT, S., WEISS, D. J., CAMERON, E., BISANZIO, D., MAPPIN, B., Dalrymple, U., BATTLE, K. E., MOYES, C. L., HENRY, A., ECKHOFF, P. A. and others. (2015). The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature* **526**, 207–211.
- CHAKRABORTY, B. AND MOODIE, E. E. (2013). *Statistical Methods for Dynamic Treatment Regimes*. New York, NY: Springer.
- CHEN, G., ZENG, D. AND KOSOROK, M. R. (2016). Personalized dose finding using outcome weighted learning. *Journal of the American Statistical Association* **111**, 1509–1521.
- DANIELS, N., BRYANT, J., CASTANO, R. A., DANTES, O. G., KHAN, K. S. AND PANNARUNOTHAI, S. (2000). Benchmarks of fairness for health care reform: a policy tool for developing countries. *Bulletin of the World Health Organization* **78**, 740–750.
- EASTMAN, R. T. AND FIDOCK, D. A. (2009). Artemisinin-based combination therapies: a vital tool in efforts to eliminate malaria. *Nature Reviews Microbiology* **7**, 864.
- GIBSON, J. L., MARTIN, D. K. AND SINGER, P. A. (2004). Setting priorities in health care organizations: criteria, processes, and parameters of success. *BMC Health Services Research* **4**, 25.
- GRIFFIN, J. T., BHATT, S., SINKA, M. E., GETHING, P. W., LYNCH, M., PATOULLARD, E., SHUTES, E., NEWMAN, R. D., ALONSO, P., CIBULSKIS, R. E. and others. (2016). Potential for reduction of burden and local elimination of malaria by reducing *Plasmodium falciparum* malaria transmission: a mathematical modelling study. *The Lancet Infectious Diseases* **16**, 465–472.
- GRIFFIN, J. T., FERGUSON, N. M. AND GHANI, A. C. (2014). Estimates of the changing age-burden of *Plasmodium falciparum* malaria disease in sub-Saharan Africa. *Nature Communications* **5**, 3136.
- GUAN, Q., REICH, B. J., LABER, E. B. AND BANDYOPADHYAY, D. (2020). Bayesian nonparametric policy search with application to periodontal recall intervals. *Journal of the American Statistical Association* **115**, 1066–1078.
- HAY, S. I., GUERRA, C. A., GETHING, P. W., PATIL, A. P., TATEM, A. J., NOOR, A. M., KABARIA, C. W., MANH, B. H., ELYAZAR, I. R. F., BROOKER, S. and others. (2009). A world malaria map: *Plasmodium falciparum* endemicity in 2007. *PLoS Medicine* **6**, e1000048.
- JONES, D. R., SCHONLAU, M. AND WELCH, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global optimization* **13**, 455–492.
- KANG, S. Y., BATTLE, K. E., GIBSON, H. S., RATSIMBASOA, A., RANDRIANARIVELOJOSIA, M., RAMBOARINA, S., ZIMMERMAN, P. A., WEISS, D. J., CAMERON, E., GETHING, P. W. and others. (2018). Spatio-temporal mapping of Madagascars Malaria Indicator Survey results to assess *Plasmodium falciparum* endemicity trends between 2011 and 2016. *BMC Medicine* **16**, 71.
- LABER, E. B. AND ZHAO, Y. (2015). Tree-based methods for optimal treatment allocation. *Biometrika* **102**, 501–514.
- LABER, E. B., MEYER, N. J., REICH, B. J., PACIFICI, K., COLLAZO, J. A. AND DRAKE, J. M. (2018). Optimal treatment allocations in space and time for on-line control of an emerging infectious disease. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **67**, 743–789.

- LENGELER, C. (1998). Insecticide treated bednets and curtains for malaria control. *Cochrane Database of Systematic Reviews*, Issue 3.
- MUGGLIN, A. S., CRESSIE, N. AND GEMMELL, I. (2002). Hierarchical statistical modelling of influenza epidemic dynamics in space and time. *Statistics in Medicine* **21**, 2703–2721.
- MURPHY, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **65**, 331–355.
- NORD, E. (2015). Cost-value analysis of health interventions: introduction and update on methods and preference data. *Pharmacoeconomics* **33**, 89–95.
- NORD, E., PINTO, J. L., RICHARDSON, J., MENZEL, P. AND UBEL, P. (1999). Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics* **8**, 25–39.
- OKELL, L. C., CAIRNS, M., GRIFFIN, J. T., FERGUSON, N. M., TARNING, J., JAGOE, G., HUGO, P., BAKER, M., DALESSANDRO, U., BOUSEMA, T. and others. (2014). Contrasting benefits of different artemisinin combination therapies as first-line malaria treatments using model-based cost-effectiveness analysis. *Nature Communications* **5**, 5606.
- ORELLANA, L., ROTNITZKY, A. AND ROBINS, J. M. (2010). Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part I: Main content. *The International Journal of Biostatistics* **6**.
- PICHENY, V., GINSBOURGER, D., RICHET, Y. AND CAPLIN, G. (2013). Quantile-based optimization of noisy computer experiments with tunable precision. *Technometrics* **55**, 2–13.
- PLUESS, B., TANSER, F. C., LENGELER, C. AND SHARP, B. L. (2010). Indoor residual spraying for preventing malaria. *Cochrane Database Systematic Review* **4**.
- RICH, B., MOODIE, E. E. M. AND STEPHENS, D. A. (2016). Optimal individualized dosing strategies: a pharmacologic approach to developing dynamic treatment regimens for continuous-valued treatments. *Biometrical Journal* **58**, 502–517.
- ROBINS, J., ORELLANA, L. AND ROTNITZKY, A. (2008). Estimation and extrapolation of optimal treatment and testing strategies. *Statistics in Medicine* **27**, 4678–4721.
- ROBINS, J. M. (2004). Optimal structural nested models for optimal sequential decisions. In: *Proceedings of the Second Seattle Symposium in Biostatistics*. New York, NY: Springer. pp. 189–326.
- ROUSTANT, O., GINSBOURGER, D. AND DEVILLE, Y. (2012). Dicekriging, Diceoptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *Journal of Statistical Software* **51**, 54p.
- SCHULTE, P. J., TSIATIS, A. A., LABER, E. B. AND DAVIDIAN, M. (2014). Q-and A-learning methods for estimating optimal dynamic treatment regimes. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics* **29**, 640.
- STUCKEY, E. M., STEVENSON, J., GALACTIONOVA, K., BAIDJOE, A. Y., BOUSEMA, T., ODONGO, W., KARIUKI, S., DRAKELEY, C., SMITH, T. A., COX, J. and others. (2014). Modeling the cost effectiveness of malaria control interventions in the highlands of western Kenya. *PLoS One* **9**, e107700.
- WALKER, P. G. T., GRIFFIN, J. T., FERGUSON, N. M. AND GHANI, A. C. (2016). Estimating the most efficient allocation of interventions to achieve reductions in *Plasmodium falciparum* malaria burden and transmission in Africa: a modelling study. *The Lancet Global Health* **4**, e474–e484.
- WHO. (2017). World malaria report 2017. *Technical Report*, World Health Organization.
- ZHANG, B., TSIATIS, A. A., LABER, E. B. AND DAVIDIAN, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics* **68**, 1010–1018.
- ZHANG, B., TSIATIS, A. A., LABER, E. B. AND DAVIDIAN, M. (2013). Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika* **100**, 681–694.

- ZHANG, Y., LABER, E. B., DAVIDIAN, M. AND TSIATIS, A. A. (2018). Interpretable dynamic treatment regimes. *Journal of the American Statistical Association* **113**, 1541–1549.
- ZHAO, Y., ZENG, D., RUSH, A. J. AND KOSOROK, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association* **107**, 1106–1118.
- ZHAO, Y.-Q., ZENG, D., LABER, E. B. AND KOSOROK, M. R. (2015). New statistical learning methods for estimating optimal dynamic treatment regimes. *Journal of the American Statistical Association* **110**, 583–598.

[Received April 5, 2020; revised March 9, 2021; accepted for publication March 14, 2021]