Published in final edited form as:

Geogr Anal. 2008 April ; 40(2): 167–188. doi:10.1111/j.1538-4632.2008.00718.x.

Developing geostatistical space-time models to predict outpatient treatment burdens from incomplete national data

Peter W. Gething^{*,1,2}, Abdisalan M. Noor³, Priscilla W. Gikandi³, Simon I. Hay^{3,4}, Mark S. Nixon¹, Robert W. Snow^{3,5}, and Peter M. Atkinson²

¹School of Electronics & Computer Science, University of Southampton, Highfield, Southampton, SO17 1BJ, UK

²School of Geography, University of Southampton, Highfield, Southampton, SO17 1BJ, UK

³Malaria Public Health & Epidemiology Group, Centre for Geographic Medicine, Kenya Medical Research Institute / Wellcome Trust Collaborative Programme, P.O. Box 43640, 00100 GPO, Nairobi, Kenya

⁴TALA Research Group, Tinbergen Building, Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, UK

⁵Centre for Tropical Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, UK

Abstract

Basic health system data such as the number of patients utilising different health facilities and the types of illness for which they are being treated are critical for managing service provision. These data requirements are generally addressed with some form of national Health Management Information System (HMIS) which coordinates the routine collection and compilation of data from national health facilities. HMIS in most developing countries are characterised by widespread under-reporting. Here we present a method to adjust incomplete data to allow prediction of national outpatient treatment burdens. We demonstrate this method with the example of outpatient treatments for malaria within the Kenyan HMIS. Three alternative modelling frameworks were developed and tested in which space-time geostatistical prediction algorithms were used to predict the monthly tally of treatments for presumed malaria cases (MC) at facilities where such records were missing. Models were compared by a cross-validation exercise and the model found to most accurately predict MC incorporated available data on the total number of patients visiting each facility each month. A space-time stochastic simulation framework to accompany this model was developed and tested in order to provide estimates of both local and regional prediction uncertainty. The level of accuracy provided by the predictive model, and the accompanying estimates of uncertainty around the predictions, demonstrate how this tool can mitigate the uncertainties caused by missing data, substantially enhancing the utility of existing HMIS data to health-service decision-makers.

INTRODUCTION

In an era in which the United Nations Millennium Development Goals (UN 2000) have focused efforts to strengthen health systems in resource-poor nations worldwide, awareness has grown of the concurrent need to improve the availability of reliable health data in these

* Author for correspondence: P.W.Gething, School of Geography, University of Southampton, Southampton SO17 1BJ, UK. pgething@soton.ac.uk.

settings (WHO/AFRO 1999; Murray et al. 2004; AbouZahr and Boerma 2005; Stansfield 2005). Effective planning and delivery of resources within a health system requires accurate and up-to-date information on the number of patients utilising different facilities and the types of illness for which they are treated. In most countries, such information requirements are addressed by some form of national Health Management Information System (HMIS) that coordinates the collection of treatment records from health facilities and the compilation of these records into a national database. A comprehensive HMIS database requires prompt monthly reporting of treatment records from all health facilities but, in many resource-poor settings, large proportions of health facilities never report or report infrequently (Al Laham et al. 2001; MoH Kenya 2001; WHO/SEARO 2002; Rudan et al. 2005). This sporadic reporting has led to spatially and temporally incomplete national data, presenting a challenge to evidence-based public health decision-making. In Africa, the widespread inadequacy of national HMIS has shown little improvement despite several decades of donor investment (Evans and Stansfield 2003; AbouZahr and Boerma 2005).

Faced with incomplete data coverage, important public health metrics are often estimated using rudimentary approaches to compensate for missing values. The purpose of this study was to define a generic modelling framework to allow prediction of the number of outpatient treatments for a given diagnoses at facilities where monthly treatment records are missing from a national HMIS database. We use the example of presumed malaria at government facilities in Kenya, where increased donor assistance and the introduction of expensive new drugs mean the government is facing an urgent need to quantify the national treatment burden for presumed malaria in its formal health sector (Kindermans 2002; WHO 2005; MoH Kenya 2005). Government health facilities in Kenya have been recently georeferenced, allowing the problem of predicting missing values to be placed in a space-time context. In this paper, three different space-time geostatistical modelling frameworks were developed, and their predictive accuracies compared. A stochastic simulation approach was adapted to provide a model for the corresponding prediction uncertainty.

THEORY

Space-time geostatistics

Geostatistics was originally developed to address problems of spatial prediction in the mining industry (Matheron 1971; Journel and Huijbregts 1978) and has been extended more recently to space-time settings (Kyriakidis and Journel 1999). In the space-time geostatistical paradigm, the value of an attribute of interest z at any location in space $\mathbf{u}_0$ and instant in time t_0 is modelled as a realisation of a spatiotemporal random variable (RV), $Z(\mathbf{u}_0, t_0)$. The set of all RVs at all possible spatial and temporal locations in the study domain represents a spatiotemporal random function (RF), $Z(\mathbf{u}, t)$. If $z(\alpha)$ is a set of observed space-time data of the attribute z at n space-time locations $((\mathbf{u}, t)_\alpha, \alpha = 1, 2, \dots, n)$, then a common problem is to predict values of $z^*(\beta)$ at a set of q unsampled space-time locations $((\mathbf{u}, t)_\beta, \beta = 1, 2, \dots, q)$, where the asterisk denotes a prediction (note that the symbols α and β are used as short-hand for $(\mathbf{u}, t)_\alpha$ and $(\mathbf{u}, t)_\beta$, respectively, from this point onwards to simplify the notation). Space-time kriging is an extension of the spatial-only kriging predictor that exploits spatiotemporal autocorrelation between dispersed values $z(\alpha)$ to make these predictions at unobserved points. Along with the data, space-time kriging algorithms require estimates of the covariance between RFs separated by different space-time lags $(\mathbf{h}_s, h_t)$ where $\mathbf{h}_s$ is a spatial vector of distance and direction, and h_t is a scalar separation in time. These estimates can be provided by estimating the covariance directly or, more commonly, the semivariance, γ , between all $i = 1, 2, \dots, p$, data pairs at a series of regular lag intervals, taking the average at each lag interval, and fitting a continuous model to these averages either manually, or automatically, according to a pre-defined criteria of

best-fit. The 2-d space-time semivariogram $\hat{\gamma}_{st}(\mathbf{h}_s, h_t)$ can be estimated as half the mean squared difference between data separated by a given space-time lag:

$$\hat{\gamma}_{st}(\mathbf{h}_s, h_t) = \frac{1}{2p(\mathbf{h}_s, h_t)} \sum_{i=1}^{p(\mathbf{h}_s, h_t)} [z((\mathbf{u}, t)_i) - z((\mathbf{u}, t)_i + (\mathbf{h}_s, h_t))]^2 \quad (1)$$

A continuous 2-d space-time semivariogram model, $\tilde{\gamma}_{st}(\mathbf{h}_s, h_t)$, can then be fitted to this surface allowing semivariance values to be estimated at any lag for input into the kriging process.

As in the spatial-only case, the principal concerns when fitting a continuous model to the sample space-time semivariogram are to ensure that the model chosen is valid (i.e. that conditionally negative semi-definiteness is ensured) and sufficiently flexible to allow adequate fitting to the data without requiring estimation of a large set of parameters. A diverse range of models have been proposed for modelling space-time autocorrelation structures (Rodriguez-Iturbe and Mejia 1974; Dimitrakopoulos and Luo 1994; Kyriakidis and Journel 1999; Stein, 2005). The most conceptually simple approach is to assume separability between the spatial and temporal components such that the space-time covariance can be factored into purely spatial and purely temporal covariances (Cressie and Huang 1999). Although separable models are convenient to construct and computationally efficient, the assumption of separability is often difficult to justify. This has led to the development of nonseparable models that allow a wider range of space-time autocorrelation structures to be represented (Cressie and Huang 1999; De Iaco et al., 2002; Gneiting 2002; Kolovos et al., 2004; Gneiting et al., 2005; Ma 2005). The product-sum model proposed by De Cesare et al. (2001, 2002) was adopted in this study because (i) it does not require the imposition of an arbitrary space-time metric, (ii) it offers a large class of flexible models that impose less constraints of symmetry between the spatial and temporal correlation components than simpler, separable, classes, and (iii) although generally nonseparable (De Iaco et al., 2002), the model can be fitted to data using relatively straightforward techniques similar to those already established for spatial-only and temporal-only semivariograms.

The product-sum space-time semivariogram model, $\tilde{\gamma}_{st}(\mathbf{h}_s, h_t)$, is defined in terms of the separate spatial and temporal semivariograms, $\tilde{\gamma}_s$ and $\tilde{\gamma}_t$, and the corresponding spatial and temporal sills, $C_s(0)$ and $C_t(0)$, where C_s and C_t are the spatial and temporal covariance, respectively, and the sill is defined as the limit value of each semivariogram, $\gamma(\infty)$:

$$\tilde{\gamma}_{st}(\mathbf{h}_s, h_t) = (k_1 C_s(0) + k_3) \gamma_t(h_t) + (k_1 C_t(0) + k_2) \gamma_s(\mathbf{h}_s) - k_1 \gamma_s(\mathbf{h}_s) \gamma_t(h_t) \quad (2)$$

The parameters k_1 , k_2 , and k_3 are defined as:

$$k_1 = [C_s(0) + C_t(0) - C_{st}(0, 0)] / C_s(0) C_t(0) \quad (3)$$

$$k_2 = [C_{st}(0, 0) - C_t(0)] / C_s(0) \quad (4)$$

$$k_3 = [C_{st}(0, 0) - C_s(0)] / C_t(0) \quad (5)$$

where $C_{st}(0,0)$ is the 'sill' of the space-time semivariogram. Various constraints are placed on these parameters to ensure model validity (see De Cesare et al. 2001). A key advantage of the product-sum model is that $\tilde{\gamma}_{st}(\mathbf{h}_s, h_t)$ is defined entirely by parameters of the separate spatial and temporal semivariograms and the space-time sill which can all be estimated from the sample space-time semivariogram surface (Equation 1).

The most commonly used spatial-only kriging predictor is ordinary kriging (OK). The space-time equivalent (space-time ordinary kriging, ST-OK) predicts unknown values $z^*(\beta)$ at each $\beta = 1, 2, \dots, q$ space-time locations as a linear combination of $\alpha = 1, 2, \dots, n$ data local in space and time to the prediction location (β):

$$z^*(\beta) = \sum_{\alpha=1}^n \lambda_{\alpha} z(\alpha) \quad \text{with} \quad \sum_{\alpha=1}^n \lambda_{\alpha} = 1 \quad (6)$$

For each prediction $z^*(\beta)$ ST-OK determines the weight λ_{α} assigned to each neighbouring datum such as to minimise the variance of the prediction error, $\sigma_{STOK}^2(\beta) = \text{var}[z^*(\beta) - z(\beta)]$ whilst maintaining unbiasedness of the predictor $z^*(\beta)$. In estimating optimum weights, kriging takes into account both the covariances between each datum and the point to be estimated, and the covariances between the data themselves.

Space-time stochastic simulation to estimate local and regional prediction uncertainty

Consider a set of predictions made at q unsampled space-time locations, $\{z^*(\beta), \beta = 1, 2, \dots, q\}$, over a space-time study region. In some cases, a global or regional summary of values may be of interest, such as the mean $\mu[z^*(\beta)]$ of the q predicted values over the entire region $\{\beta = 1, 2, \dots, q\}$, or of a subset of v points within a space-time sub-region, $\{\beta = 1, 2, \dots, v\}$. In addition to making kriged predictions of these global or regional means, it is necessary to provide estimates of the uncertainty associated with these predictions. Although kriging systems provide ‘optimum’ local predictions by minimising the variance of the error of each prediction, a set of kriging predictions appears ‘smoother’ than the original data due to a missing error component. Conceptually, the RF $Z(\mathbf{u}, t)$ can be decomposed into the predictor $z^*(\beta)$, as provided by kriging, and the corresponding unknown prediction error $R(\beta)$: $Z(\beta) = z^*(\beta) + R(\beta)$. Estimates of the uncertainty associated with predictions of regional or global means must take into account the variance introduced by this unknown error component in order to restore the full variance of the RF model.

One approach to the above problem is to simulate, for each of the $\beta = 1, 2, \dots, q$ prediction points, l realisations $\varepsilon^{(l)}(\beta)$ of the error component with zero mean and the correct variance and covariance which can then be added to the original prediction, $z^*(\beta)$, to give a conditional simulated prediction (Deutsch and Journel 1998, p. 127):

$$z^{(l)}(\beta) = z^*(\beta) + \varepsilon^{(l)}(\beta) \quad (7)$$

If $z^{(l)}(\beta)$ is to have the same variance as the true value $z(\beta)$, this approach requires that the error component is orthogonal to the predictor and has at least the same covariance, if not spatial distribution, as the actual error. A procedure to generate realisations of the error component under these conditions was proposed originally by Journel and Huijbregts (1978, p. 495) for a spatial-only setting and is presented here in a space-time context. $l = 1, 2, \dots, L$ nonconditional realisations $z_{nc}^{(l)}(\nu)$ that share the same covariance as the RF $Z(\mathbf{u}, t)$ are simulated at all data and prediction locations $\nu = 1, 2, \dots, n+q$. The original space-time kriging exercise performed on the data is then repeated using the simulated values at the n data locations $\{z_{nc}^{(l)}(\alpha), \alpha = 1, 2, \dots, n\}$ to obtain simulated predictions at the q unsampled prediction locations $\{z^{*(l)}(\beta), \beta = 1, 2, \dots, q\}$ to compare to the simulated values at these locations $\{z_{nc}^{(l)}(\beta), \beta = 1, 2, \dots, q\}$. Simulated errors are then defined for each prediction location as the difference between simulated values and simulated predictions, $\varepsilon^{(l)}(\beta) = z^{*(l)}(\beta) - z_{nc}^{(l)}(\beta)$, and these can be added to the original predictions $z^*(\beta)$ to give conditional simulated predictions, $z^{(l)}(\beta)$:

$$z^{(l)}(\beta) = z^*(\beta) + [z^{*(l)}(\beta) - z_{nc}^{(l)}(\beta)] \quad (8)$$

The distribution of a set of L realisations $\{z^{(1)}(\beta), z^{(2)}(\beta), \dots, z^{(L)}(\beta)\}$ at each prediction location represents the uncertainty of that prediction which can be summarised by the standard deviation of the L realisations, $\sigma_{\text{sim}}[z^*(\beta)] = \sigma[z^{(l)}(\beta)]$, $l = 1, 2, \dots, L$. Where the value of interest is the mean, $\mu[z^*(\beta)]$, of a set of $\beta = 1, 2, \dots, q$ predicted values within a space-time region, simulated realisations of the mean, $\mu[z^{(l)}(\beta)]$, can also be defined:

$$\mu[z^{(l)}(\beta)] = \frac{1}{q} \sum_{\beta=1}^q z^{(l)}(\beta) \quad (9)$$

and the distribution of the set of these L realisations $\{\mu[z^{(1)}(\beta)], \mu[z^{(2)}(\beta)], \dots, \mu[z^{(L)}(\beta)]\}$ represents the uncertainty of the predicted mean, $\mu[z^*(\beta)]$. Again, this uncertainty can be summarised by the standard deviation of the L realisations, $\sigma_{\text{sim}}[\mu[z^*(\beta)]] = \sigma[\mu[z^{(l)}(\beta)]]$, $l = 1, 2, \dots, L$.

The remaining issue is the choice of simulation algorithm to generate the L nonconditional simulated realisations of the RF $Z(\mathbf{u}, t)$. Sequential Gaussian simulation (sGs) is one such algorithm that creates realisations under the assumption of a multiGaussian RF model and is presented for spatial-only settings in Goovaerts (1997, pp. 380-393).

DATA

The data used in this study were obtained from the Kenyan Ministry of Health (Department of Health Management Information Systems). Data consisted of monthly records from 1765[†] government outpatient health facilities over an 84 month period (January 1996 - December 2002). Each record included the total number of treatment events made at each facility each month (termed total cases, TC) and the number of treatment events resulting from a diagnosis of malaria (termed malaria cases, MC). The records were not structured by age, sex or distinguished as initial or follow-up visits, and diagnoses were generally not slide-confirmed. MC therefore represented the count of presumed malaria cases seen as outpatients each month. A complete set of 84 monthly records from each of the 1765 facilities would consist of data for 148,260 facility-months. The dataset contained data for 63,542 facility-months (43%), meaning 84,718 (57%) were unsampled.

Data were linked to a georeferenced database that contained the longitude and latitude coordinates of each facility (Noor et al. 2004; Noor 2005). Government health facilities included in this study were categorized as dispensaries (1194 facilities, 40,191 records), health centres (445 facilities, 18,669 records), or hospitals (126 facilities, 4,682 records) according to the level of services offered, in line with Noor et al. (2004).

METHODOLOGY

Defining three alternative modelling frameworks to predict malaria cases

The aim of this study was to define and test a modelling framework to predict the variable MC, as defined above, at unsampled facility-months within the national HMIS database. Data in the form of counts of disease cases are rarely used for epidemiological inference or prediction in their raw format but are generally standardised by some measure of the

[†]This set of 1765 facilities is not exhaustive of all government health facilities in Kenya. Rather, this figure represents those located, identified, and georeferenced at the time of this study. The process of defining a comprehensive facility database is ongoing and recent studies have included additional facilities (e.g. Gething et al., 2006).

population from which the counts were generated (Webster et al., 1994; Goovaerts et al., 2005). Where count data are generated at health facilities, one appropriate denominator is the at-risk section of the catchment population of each facility. Such standardisation is of particular interest in a space-time modelling context since it may allow underlying spatial structure to be revealed. Although models have been developed recently to characterise the spatial pattern of government health facility use in four Kenyan districts (Noor et al. 2003; Gething et al. 2004; Noor et al. 2006), data are currently unavailable by which these models can be scaled up to a national level. As such, catchment population estimates are unavailable for most facilities included in the HMIS. In response to this, the current study has defined three alternative space-time modelling frameworks to implement ST-OK as a prediction algorithm to predict MC.

Model 1 represented the null case, in which MC at the q missing facility-months $z_{MC}^*(\beta)$, $\beta = 1, 2, \dots, q$, were predicted directly from the n data $z_{MC}(\alpha)$, $\alpha = 1, 2, \dots, n$, using ST-OK. Models 2 and 3 incorporated the use of TC data as a denominator to MC data. This approach recognised that TC represented a proxy measure of the catchment population of each facility. In Model 2, MC data, $z_{MC}(\alpha)$, were divided by the corresponding TC data, $z_{TC}(\alpha)$, at each known facility-month, to create a new variable termed malaria proportion (MP), $z_{MP}(\alpha) = z_{MC}(\alpha) / z_{TC}(\alpha)$, simply the proportional monthly case load of presumed malaria. ST-OK was then implemented using $z_{MP}(\alpha)$ to obtain predictions $z_{MP}^*(\beta)$ at missing facility-months. The back-conversion of these predictions to MC required corresponding predictions of TC. As such, the TC data $z_{TC}(\alpha)$ were used in a separate ST-OK exercise to predict $z_{TC}^*(\beta)$. MC was then predicted as $z_{MC}^*(\beta) = z_{MP}^*(\beta) \times z_{TC}^*(\beta)$.

Model 3 used TC data in a different way to Model 2. Instead of using individual TC values as denominators for every facility-month, a single, time-invariant, denominator was defined for each facility, referenced by the $k = 1, 2, \dots, K$ facility spatial locations ($\mathbf{u}_k$). This value was the mean monthly total cases (MMTC) per facility, $z_{MMTC}^*(k)$. The rationale for developing MMTC as an alternative denominator to TC was that, by representing an 84-month average rather than a series of monthly values, MMTC may act as a more suitable proxy measure of facility catchment populations than TC. As with Model 2, ST-OK was implemented with the TC data, $z_{TC}(\alpha)$, to predict $z_{TC}^*(\beta)$. 84 monthly TC values were now available for each facility, consisting of $d = 1, 2, \dots, D$ data and $p = 1, 2, \dots, P$ predictions, where $D + P = 84$. $z_{MMTC}^*(k)$, was then calculated for each facility as the temporal mean of these combined data and prediction sets:

$$z_{MMTC}^*(k) = \frac{1}{D+P} \left[\sum_{d=1}^D z_{TC}(\mathbf{u}_k, t_d) + \sum_{p=1}^P z_{TC}^*(\mathbf{u}_k, t_p) \right] \quad (10)$$

Each MC datum, $z_{MC}(\alpha)$, was then divided by the MMTC value for the facility in question, $z_{MMTC}^*(\alpha)$, to create a new variable termed 'standardised malaria cases' (SMC):

$$z_{SMC}(\alpha) = \frac{z_{MC}(\alpha)}{z_{MMTC}^*(\alpha)} \quad (11)$$

ST-OK was then implemented using $z_{SMC}(\alpha)$ to obtain predictions, $z_{SMC}^*(\beta)$, at missing facility-months. The existing $z_{MMTC}^*(\beta)$ values were then used to back-transform SMC predictions to MC, $z_{MC}^*(\beta) = z_{SMC}^*(\beta) \times z_{MMTC}^*(\beta)$.

Implementation of modelling frameworks

Before implementing the three frameworks outlined above, it was necessary to decide whether data from each of the three facility classes (hospitals, health centres, and

dispensaries) should be modelled together or separately. The distributional characteristics (characterised by the mean and variance) of data from hospitals, health centres, and dispensaries were compared and were found to differ substantially between the facility types, suggesting that these data may be most appropriately modelled as belonging to three separate populations. Sample spatial semivariograms were estimated for MC using data from each facility class separately, and for the three classes combined (not shown). Whilst the class-specific semivariograms were indicative of spatial autocorrelation in the MC values, the semivariogram of the combined MC data suggested zero spatial autocorrelation. Modelling frameworks were, therefore, implemented separately and independently for each facility class.

In total, the three modelling frameworks described in the previous section comprised, for each facility class, four individual prediction exercises to predict MC directly (Model 1), TC (Models 2 and 3), MP (Model 2) and SMC (Model 3). Each prediction exercise followed a similar procedure, as follows.

- i. Firstly, the space-time sample semivariogram surface, $\hat{\gamma}_{st}(\mathbf{h}_s, h_t)$ (Fig. 1a) was estimated. This procedure (Equation 1) was executed using a modified space-time GSLIB *gamv* routine (Deutsch and Journel 1998; De Cesare et al. 2002). Semivariograms were modelled up to spatial lags of 100 km and temporal lags of 24 months. Directional (i.e. anisotropic) spatial semivariograms were estimated and found not to vary substantially from the non-directional (isotropic) case. As such, space was considered isotropic. Since the objective was to interpolate (fill in gaps), rather than to extrapolate (predict into the future), time was also considered isotropic (i.e. temporal lag was defined only by the number of months, and not by direction in time). This meant that, when predicting a given missing monthly value, data could be used from months both before and after it in time, thus increasing substantially the number of data available for prediction, especially for those months that occurred early in the study period.
- ii. The product-sum space-time semivariogram model (De Cesare et al. 2001, 2002) was then fitted to the sample semivariogram surface, $\hat{\gamma}_{st}(\mathbf{h}_s, h_t)$. This proceeded in a number of steps. Firstly, the sample spatial, $\hat{\gamma}_s(\mathbf{h}_s)$, and temporal, $\hat{\gamma}_t(h_t)$, semivariograms were estimated from the sample space-time semivariogram surface by setting $h_t = 0$ and $\mathbf{h}_s = 0$, respectively, and conventional semivariogram models were fitted manually to each (Fig. 1b and c). These models were generally nested structures constructed as linear combinations of two or more model components. All models included a nugget component and some combination of spherical, exponential, and hole effect components (see Deutsch and Journel (1998, p. 25) for definitions of these model components). The space-time sill, $C_{st}(0,0)$, was also estimated manually from the sample space-time semivariogram surface and this, along with the parameters of the spatial and temporal semivariograms, provided all the necessary parameters to construct the space-time semivariogram model, $\tilde{\gamma}_{st}(\mathbf{h}_s, h_t)$ (Equations 2-5).
- iii. The space-time semivariogram model (Fig. 1d) was used as input into an ST-OK prediction carried out using the space-time GSLIB *kt3d* routine (Deutsch and Journel 1998) modified to allow prediction of space-time points using product-sum space-time covariance structures (De Cesare et al. 2002).

Comparison of modelling frameworks

Each modelling framework was implemented as a cross-validation whereby the output was a set of n predicted values of MC at the data location, $\{z_{MC^*}(a), a = 1, 2, \dots, n\}$, that could be compared to the MC data themselves, $\{z_{MC}(a), a = 1, 2, \dots, n\}$, at the same locations in order

to assess the predictive accuracy of each model. For Model 1, cross-validation was applied as described to create the cross-validation set $z_{MC}^*(\alpha)$ to compare to $z_{MC}(\alpha)$. For Model 2, cross-validation sets were required for both MP and TC to define the cross-validation set $z_{MC}^*(\alpha) = z_{MP}^*(\alpha) \times z_{TC}^*(\alpha)$. For Model 3, the cross-validation procedure could not be based entirely on predictions made at data locations since the MMTC variable, by definition, requires predictions at unsampled facility-months (Equation 10). As such, MMTC was calculated using both the available TC data $z_{TC}(\alpha)$ and predictions of TC at unsampled facility-months $z_{TC}^*(\beta)$, and a cross-validation set for MC was predicted using the resulting $z_{MMTC}^*(\alpha)$ values in the forward and back-transform between MC and SMC such that $z_{MC}^*(\alpha) = z_{SMC}^*(\alpha) \times z_{MMTC}^*(\alpha)$.

For each cross-validation set defined above for the three modelling frameworks, three summary statistics were calculated to assess the relative performance of each model in predicting MC. These statistics were the correlation coefficient between the predicted and actual set, $\rho[z_{MC}^*(\alpha), z_{MC}(\alpha)]$, the mean prediction error (ME), and the mean absolute prediction error (MAE) (Saito and Goovaerts 2000):

$$MAE = \frac{1}{n} \sum_{\alpha=1}^n |z_{MC}^*(\alpha) - z_{MC}(\alpha)| \quad (12)$$

The correlation coefficient provides a straightforward measure of linear association between the data and prediction sets, the MAE provides a measure of the mean accuracy of individual predictions, and the ME provides a measure of the bias of the predictor.

Space-time stochastic simulation to estimate prediction uncertainty

In addition to providing a modelling framework to predict MC at unsampled facility-months, a further aim of this study was to provide a measure of the uncertainty of these predictions both individually and over aggregated sets of predictions within different space-time regions. Comparison of the cross-validation statistics for the three modelling frameworks indicated that Model 3 produced the most accurate predictions of MC, and was therefore chosen as the framework for which to develop an accompanying uncertainty model. A space-time sGs (ST-sGs) algorithm was implemented to simulate and restore the variance associated with the unknown prediction error component, $R(\mathbf{u}, t)$, by modifying the GSLIB routine *sgsim* (Deutsch and Journel 1998) to incorporate a product-sum space-time covariance structure and to simulate values at space-time locations.

The simulation exercise was implemented in a cross-validation mode that replicated the cross-validation prediction carried out for Model 3, with the two ST-OK procedures that gave predictions of TC at unsampled points, $z_{TC}^*(\beta)$, and cross-validation predictions of SMC at known points, $z_{SMC}^*(\alpha)$, replaced with ST-sGs procedures. The output of the simulation exercise was a set of $l = 1, 2, \dots, L$ conditional realisations of MC, $z_{MC}^{(l)}(\alpha)$, at the $\alpha = 1, 2, \dots, n$ data locations. The ST-sGs algorithm required substantial computation and the number of realisations was therefore limited to $L = 100$. These L simulated sets provided a model of uncertainty for each prediction that could be compared to the known prediction errors determined in the cross-validation for Model 3, allowing assessment of the accuracy of the uncertainty model itself.

Testing the accuracy of the uncertainty model

The L simulated sets were tested as a model for (i) local uncertainty, that is, of predictions of MC at individual facility-months, and (ii) regional uncertainty, that is, of predictions of the regional mean MC per facility-month over aggregated sets of cross-validation predictions within space-time regions.

Each local uncertainty model was summarised by its simulated error standard deviation, $\sigma_{\text{sim}}[\varepsilon(\alpha)]$:

$$\sigma_{\text{sim}}[\varepsilon(\alpha)] = \sigma[\varepsilon^{(l)}(\alpha)], l=1, 2, \dots, L \quad (13)$$

where simulated errors $\varepsilon^{(l)}(\alpha)$ were defined as the difference between each conditional realisation and the corresponding original prediction, $\varepsilon^{(l)}(\alpha) = mc^{(l)}(\alpha) - mc^*(\alpha)$.

It was then necessary to compare the simulated error standard deviations, $\sigma_{\text{sim}}[\varepsilon(\alpha)]$, to estimates of the corresponding actual error standard deviation, $\widehat{\sigma}[\varepsilon(\alpha)]$, where actual error was defined as the difference between each cross-validation prediction and the true data value, $\varepsilon(\alpha) = z_{MC}^*(\alpha) - z_{MC}(\alpha)$. The set of n errors $\varepsilon(\alpha)$, $\alpha = 1, 2, \dots, n$, was partitioned into $b = 1, 2, \dots, B$ subsets or 'bins' according to the magnitude of their corresponding simulated error standard deviations, $\sigma_{\text{sim}}[\varepsilon(\alpha)]$. Each bin spanned $1/B^{\text{th}}$ of the range of values of $\sigma_{\text{sim}}[\varepsilon(\alpha)]$ and B was chosen as 40. Each bin therefore contained a set $\varepsilon(j)$ of $j=1, 2, \dots, J$ error values, each with a corresponding simulated error standard deviation value, $\sigma_{\text{sim}}[\varepsilon(j)]$. For each bin, the median of the J simulated error standard deviation values was compared to the estimated actual error standard deviation, $\widehat{\sigma}_b[\varepsilon(j)]$. This pair of values was obtained for each of the B bins and plotted on a scatter plot to allow visual comparison

A large number of regionally-aggregated sets of $\alpha = 1, 2, \dots, m$ prediction locations were defined using moving space-time windows with spatial radii of between 12.5 km and 100 km and temporal radii of between 3 and 24 months. The size of aggregated sets varied from $m = 2$ to $m = 1000$ individual predictions. For each set, the true regional MC mean,

$\mu[z_{MC}(\alpha)]$, and predicted mean, $\mu[z_{MC}^*(\alpha)]$ were calculated from the data and cross-validation predictions, respectively, and the model of prediction uncertainty was defined by the distribution of the corresponding means of the $l = 1, 2, \dots, L$ simulated realisations of the m predictions, $\mu[z_{MC}^{(l)}(\alpha)]$. Each regional model of uncertainty was summarised by the simulated mean error standard deviation, $\sigma_{\text{sim}}[\mu[\varepsilon(\alpha)]]$:

$$\sigma_{\text{sim}}[\mu[\varepsilon(\alpha)]] = \sigma[\mu[\varepsilon^{(l)}(\alpha)]], l=1, 2, \dots, L \quad (14)$$

where each simulated mean error, $\mu[\varepsilon^{(l)}(\alpha)]$, was defined as the difference between the simulated mean, $\mu[z_{MC}^{(l)}(\alpha)]$, and the corresponding predicted mean, $\mu[z_{MC}^*(\alpha)]$:

$$\mu[\varepsilon^{(l)}(\alpha)] = \mu[z_{MC}^{(l)}(\alpha)] - \mu[z_{MC}^*(\alpha)] \quad (15)$$

The large set of regional simulated error standard deviations for different aggregated sets was compared to estimates of the actual error standard deviation using exactly the same 'binning' approach described above for the local case, resulting in a corresponding scatter plot for the regional case.

RESULTS

Variography

Semivariograms where the nugget component (the intercept on the ordinate of the y -axis) is large relative to the structured component (the distance on the ordinate from the intercept to sill) have a large nugget ratio (NR) which is indicative of a relative lack of autocorrelation in the variable of interest. In these circumstances, kriging variances are expected to be large, and predictions imprecise. Semivariograms in which the sill is reached at short lags are

indicative of autocorrelation only being present over short distances, again generally resulting in relatively imprecise kriging predictions. Spatial and temporal sample semivariograms, along with the fitted models are shown (Fig. 2) for each variable that underwent ST-OK (TC, MC, MP, and SMC) and each facility class. In all cases, the temporal semivariograms differ substantially in structure to the corresponding spatial semivariograms, with temporal semivariograms generally having smaller NRs and smaller sills. Most temporal semivariograms were modelled with a hole effect component to account for a pseudo-periodic structure. This can be interpreted as corresponding to the seasonal nature of malaria transmission in Kenya, where many areas experience distinct increases in malaria incidence in the same months each year. This feature means that values separated in time by 12 months are likely to be more similar than, say, values separated by six months. TC showed the smallest amount of spatial autocorrelation, with the semivariogram for hospitals modelled as a pure nugget effect (zero spatial autocorrelation), although there was substantially larger temporal autocorrelation for all facility classes. MC spatial semivariograms indicated relatively greater spatial autocorrelation (lower NRs) over larger lags than TC and MC temporal semivariograms had less pronounced periodicity than those for TC. MP spatial semivariograms had smaller NRs and displayed autocorrelation over larger lags than TC, but had larger NRs than MC spatial semivariograms for hospitals and health centres. MP temporal semivariograms displayed weak periodicity and had NRs that were similar to those for TC and marginally larger than those for MC. SMC semivariograms generally displayed the largest amount of spatial autocorrelation of the four variables, with the smallest NRs. The difference between the spatial and temporal sills was also least for SMC.

Comparison of modelling frameworks

The results of the cross-validation for each of the three modelling frameworks are shown in Table 1. Model 3 produced predictions of MC which had the smallest mean inaccuracy (smallest MAE) for all three facility classes. Model 2 performed better by this criterion than Model 1 for health centres and dispensaries and worse for hospitals. Results for overall bias (ME) were more mixed. The least biased prediction (smallest ME) was provided by Model 1 for health centres, Model 2 for dispensaries, and Model 3 for hospitals. The largest values of ρ (largest linear associations between predicted and actual values) were provided by Model 1 for hospitals, Model 2 for dispensaries and Model 3 for health centres, although differences in values of ρ between the three models were not substantial. Given these results it was decided that Model 3 was the best overall choice of predictor for MC because it resulted in the smallest mean inaccuracy for all three facility classes and, although its predictions were not the least biased for health centres and dispensaries, the bias in these cases was nevertheless very small.

Uncertainty assessment

The results of the procedure to test the accuracy of the simulated uncertainty model are shown in Figure 3 for both local predictions of MC at individual facility-months and regional predictions of mean MC for sets of between 2 and 1000 facility-months aggregated over space-time neighbourhoods. In the local case (Fig. 3a), simulated error standard deviations replicated closely actual values with no overall tendency for over or under-estimation. Points plotted for smaller standard deviations were progressively less scattered around the 1:1 line, which is indicative of the larger number of values in these bins.

In the regional case (Fig. 3b), there was, again, a strong linear association between simulated and actual error standard deviations in each bin, although there was a tendency for simulated values to be slightly overestimated. This overestimation was more pronounced for larger standard deviations, although was less in relative terms. A simulated error standard

deviation of 19.2 cases, for example, corresponded to an actual error standard deviation of 15.2 cases, representing an over-estimation of 4.0 cases or 26%, whilst a simulated error standard deviation of 71.7 cases corresponded to an actual error standard deviation of 59.6 cases, representing an over-estimation of 12.1 cases or 17%.

DISCUSSION

This study has presented three alternative modelling frameworks in which space-time geostatistical prediction algorithms can be used to predict MC values at missing facility-months within the Kenyan HMIS. Whilst Model 1 used these data in their raw form, Models 2 and 3 used accompanying facility attendance data (TC) to construct a denominator and predictions were made on the resulting standardised variables, MP and SMC, respectively. The rationale was that the spatial structure of the standardised variables may be more pronounced than that of the raw count data, thus, yielding more accurate predictions from the geostatistical algorithms. Since the presence or absence of TC data matched that of MC, however, predictions of MP and SMC required back-transformation by corresponding predictions of the relevant denominator (TC and MMTC, respectively) at unsampled facility-months and, as such, the accuracy of the ultimate predictions of MC was dependent on the prediction accuracies of both the standardised variables and the denominator. Model 2 did not offer a substantial increase in predictive accuracy over Model 1, indicating that the large uncertainty associated with modelling TC negated any benefit of modelling a standardised variable. The modelling framework for Model 3, however, did result in modest increases in prediction accuracy over Model 1. The temporal semivariograms for MC and SMC had almost identical structure which is to be expected since the denominator, MMTC, is constant through time at each spatial location. The benefit of standardising MC by MMTC to obtain SMC can be explained partly by the spatial semivariograms for SMC (Fig. 2) which had smaller NRs and sill values that were much nearer to the corresponding temporal sills than was the case for the MC semivariograms, indicating a relative reduction in the overall variance of the variable across space, of which a greater proportion was autocorrelated. These factors meant SMC could be predicted directly with greater accuracy than could MC. Although the back-transform by MMTC involved further uncertainty, the net effect was that MC was predicted with slightly greater accuracy under this framework than using raw MC data directly in Model 1. The greater spatial structure displayed by SMC emphasises the potential benefit of incorporating measures of facility size and utilisation in models to predict MC. However, this study has shown that, when the only such measures available are themselves incomplete and subject to substantial uncertainty, their inclusion in a predictive model can offer only modest increases in prediction accuracy. Efforts are currently underway in Kenya to both define nationwide facility catchment and utilisation models and compile existing data on individual facility resources such as medical staff and equipment. Model 3 represents a prediction framework to which these new data sources can be added to refine estimation of facility-specific denominators, allowing the more accurate definition and prediction of SMC, and more reliable back-transformation to MC.

As discussed above, the standardisation of raw MC data in models 2 and 3 altered the spatial autocorrelation characteristics of the resulting standardised variables, MP and SMC. The more pronounced spatial autocorrelation of these variables partly explains their potential for more accurate prediction using geostatistical techniques. A further effect of the standardisation procedure was to alter the frequency distributions of the resulting variables such that histograms of both MP and SMC were less positively skewed than that for the raw MC data (not shown). Whilst kriging does not explicitly require distributional assumptions (e.g. that the variable of interest is Normally distributed), the reduction of skewness in the variable of interest can increase prediction accuracy (Saito and Goovaerts, 2000), and this

effect is likely to have contributed further to the increases in prediction accuracy observed in Model 3.

It is important to note that the standardised variables MP and SMC represent subtly different outcomes. MP (Model 2) represents the proportional monthly case load of presumed malaria at a given facility-month, affected both by the count of presumed malaria cases seen at a facility in a given month, and by monthly fluctuations in the total number of outpatients attending. In contrast, SMC (Model 3) is derived by dividing monthly MC values by MMTTC, a time-invariant denominator indicative of the overall level of outpatient use of a given facility. In this study, the two standardised variables were defined as a means to an end, used in modelling frameworks in which the ultimate goal was the prediction of unknown MC values. As such, it was their contrasting statistical properties that were of principal interest. In other applications, however, the interest may lie in the standardised variables themselves, in which case their differing interpretations may be of principal importance.

The decision to perform the modelling procedure separately for each of the three facility classes was taken because, in contrast to the sample spatial semivariograms estimated for each facility class separately, the equivalent semivariogram for the combined data displayed zero autocorrelation. This effect occurred because, whilst monthly MC values from facilities of the same class tended to be spatially dependent (spatially proximate values were likely to be more similar than those more separate), values from different facility classes often varied substantially, even over very short distances. Treating data from hospitals, health centres, and dispensaries separately allowed this confounding effect to be avoided. It is unlikely, however, that no dependencies exist between data from the three sets. Similarities in population, malaria, and health system factors in a given locality mean that MC values from a given facility are likely to contain information about those in neighbouring facilities, even if it is of a different facility class. A refinement to the modelling approach presented here would be to attempt to exploit this information; for example, using a space-time cokriging system to predict missing values.

Modelling output in context

Predictions of space-time variables derived from developing-world outpatient data are likely to have a large inherent uncertainty associated with them and this is reflected in this study in both the semivariograms (Fig. 2) and model outputs. At the level of individual facility-months, predictions with the accuracies presented in Table 1 are likely to be of only limited use to health system decision-makers (MAE is 26.8%, 27.6%, and 22.9% of the mean MC value for hospitals, health centres and dispensaries, respectively). Strategic decision-making is rarely made at this level, however, and the accuracy of predictions of mean MC burdens (which can then be translated directly into total MC burdens) at monthly and annual district, provincial, and national levels are of greater importance. Predictions of mean MC at these levels entail the aggregation of many individual predictions (along with existing data) across space-time regions. Although the expectation of the mean prediction error (bias, or ME) remained constant (0.4 cases across all facility classes) at different levels of aggregation, the variance of this mean decreased (as is expected) as progressively larger aggregated sets were considered. The overall prediction error standard deviation of unaggregated predictions was 181.4 cases. The error standard deviation for predictions of mean MC for aggregated sets of 200 predictions (corresponding approximately to aggregation at the district-year level) was 6.3 cases (2.2% of mean MC), and for sets of 1500 predictions (corresponding approximately to aggregation at the province-year level) this value was 1.8 cases (0.6% of mean MC). As 95% of prediction errors at these levels can be expected to fall within two such standard deviations of the mean error, predictions at these levels of aggregation are sufficiently accurate to be of real value to decision-makers.

Of equal importance to providing accurate predictions, however, is the provision of accompanying estimates of this prediction uncertainty. These were provided in this study by an uncertainty model derived from an adapted stochastic simulation algorithm. The results presented above indicated that this model provides accurate estimates of individual (local) prediction uncertainty. For aggregated (regional) predictions, the model marginally over-estimates uncertainty such that, where the true error standard deviation in the above two examples was 6.3 and 1.8 cases (representing 2.2% and 0.6% of the MC mean), the model predicted corresponding standard deviations of 9.4 and 4.4 cases (3.2% and 1.5%). Given that the over-estimation of prediction uncertainty is likely to be preferable to under-estimation, and that the difference between actual and modelled uncertainty is small, this model can provide measures of prediction uncertainty that further enhance the value of the predictions to decision makers.

Having presented methods for predicting MC values at facility-months with missing HMIS records, and assessed the likely accuracy of these predictions, it is important to reconsider the interpretation of these values. Of crucial importance is the distinction between MC (counts of presumed malaria at government health facilities) and the burden of malaria in the population. The issue of low utilisation of formal health facilities by care seekers is well documented in many developing-world settings (Foster, 1995; Molyneux et al., 2002). In the current context, this leads to only a small proportion of malaria episodes in the population resulting in a visit to a government health facility (McCombie, 2002; Molyneux et al., 1999; Amin et al., 2003). A further issue is misdiagnosis. The widespread lack of laboratory resources with which to confirm diagnoses in many developing-world settings (Zurovac et al., 2002, 2006), and the generally accepted alternative that all febrile children in high risk areas be considered to have, and be treated for, malaria (WHO, 1997; Bloland et al., 2003), mean that malaria is often over-diagnosed at health facilities, with many patients given false-positive diagnoses. Both under-utilisation and misdiagnosis mean that predictions of MC should not be used directly to evaluate the burden of malaria in a given population. However, these factors do not reduce the importance of MC values as a metric for resource planning. MC should be interpreted as quantifying the number of diagnoses that have been made for malaria and, importantly, the number of malaria treatments that have been administered. Despite the disparities between MC and the true pattern of population and outpatient malaria morbidity, MC remains critical for health-service planning because it determines the level of resources required to treat patients under this diagnosis. There remains an urgent need to upgrade HMIS systems in many developing-world nations by reducing the extent of missing data to allow robust quantification of treatment burdens. Whilst the methods presented in this paper in no way reduce this ultimate requirement, we believe that the development of statistical models that can mitigate the uncertainties caused by missing data substantially enhances the utility to health-service decision-makers of existing HMIS data.

CONCLUSION

The Kenyan HMIS database on the number of outpatients treated for malaria across the country is 57% incomplete, preventing the quantification of key public-health statistics. This study has presented a geostatistical space-time modelling framework that uses the available data to predict the monthly count of treatments for malaria at all government health facilities where data are missing. Three different modelling frameworks were developed and tested, each comprising one or more space-time kriging procedures. The model that predicted malaria cases most accurately in a cross-validation procedure was identified and a space-time stochastic simulation approach was used to develop a corresponding model of prediction uncertainty. These models were tested and found to give predictions and uncertainty estimates at an accuracy acceptable for public-health decision-making at district,

provincial, and national scales. Work undertaken to implement the modelling approaches developed in this paper (Gething et al. 2006) has provided recent estimates of the national treatment burden for outpatient malaria in the Kenyan government's formal health sector.

Acknowledgments

This study received financial support from The Wellcome Trust, UK (#058992), the Roll Back Malaria Initiative, AFRO (AFRO/WHO/RBM # AF/ICP/CPC/400/XA/00) and the Kenya Medical Research Institute (KEMRI). The authors are grateful to Dr. James Nyikal, the Director of Medical Services, for his support and policy framework for our work and Dr. Esther Ogara, MoH-HMIS, for facilitating the acquisition of the outpatient data. The authors are also grateful to Briony Tatem for her dedicated assistance in formatting the dataset and Professor David Rogers for helping with a Quick Basic programme to ordinate digital HMIS records for import into Access. PW Gething gratefully acknowledges support from the EPSRC through the School of Electronics and Computer Science and from the School of Geography, University of Southampton. SIH is a Research Career Development Wellcome Trust Fellow (#056642). RWS is a Senior Wellcome Trust Fellow (#058992). This paper is published with the permission of the director, KEMRI.

LITERATURE CITED

- AbouZahr C, Boerma T. Health Information Systems: The Foundations of Public Health. Bulletin of the World Health Organization. 2005; 83:578–583. [PubMed: 16184276]
- Al Laham H, Khoury R, Bashour H. Reasons for Underreporting of Notifiable Diseases by Syrian Paediatricians. Eastern Mediterranean Health Journal. 2001; 7:590–596. [PubMed: 15332753]
- Amin AA, Marsh V, Noor AM, Ochola SA, Snow RW. The use of formal and informal curative services in the management of paediatric fevers in four districts of Kenya. Tropical Medicine and International Health. 2003; 8:1143–1152. [PubMed: 14641851]
- Bloand PB, Kachur SP, Williams HA. Trends in antimalarial drug deployment in subSaharan Africa. Journal of Experimental Biology. 2003; 206:3761–3769. [PubMed: 14506211]
- Cressie N, Huang HC. Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions. Journal of the American Statistical Association. 1999; 448:1330–1340.
- De Cesare L, Myers DE, Posa D. Estimating and Modeling Space-Time Correlation Structures. Statistics and Probability Letters. 2001; 51:9–14.
- De Cesare L, Myers DE, Posa D. FORTRAN Programs for Space-Time Modeling. Computers and Geosciences. 2002; 28:205–212.
- De Iaco S, Myers DE, Posa D. Nonseparable space-time covariance models: some parametric families. Mathematical Geology. 2002; 34:23–42.
- Deutsch, CV.; Journel, AG. GSLIB: geostatistical software library and user's guide. 2nd ed.. Oxford University Press; New York: 1998.
- Dimitrakopoulos, R.; Luo, X. Spatiotemporal Modeling: Covariances and Ordinary Kriging Systems. In: Dimitrakopoulos, R., editor. Geostatistics for the Next Century. Kluwer Academic Publishers; Dordrecht: 1994. p. 88–93.
- Evans T, Stansfield S. Health Information in the New Millennium: A Gathering Storm? Bulletin of the World Health Organization. 2003; 81:856–856. [PubMed: 14997237]
- Foster S. Treatment of malaria outside the formal health services. Journal of Tropical Medicine and Hygiene. 1995; 98:29–34. [PubMed: 7861477]
- Gething PW, Noor AM, Gikandi PW, Ogara E, Hay SI, Nixon MS, Snow RW, Atkinson PM. Improving Imperfect Health Management Information System Data In Africa Using Space-Time Geostatistics. PLoS Medicine. 2006; 3:825–831.
- Gething PW, Noor AM, Zurovac D, Atkinson PM, Hay SI, Nixon MS, Snow RW. Empirical Modelling of Government Health Service Use by Children With Fevers in Kenya. Acta Tropica. 2004; 91:227–237. [PubMed: 15246929]
- Gneiting T. Nonseparable, stationary covariance functions for space-time data. Journal of the American Statistical Association. 2002; 97:590–600.

- Gneiting, T.; Genton, MG.; Guttorp, P. Geostatistical Space-Time Models, stationarity, separability and full symmetry. Department of Statistics, University of Washington; 2005. Technical Report no. 475
- Goovaerts, P. Geostatistics for Natural Resource Evaluation. Oxford University Press; New York: 1997.
- Goovaerts P, Jacquez GM, Greiling D. Exploring Scale-Dependent Correlations Between Cancer Mortality Rates Using Factorial Kriging and Population-Weighted Semivariograms. *Geographical Analysis*. 2005; 37:152–182. [PubMed: 16915345]
- Journel, AG.; Huijbregts, CJ. Mining Geostatistics. Academic Press; London: 1978.
- Kindermans, J. Changing National Malaria Treatment Protocols in Africa: What Is the Cost and Who Will Pay?; RBM Partnership Meeting on Improving Access to Antimalarial Treatment; Medicins Sans Frontieres: Geneva. 30 September-2 October 2002; 2002.
- Kolovos A, Christakos G, Hristopulos DT, Serre ML. Methods for generating non-separable spatiotemporal covariance models with potential environmental applications. *Advances in Water Resources*. 2004; 27:815–830.
- Kyriakidis PC, Journel AG. Geostatistical Space-Time Models: a Review. *Mathematical Geology*. 1999; 31:651–684.
- Lawson, AB. Statistical Methods in Spatial Epidemiology. Wiley; Chichester: 2001.
- Ma C. Spatio-temporal variograms and covariance models. *Advances in Applied Probability*. 2005; 37:706–725.
- Matheron, G. Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau. Ecole Nationale Supérieure des Mines de Paris; Paris: 1971. The Theory of Regionalized Variables and Its Applications.
- McCombie SC. Self-treatment for malaria: the evidence and methodological issues. *Health Policy and Planning*. 2002; 17:333–344. [PubMed: 12424205]
- MoH Kenya. Health Management Information Systems: Report for the 1996 to 1999 Period. Ministry of Health; Republic of Kenya: Apr. 2001 2001
- MoH Kenya. Transition Plan for Implementation of Artemisinin-Based Combination Therapy (ACT) Malaria Treatment Policy in Kenya. Ministry of Health; Republic of Kenya: Jul. 2005 2005
- Molyneux CS, Murira G, Masha J, Snow RW. Intra-household relations and treatment decision-making for childhood illness: A Kenyan case study. *Journal of Biosocial Science*. 2002; 34:109–131. [PubMed: 11814209]
- Molyneux CS, Mung'ala-Odera V, Harpham T, Snow RW. Maternal responses to childhood fevers: a comparison of rural and urban residents in coastal Kenya. *Tropical Medicine and International Health*. 1999; 4:836–845. [PubMed: 10632992]
- Murray CJL, Lopez AD, Wibulpolprasert S. Monitoring Global Health: Time for New Solutions. *British Medical Journal*. 2004; 329:1096–1100. [PubMed: 15528624]
- Noor, AM. PhD. Thesis Submitted to The Open University. 2005. Developing Spatial Models of Health Service and Utilisation to Define Health Equity in Kenya. 2005, 258 Pages
- Noor AM, Amin AA, Gething PW, Atkinson PM, Hay SI, Snow RW. Modelling Distances Travelled to Government Health Services in Kenya. *Tropical Medicine and International Health*. 2006; 11:188–196. [PubMed: 16451343]
- Noor AM, Gikandi PW, Hay SI, Muga RO, Snow RW. Creating Spatially Defined Databases for Equitable Health Service Planning in Low-Income Countries: the Example of Kenya. *Acta Tropica*. 2004; 91:239–251. [PubMed: 15246930]
- Noor AM, Zurovac D, Hay SI, Ochola SA, Snow RW. Defining Equity in Physical Access to Clinical Services Using Geographical Information Systems As Part of Malaria Planning and Monitoring in Kenya. *Tropical Medicine and International Health*. 2003; 8:917–926. [PubMed: 14516303]
- Rodriguez-Iturbe I, Mejia JM. Design of Rainfall Networks in Time and Space. *Water Resources Research*. 1974; 10:713–728.
- Rudan I, Lawn J, Cousens S, Rowe AK, Boschi-Pinto C, Tomaškovi L, Mendoza W, Lanata CF, Roca-Feltre A, Carneiro I, Schellenberg JA, polašek O, Weber M, Bryce J, Morris SS, Black RE, Campbell H. Gaps in Policy-Relevant Information on Burden of Disease in Children: A Systematic Review. *Lancet*. 2005; 365:2031–2040. [PubMed: 15950717]

- Saito H, Goovaerts P. Geostatistical Interpolation of Positively Skewed and Censored Data in a Dioxin-Contaminated Site. *Environmental Science and Technology*. 2000; 34:4228–4235.
- Stansfield S. Structuring Information and Incentives to Improve Health. *Bulletin of the World Health Organization*. 2005; 83:562–563. [PubMed: 16184268]
- Stein ML. Space-time covariance functions. *Journal of the American Statistical Association*. 2005; 100:310–321.
- UN. United Nations Millennium Declaration. United Nations General Assembly; 2000. A/RES/55/2. (www.un.org)
- Webster R, Oliver MA, Muir KR, Mann JR. Kriging the local risk of a rare disease from a register of diagnoses. *Geographical Analysis*. 1994; 26:168–185.
- WHO. Integrated Management of Childhood Illnesses Adaptation Guide. Part 2. C. Technical basis for adapting clinical guidelines, feeding recommendations, and local terms. World Health Organisation; Geneva: 1997.
- WHO. World Malaria Report 2005. Prepared by Roll Back Malaria, World Health Organization and United Nations Children Fund; Geneva, Switzerland: 2005. WHO/HTM/MAL/2005.1102
- WHO. AFRO. Integrated Disease Surveillance Strategy, a Regional Strategy for Communicable Diseases 1999-2003. World Health Organisation Regional Office for Africa; Geneva: 1999.
- WHO. SEARO. Strengthening of Health Information Systems in Countries of the South-East Asia Region. World Health Organization Regional Office for South -East Asia; New Delhi: 2002. Report of an Intercountry Consultation
- Zurovac D, Midia B, Ochola SA, English M, Snow RW. Microscopy and outpatient malaria case management among older children and adults in Kenya. *Tropical Medicine and International Health*. 2006; 11:432–440. [PubMed: 16553926]
- Zurovac, D.; Midia, B.; Ochola, SA.; Barake, Z.; Snow, RW. Evaluation of Malaria Case Management of Sick Children Presenting in Outpatient Departments in Government Health Facilities in Kenya. Division of Malaria Control, Ministry of Health; Republic of Kenya, Nairobi: 2002.

Figure 1.

Space-time variography of the malaria cases (MC) variable for health centres as an example of the modelling procedure for a product-sum space-time semivariogram. Firstly, a sample space-time semivariogram surface (a) is estimated from the available data. The sample spatial (b) and temporal (c) semivariograms are then estimated from this surface (circles), and a 1-d semivariogram model (lines) is fitted to each. The parameters of these models, along with an estimate of the 'sill' of the space-time semivariogram are then used to define the 2-d product-sum semivariogram model (d) as defined in the text. For each of the four plots shown, vertical axes measure semivariance, γ , and horizontal axes measure either spatial lag (h_s) or temporal lag (h_t).

Figure 2.

Spatial and temporal sample semivariograms (dots) derived from space-time sample semivariogram surfaces and 1-d semivariogram models (lines) for (a) hospitals, (b) health centres, and (c) dispensaries. The variables shown for each facility class are monthly outpatient total cases (TC), malaria cases (MC), malaria proportion (MP) and standardised malaria cases (SMC) as described in the text for government health facilities across Kenya.

Figure 3.

Comparison of simulated and actual standard deviations of prediction errors for malaria cases (MC). Simulated standard deviations were derived for individual and aggregated prediction of MC via sequential-Gaussian-simulation and corresponding actual errors were obtained from a cross-validation exercise. Prediction errors were divided into bins according to their simulated standard deviation, and the actual standard deviation of the set of errors in each bin was calculated (circles) along with the 95% confidence interval (vertical bars). Results are shown for (a) predictions of MC at individual facility-months and (b) predictions of mean MC within sets of between 2 and 1000 facility-months created by aggregating points within progressively larger space-time neighbourhoods.

Table 1

Comparison of summary statistics for cross-validation predictions of malaria cases using three different modelling frameworks. Predictions were made separately for hospitals, health centres and dispensaries. The statistics shown are the correlation coefficient, ρ , the mean error (ME) and mean absolute error (MAE), as described in the text. Model 3 (highlighted in bold text) was chosen as the best overall predictor of malaria cases

Facility type	Model	ρ	ME	MAE
Hospitals	Model 1	0.859	4.439	193.188
	Model 2	0.848	6.244	205.730
	Model 3	0.856	2.822	192.423
Health Centres	Model 1	0.779	0.416	92.067
	Model 2	0.783	-2.179	90.240
	Model 3	0.789	-1.050	89.042
Dispensaries	Model 1	0.764	0.530	69.527
	Model 2	0.776	-0.397	67.156
	Model 3	0.774	-0.638	66.903