An Application of the Holistochastic Human Exposure Methodology to Naturally Occurring Arsenic in Bangladesh Drinking Water

M. L. Serre,[1]* A. Kolovos,[1] G. Christakos,[1] and K. Modis[2]

The occurrence of arsenic in drinking water is an issue of considerable interest. In the case of Bangladesh, arsenic concentrations have been closely monitored since the early 1990s through an extensive sampling network. The focus of the present work is methodological. In particular, we propose the application of a holistochastic framework of human exposure to study lifetime population damage due to arsenic exposure across Bangladesh. The Bayesian Maximum Entropy theory is an important component of this framework, which possesses solid theoretical foundations and offers powerful tools to assimilate a variety of knowledge bases (physical, epidemiologic, toxicokinetic, demographic, etc.) and uncertainty sources (soft data, measurement errors, etc.). The holistochastic exposure approach leads to physically meaningful and informative spatial maps of arsenic distribution in Bangladesh drinking water. Global indicators of the adverse health effects on the population are generated, and valuable insight is gained by blending information from different scientific disciplines. The numerical results indicate an increased lifetime bladder cancer probability for the Bangladesh population due to arsenic. The health effect estimates obtained and the associated uncertainty assessments are valuable tools for a broad spectrum of end-users.

KEY WORDS: Holistochastic; exposure; arsenic; BME; health effect

1. INTRODUCTION

Arsenic (*As*) is a known toxic and carcinogenic substance that can be found in groundwater in reduced (As_{III}) and oxidized (As_V) forms. This can be the result of human-induced pollution (e.g., Cherry Point, NC[1]), or a natural event, as is the case of *As* in Bangladesh where the *As* contained in the sediments dissolves in water aquifers.[2] If dissolved or desorbed in sufficient quantity in drinking water, it can pose a threat for public health. Long-term exposure to *As* via drinking water is known to cause skin, lung, bladder, and kidney cancers, as well as skin lesions, namely, pigmentation changes and hyperkeratosis.[3]

As distributions in space and time and the consequent health effects due to human exposure to the contaminant have been at the epicenter of interest in various studies.[1,4,5] The development of the *Bayesian Maximum Entropy* theory (BME)[6] has offered a set of powerful techniques for a *holistochastic* human exposure analysis across space-time. The term "holistochastic" was coined by Christakos and Hristopulos[7] to refer to a modeling framework that accounts for both the holistic and the stochastic features of the human exposure situation. The holistic component expresses the multidisciplinary nature of

[1] Center for the Advanced Study of the Environment, University of North Carolina at Chapel Hill, NC 27599-7431.

[2] Department of Mining and Metallurgical Engineering, National Technical University of Athens, Greece.

* Address correspondence to A. Kolovos, University of North Carolina at Chapel Hill, Environmental Sciences and Engineering, Room 104 Rosenau Hall, CB#7431, Chapel Hill, NC 27599-7431; tel.: 919-966-1173; fax: 919-966-7911; kolovos@email.unc.edu.

human exposure such that the whole is greater than the sum of its parts; the stochastic component, on the other hand, emphasizes the importance of uncertainty characterization. The BME theory allows meaningful spatiotemporal inferences to be drawn and exposure maps to be produced based on a sound logical system that assimilates a wealth of *knowledge bases* (KB), many of which were previously unaccounted for in human exposure studies. These KB include physical and biological laws, toxicokinetic principles, empirical relations, multiple-point and nonlinear statistics, as well as uncertain data available in the form of soft information. The BME viewpoint offers valuable tools in a variety of environmental health science applications. In this work, we apply BME to study the spatial distribution of high *As* concentrations in the Bangladesh drinking water and assess the resulting population health damages.

The presence of *As* in Bangladesh groundwater has gained considerable attention during the last years,[8,9] mainly since the British Geological Survey (BGS) started monitoring the problem in the 1990s. In particular, during 1998–1999, BGS gathered samples from 3,534 wells throughout Bangladesh by means of local and national surveys (see the BGS Bangladesh website http://www.bgs.ac.uk/ arsenic/Bangladesh.html). The data sets mainly focus on shallow aquifers (i.e., depths smaller than 200 m). Deeper aquifers have been found to be almost *As*-free, and according to the final report there is no reason to believe that they will become seriously contaminated.[2] The World Health Organization (WHO) standards for *As* concentration in drinking water include a maximum of 10 μg/L. However, even with the significantly looser Bangladesh standard of 50 μg/L, contaminant concentrations were found to exceed the acceptable limits in many sampling locations.

About 4,000 tubewell and borehole samples from four different surveys comprise the data volume of the Bangladesh study.[2] These samples provided a total of 3,373 *As* concentration measurements (after excluding the ones from deep aquifers). The spatial distribution of the samples is rather uniform. There also appears to be some short-term temporal variation in *As* concentration in the scale of weeks or less, possibly due to changes in water-flow patterns and strong stratification of water quality within the same aquifer. Furthermore, an intense spatial nonhomogeneity of *As* concentrations has been observed, resulting in samples that greatly vary even when located a few meters apart. This effect has been local in some areas, whereas in other areas the variation is

smoother. The situation is attributed to the nature of the underlying sediments where *As* is found.

In the majority of the collected BGS samples, the *As* concentration was measured using a typical method (atomic fluorescence spectrometry with HG-AFS hybrid generation) at the BGS laboratories. These values are considered free of significant measurement errors. However, some *As* measurements have been reported to lie between 0 μg/L and some given threshold value (equal to either 6 or 30 μg/L, depending on the BGS sample considered). Moreover, a significant number of measurements were below the detection limit of 1 μg/L. These sorts of measurements contain a higher degree of uncertainty and constitute prime examples of soft data. Naturally, single-value measurements may also range within some narrow limits (depending on the instrumentation accuracy), but this level of uncertainty is negligible compared to that of soft data. In view of the above considerations, accounting for soft information is a crucial matter because of the technical difficulties in measuring accurately low *As* levels in water samples, and so forth. Nevertheless, soft data is very valuable because it can improve exposure estimates that would have been otherwise obtained with less information at hand (e.g., using only single-valued hard data). Indeed, as has been noticed in the relevant literature, the uncertain knowledge obtained about important exposure parameters could be more valuable than the certain knowledge obtained about less important parameters.[6] Handling soft data requires a scientifically different treatment than hard data because of the soft data form and uncertainty expression, for example, some soft data available in the present study are in the form of interval data (i.e., it is assumed that all values in an interval have the same probability of occurrence). The BME techniques of space-time exposure mapping possess some unique features capable of integrating and processing a variety of knowledge bases (including hard and soft data) in a scientifically meaningful and mathematically rigorous manner. In the following section we briefly describe some basic features of BME and place it within the general framework of holistochastic human exposure and population health effect considered in this work.

2. THE HOLISTOCHASTIC HUMAN EXPOSURE FRAMEWORK

2.1. The Spatiotemporal BME Mapping Method

The *As* data set is characterized by considerable natural variability in space-time, as well as by varying

levels of uncertainty in the measured values (records contaminated with measurement noise, values below detection limits, etc.). In such situations it is useful to adopt a stochastic representation of the *As* distribution across space-time in terms of the random field model, $X(\boldsymbol{p})$, where the vector $\boldsymbol{p} = (\boldsymbol{s}, t)$ defines a point in the space-time domain ($\boldsymbol{s} = (s_1, s_2)$ is the spatial location vector with coordinates s_1 and s_2, and t denotes time).[6] The uncertainty of the *As* distribution manifests itself as an ensemble of realizations $\{\chi\}$ of the possible $X(\boldsymbol{p})$ values. To each of these realizations the S/TRF assigns a probability depending on $\boldsymbol{p}$. BME considers all available KB and using the provisions of the underlying theory it generates informative probability distributions of *As* concentrations at unsampled locations. This is achieved through an epistemic process that distinguishes between three main BME stages.[6]

1. At the first (structural) stage, the general KB (briefly denoted as $\mathcal{G}$-KB) describes the structural characteristics of the $X(\boldsymbol{p})$ field. The $\mathcal{G}$-KB includes theoretical and empirical expressions of these characteristics. For example, $\mathcal{G}$-KB may include the mean trend $m_x(\boldsymbol{p}) = \overline{X(\boldsymbol{p})}$ (the bar denotes stochastic expectation), which characterizes systematic spatiotemporal *As* patterns in the Bangladesh groundwater. Often, space-time dependencies between *As* values are expressed for sets of points considered simultaneously. The $\mathcal{G}$-KB includes multiple-point statistics across space and time. For example, a two-point statistic is the covariance function $c_x(\boldsymbol{p}, \boldsymbol{p}') = \overline{[X(\boldsymbol{p}) - m_x(\boldsymbol{p})][X(\boldsymbol{p}') - m_x(\boldsymbol{p}')]}$, which expresses spatiotemporal correlation between points $\boldsymbol{p}$ and $\boldsymbol{p}'$. As a result of accounting for $\mathcal{G}$-KB, the relevant probabilities of the *As* values are assigned to each point $\boldsymbol{p}$. The outcome is a map of probability density functions (PDF) that quantifies the distribution of these probabilities. The structural PDF is defined by

$$f_\mathcal{G}(\chi; \boldsymbol{p})d\chi = \text{Prob}[\chi \leq X(\boldsymbol{p}) \leq \chi + d\chi];$$
(1)

where the subscript "$\mathcal{G}$" denotes that the general KB available regarding the *As* distribution has been used to construct the PDF in Equation (1).

2. At the second (specificatory) stage, BME makes an assessment of the monitoring data χ_{data} obtained at a specific set of sampling points $\boldsymbol{p}_i$ ($i = 1, \ldots, m$). The χ_{data} set constitutes the so-called specificatory (or site-specific) KB ($\mathcal{S}$-KB). The $\chi_{\text{data}} = (\chi_{\text{hard}}, \chi_{\text{soft}})$ consists of the hard data χ_{hard} (exact *As* measurements) and the soft data χ_{soft} at sampling points where *As* could not be measured accurately. Soft data may consist of intervals (e.g., below-detect measurements indicating *As* concentrations between 0 and the detection limit of the measuring device); or they may involve probabilistic descriptions of the measured values (as provided by devices with calibrated measurement errors). BME is the only method that can incorporate all types of information conveyed by soft data (as opposed, e.g., to considering merely the middle point of a soft interval datum as a hard datum and ignoring the interval range and the associated uncertainty) and, thus, it can provide higher quality estimates.

3. At the final (integration) stage BME blends the total KB, $\mathcal{K} = \mathcal{G} \cup \mathcal{S}$, described above and produces an updated (integration or posterior) PDF, $f_\mathcal{K}(\chi_k)$, at each mapping point $\boldsymbol{p}_k$. Then, estimates $\hat{\chi}_k$ of the *As* field χ_k at any point $\boldsymbol{p}_k$ can be derived from $f_\mathcal{K}(\chi_k)$. The $f_\mathcal{K}(\chi_k)$ provides a complete stochastic description of the *As* distribution that accounts for the total knowledge $\mathcal{K}$ available. To produce spatiotemporal exposure maps, a variety of estimators can be derived at the integration stage using f_k. The choice depends on the intended use, e.g., if the most probable estimate is sought for at each location $\boldsymbol{p}_k$, the estimator of choice will be the mode (*BMEmode*) of the distribution f_k. In this work we used the median of f_k (*BMEmedian*) to obtain the map of the *As* distribution over Bangladesh. In addition, the estimation standard deviation at $\boldsymbol{p}_k$ was chosen as a measure to characterize the corresponding mapping uncertainty.

The different processing stages of BME mapping are outlined in Fig. 1. At the structural stage, the $\mathcal{G}$-KB is processed and the structural (or prior) PDF $f_\mathcal{G}$ representing the general characteristics of subsurface *As* are calculated in the proper space-time continuum. At the specificatory stage, the $\mathcal{S}$-KB is considered and transformed into appropriate operators $\Xi_\mathcal{S}$.[6] For example, in the case of interval soft data the $\Xi_\mathcal{S}$ is such that $d\Xi_\mathcal{S} = d\chi_{\text{soft}}$, whereas for probabilistic
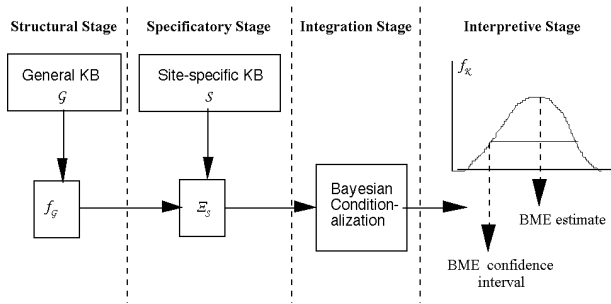
Fig. 1. The BME-based knowledge assimilation and processing stages.

data $d\Xi_S = f_S(d\chi_{soft})d\chi_{soft}$. The S-KB is properly assimilated at the integration stage, thus resulting in the posterior PDF, f_k. BME theory involves elaborate numerical computations to obtain f_k at a set of estimation locations $\boldsymbol{p}_k$, which are carried out using the *BMElib*.[10]

2.2. The Stages of the Holistochastic Approach

In this work we employ a comprehensive spatiotemporal holistochastic framework of human exposure due to high *As* concentrations in Bangladesh drinking water. The framework is summarized in Fig. 2. In the previous section we examined the BME
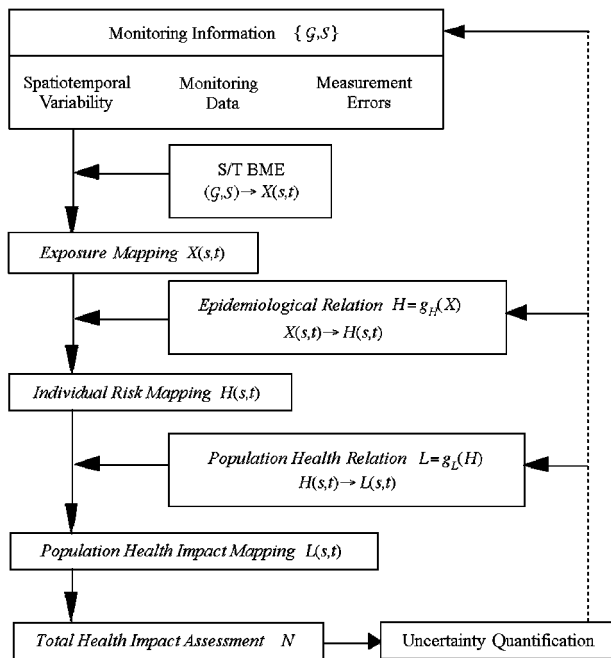

Fig. 2. A comprehensive holistochastic framework for human exposure and health risk assessment across space and time.

method, which is part of this framework and provides valuable information regarding the contaminant distribution across space and time based on the available KB. Each stage of the holistochastic framework produces maps that allow a better assessment of *As* health risk and its impact at the population level, thus providing useful information for health policy planning, administration, and financial management. In brief, the holistochastic framework consists of the following stages.

1. Generate exposure maps via the BME approach by assimilating all relevant physical KB (Fig. 1).
2. Integrate epidemiologic KB about *As* adverse health effects to describe the expected health risk across space and time.
3. Account for demographics and population dynamics to obtain maps of the distribution of the population health damage (or loss). These maps are readily used to assess the global health impact of *As* exposure in Bangladesh.

This holistochastic approach also accounts for an important aspect of risk assessment, which is to properly quantify the uncertainties involved in each stage above.

The population of Bangladesh is exposed to *As* in drinking water coming from shallow aquifers. Due to the lack of comprehensive water-treatment facilities, we can reasonably assume that the *As* concentration field $X(\boldsymbol{p})$ adequately represents the actual drinking water *As* exposure. At the Stage 1 above, the BME mapping method rigorously incorporates and processes the physical $\mathcal{G}$- and $\mathcal{S}$-KB available about *As*. This results in maps of the *As* exposure distribution $X(\boldsymbol{p})$, that is:

$$\{\mathcal{G}, \mathcal{S}\} \rightarrow X(\boldsymbol{p}) \qquad (2)$$

The estimated exposure field in Equation (2) is described at any point $\boldsymbol{p}_k$ of the mapping grid by the BME posterior PDF $f_k(\chi_k)$, which offers a complete stochastic description of the associated uncertainty. This uncertainty is owed to the natural spatial variability of *As* and the measurement errors of the available data. The BME method leads to more accurate maps than those obtained by classical methods, as has been previously demonstrated in similar situations involving soft data.[10,11]

In Stage 2, we focus on the health risk posed by the *As* exposure in drinking water from shallow wells. Normally, this task is undertaken using toxicokinetic (or pollutokinetic) models that evaluate

the contaminant burden on target organs, as well as models providing burden-health response relationships (see the approach proposed by Christakos and Hristopulos[7]). However, currently there exists insufficient knowledge of the *As* mode of action, and the main information sources are epidemiologic or toxicologic exposure-health response studies.[12] On the basis of these studies, empirical models of the form $H = g_H(X)$ can be constructed describing the stochastic relationship between *As* exposure, $X(\boldsymbol{p})$, and the resulting health risk, $H(\boldsymbol{p})$, across space and time, that is:

$$X(\boldsymbol{p}) \xrightarrow{g_H(\cdot)} H(\boldsymbol{p}) \qquad (3)$$

The health risk $H(\boldsymbol{p})$ employed in this work is the lifetime probability that an individual at a certain point $\boldsymbol{p} = (\boldsymbol{s}, t)$ will develop the specified health outcome (e.g., lifetime incidence of bladder cancer resulting from chronic exposure to *As* in drinking water). The uncertainty in the estimated health risk $H(\boldsymbol{p})$ is the combination of exposure $X(\boldsymbol{p})$ mapping uncertainty with uncertainty in the empirical model $g_H(\cdot)$.

In Stage 3, we integrate population information (demographics, population dynamics and migration, water consumption habits, etc.) to study the population health impact resulting from *As* exposure. In this case, the population health impact is considered to be the damage (or loss) $L(\boldsymbol{p})$ on the population due to the health risk $H(\boldsymbol{p})$. We also focus on $L_{As}(\boldsymbol{p})$, which is the population health damage taken as the incidence density of the lifetime health outcome attributed purely to *As* (e.g., the increase in the number of people per km^2 who will develop lifetime bladder cancer due to *As* exposure). The relationship $L = g_L(H)$ describing the population loss due to health risk is a function of population density, and it may also account for other population characteristics (such as daily range of displacement or water consumption habits). Once the relationship $g_L(\cdot)$ has been established, we infer the distribution of population health damage as follows:

$$H(\boldsymbol{p}) \xrightarrow{g_L(\cdot)} L(\boldsymbol{p}). \qquad (4)$$

The uncertainty of the $H(\boldsymbol{p})$ distribution is transferred, through model $g_L(\cdot)$ to the relevant uncertainty of the $L(\boldsymbol{p})$ distribution.

3. NUMERICAL RESULTS

3.1. Arsenic Exposure Mapping

At Stage 1 of the holistochastic analysis we are interested in assessing the *As* exposure on the popu-

lation through drinking water. We consider the BGS data set of *As* concentration measurements at shallow aquifers described above. The data were collected within a two-year interval, which was too small to detect significant temporal trends. Therefore, we assumed that the *As* concentration at any given spatial location is not varying with time. This is a reasonable assumption since naturally occurring *As*—which is the case of Bangladesh—does not display considerable temporal fluctuations at time frames of such small magnitude. Hence, herein we focus on a purely spatial analysis where the point vector $\boldsymbol{p}$ of Equations (1)–(4) now corresponds to the spatial location (time is fixed).

According to the BME method, the *As* estimates at unsampled locations are obtained from the available KB following an exploratory analysis. The spatial estimation process operates on de-trended data; therefore, exploratory analysis removes any data trends and identifies the underlying correlation structure by means of covariance models. The spatial trends are restored after the estimation process has been completed. In particular, the available data are scanned for trend determination, the trend is removed, and the de-trended *As* data roughly follow a log-normal distribution (Fig. 3), that is, the logarithmic values of the data are normally distributed. We choose to work with the normally distributed logarithmic values, keeping in mind that the estimates will eventually need to be back-transformed to the normal space values. The next step in the exploratory analysis is the investigation of the systematic dependencies in
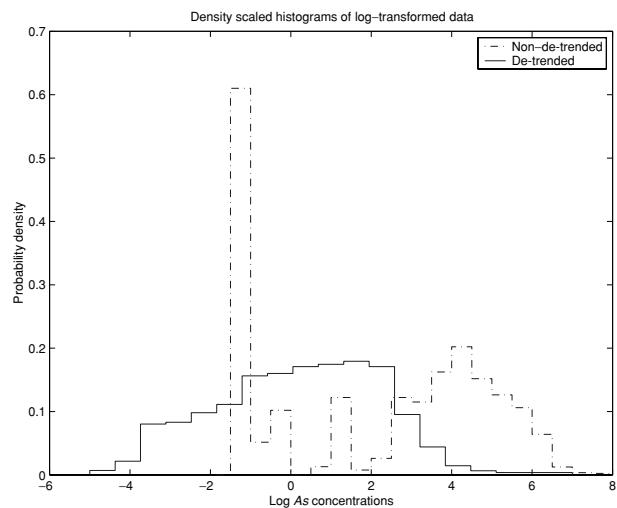

Fig. 3. Density scaled histograms of log-transformed data.

the *As* data. A physically and statistically acceptable covariance model is sought to describe the correlation among the data. A theoretical covariance model suitable for spatial *As* fields involves a nested structure of two exponential and one gaussian components. This model is fitted to experimentally derived covariance values for *As* log-concentrations, as shown in Fig. 4. The covariance model is a function of the distance r between any two spatial points as follows:

$$c_x(r) = c_1 \exp\left\{-3\frac{r}{a_1}\right\} + c_2 \exp\left\{-3\frac{r}{a_2}\right\}$$
$$+ c_3 \exp\left\{-3\left(\frac{r}{a_3}\right)^2\right\}, \tag{5}$$

where the sill values are $c_1 = 3.194$, $c_2 = 0.29$, and $c_3 = 0.63$, and the corresponding range values are $a_1 = 2$ km, $a_2 = 35$ km, and $a_3 = 57$ km. Equation (5) accounts for the intense spatial variability of *As* in Bangladesh as has been reported.[13] Moreover, Equation (5) was tested for anisotropy by calculating covariances along the north-south and east-west directions, in addition to the all-directional model of Fig. 4. The results suggest a fairly isotropic spatial distribution of the *As* log-values so that Equation (5) is an adequate model. The means and covariance models constitute the *G*-KB of the Bangladesh *As* situation.

The available *As* data set includes mostly hard data values (about two-thirds of the total number of measurements) that have been analyzed at the BGS laboratories. On the other hand, about one-third of the values (1,240 out of 3,373) are uncer-
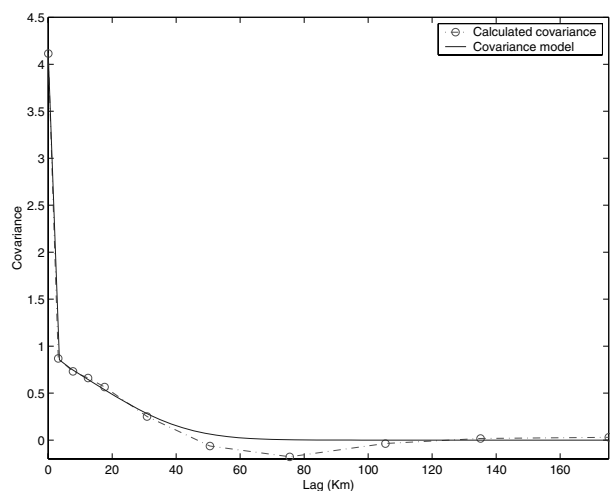
tain data that are either below some value (reported as being <6 μg/L or <30 μg/L) or below the measuring limit of the sampling device. These data are considered as soft interval data, where the interval values range accordingly from 0 to the measuring limit or the reported threshold. In BME terminology, the hard and soft data values together constitute the *S*-KB of the current study. At the BME integration (or posterior) stage, the structural PDF f_g are conditioned on the *S*-KB above to yield the integration PDF f_k at a dense set of mapping locations throughout Bangladesh. The BME calculations were performed using the *BMElib* software library,[10] which generates f_k throughout Bangladesh. As was mentioned in a previous section, from these PDF a variety of spatial *As* maps can be derived representing, for example, estimation means (*BMEmean*), medians (*BMEmedians*), or modes (*BMEmodes*). Fig. 5 presents a spatial map of *As* concentrations based on the *BMEmedian* estimates, after the logarithmic values have been back-transformed to the normal space and the spatial trend has been restored. The figure shows iso-concentration contours, that is, contour lines along which the *As* concentrations remain constant. The contours display concentrations in increments of 100 $\mu g/L$. For illustration, the 10 $\mu g/L$ and 50 $\mu g/L$ contours have been additionally plotted, which correspond to the WHO and Bangladesh maximum

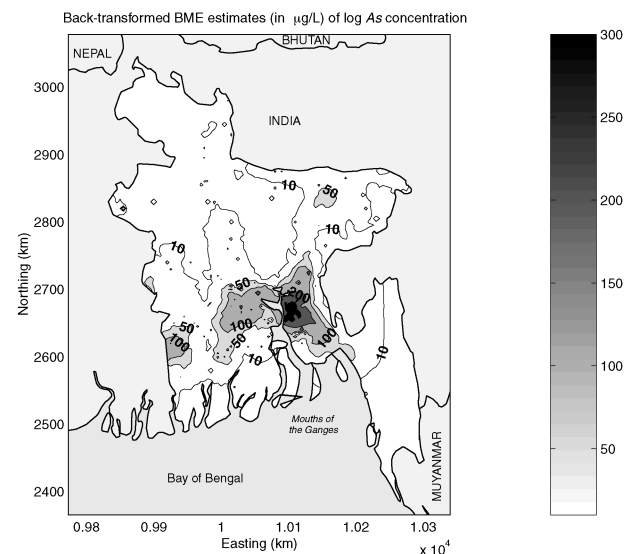**Fig. 5.** Map of the *BMEmedian* estimates of the *As* concentrations (in μg/L). The concentration contours of 10 μg/L (World Health Organization standard) and 50 μg/L (Bangladesh standard) are shown along with contours in increments of 100 μg/L.

Fig. 4. Modeled (solid line) and experimental covariances (circles connected by dash-dotted line) of the *As* log-concentration field.
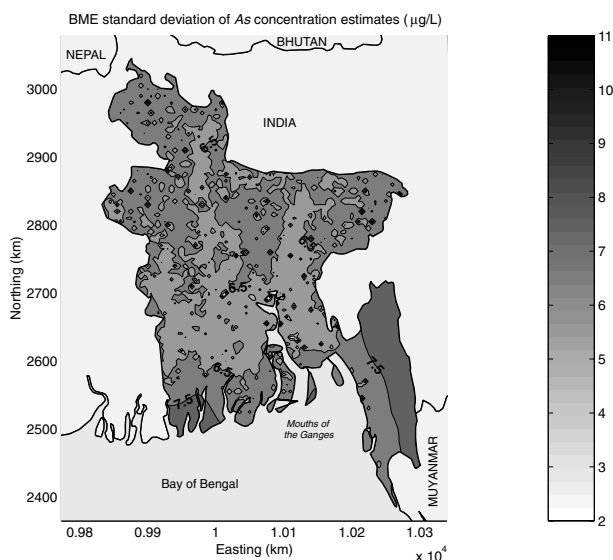
Fig. 6. Map of BME standard deviation of the *As* concentrations (in μg/L). At each location the displayed value is the standard deviation from the corresponding *BMEmedian* estimate shown in Fig. 5. Some darker spots that appear suggest the fast increase of uncertainty from lower values to the level of the surrounding values at clustered data locations.

allowed *As* concentration standards, respectively. Note that according to the estimates mapped in Fig. 5, there exist areas in Bangladesh where the *As* concentration is in the neighborhood of 300 μg/L. In fact, the maximum concentration value estimated for a single location in Bangladesh tops at 508 μg/L. The f_k also provide measures of the uncertainty of these estimates. Such an uncertainty measure is the estimation standard deviation at each location, which is mapped in Fig. 6. A different uncertainty measure will be discussed later in Section 3.2.2. These values provide a range within which the estimates may vary at a particular location. The standard deviation is smaller closer to data sampling locations and increases as one moves further away from the measured samples. Overall, the map in Fig. 6 provides a stochastic description of the uncertainty associated with *As* concentrations in shallow wells due to the high spatial variability of *As* and the detection limits of measurement instruments.

3.2. Health Risk Mapping of Bladder Cancer

In Stage 2 of the holistochastic human exposure analysis we use a model-based approach to assess the health risk resulting from chronic exposure to *As* in the drinking water. We demonstrate the holistochastic framework in terms of bladder cancer health risks,

since there is sufficient evidence from studies in various countries that *As* is a possible causal factor of bladder cancer.[14] Although it is generally known that this particular health outcome tends to occur among the rather older ages of population, we make use of the bladder cancer health effect for the following reason. The proposed model-based approach makes use of direct exposure-health response relationships, which can be constructed on the basis of one's experience with related epidemiologic or toxicologic studies. In the present case, an exposure-health response curve on bladder cancer was directly available allowing an instructive demonstration of the function and strengths of the holistochastic approach. This demonstration can serve as a guide for performing similar analyses in terms of other health outcomes.

Once ingested, inorganic *As* is rapidly absorbed by the organism and distributed to human organs. Next, a methylation process follows and approximately 5–25% of inorganic *As* is excreted unmethylated. This unmethylated portion, which passes through the urinary tract and is briefly stored in the bladder, is believed to be responsible for the increased rates of bladder cancer seen in exposed populations.[15] Currently, the main information on *As* health models are epidemiologic and toxicologic exposure-response studies. Interaction between scientific disciplines can provide modelers with:

1. Purely statistical regression models.
2. Mechanistic carcinogenesis models describing the potential *As* interference with background DNA damage.

The models of Group 1 are typically built using the lifetime probability data of epidemiologic studies for cancers caused by *As* exposure. In the following we examine one model from each category, namely, a linear exposure-response model derived from statistical regression of epidemiologic data, and a nonlinear one that is based on a more elaborate approach dealing with the mechanisms that cause cancer.

3.2.1. A Linear Exposure-Response Model

We first consider a linear regression model calibrated to the epidemiologic data proposed by Morales *et al.*[16] These experimental data are shown in Fig. 7 and suggest no clear trend in the dose-response behavior. It is common that such measurements are fitted using a linear model, according to which the lifetime probability *P*[*cancer*] of developing bladder cancer is
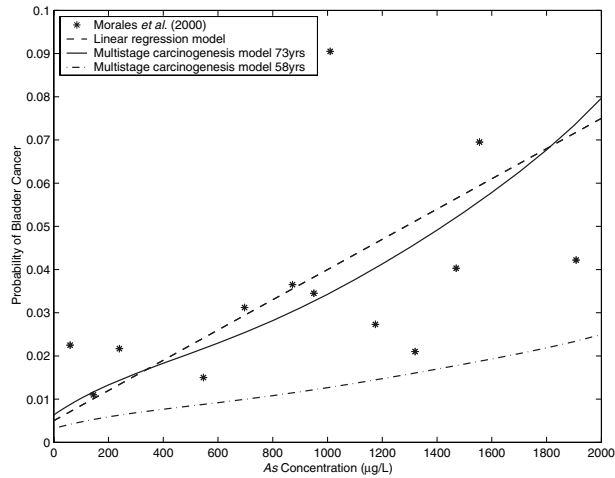
Fig. 7. Comparison of two models for the health effect $H(p)$ (probability of lifetime bladder cancer) vs. *As* concentrations. The stars are epidemiologic data from Morales *et al.*[16] The dashed line refers to the linear exposure-response model (see Equation (6)). The other two lines refer to the multistage carcinogenesis model with variable repair rate, see Equation (A1); specifically, the solid line is the outcome of the model for an expected lifetime of 78 years (corresponding to the Taiwanese population), and the dash-dotted line shows the model response for a lifetime expectancy of 58 years, as assumed for Bangladesh.

related to the *As* concentration $X(p)$ as follows:

$$P[cancer; p] = P_B + k X(p), \qquad (6)$$

where $X(p)$ expresses the *As* concentration (in μg/L) in the drinking water that people are exposed to, and k is a linear slope constant. Note that in this particular model there is an underlying background bladder cancer probability of $P_B = 0.005$ not depending on *As* as a causing factor. The k value obtained by fitting Equation (6) to the Morales *et al.*[16] data is $k = 3.5 \times 10^{-5} [\mu\text{g/L}]^{-1}$. The epidemiologic data in Morales *et al.* and the corresponding linear regression model are shown in Fig. 7 as stars and a dotted line, respectively. The health effect $H(p)$ is defined as the probability in Equation (6), that is:

$$H(p) \equiv P[cancer; p]. \qquad (7)$$

We can further write:

$$H(p) = H_B + H_{As}(p), \qquad (8)$$

where $H_B = P_B$ and $H_{As} = kX(p)$. This is a flexible approach that can consider any other health effect as well (skin cancer, etc.). Equation (8) distinguishes between the health effect H_B due to the existing background bladder cancer and the increased bladder cancer effect $H_{As}(p)$ purely due to *As* exposure.
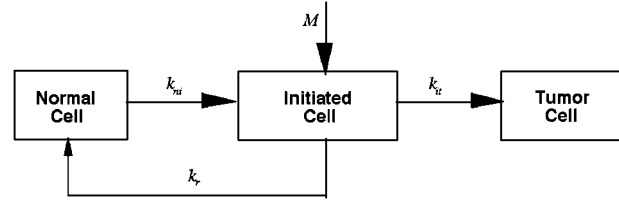

Fig. 8. The Moolgavkar model (modified by Crawford-Brown and Hoffman). Experimental parameters include the cell transfer rates k_{ni} (normal-to-initiated), k_{it} (initiated-to-tumor), k_r (repair), and the mitotic rate M at which initiated cells divide.

As defined above, $H(p)$ is a probability expressing the rate of the population expected to suffer a health effect based on *As* exposure to contaminated drinking water.

3.2.2. A Nonlinear Exposure-Response Model

As an example from Group 2 we employed a multistage carcinogenesis model from Moolgavkar *et al.*[17] and modified by Crawford-Brown and Hoffman.[18] A conceptual representation of this mechanistic model is given in Fig. 8 (boxes represent the three assumed states of cells, and arrows represent the transitions between these states). The model expresses the probability that a normal cell will turn into a tumor cell during an individual's lifetime. Experimental parameters describe the normal-to-initiated cell transfer rate (k_{ni}), the initiated-to-tumor cell transfer rate (k_{it}), and the mitotic rate (M) at which initiated cells divide. In addition, the repair transfer rate (k_r) is a decreasing function of *As* concentration representing the *As* interference with DNA repair mechanisms. Using these parameters, a set of equations is derived describing the growth of cells over the lifetime of an individual (see Appendix A). This set of equations can be solved for some initial number of normal cells in the infant, assuming that the lifetime probability P of bladder cancer is equal to the number of tumor cells generated for each initial normal cell at a given *As* concentration during an individual's lifetime. In this manner, the model effectively establishes a composite relationship between *As* concentrations and the lifetime probability of cancer caused by *As*. Just as for the linear Equation (7), the health effect $H(p)$ is defined as the resulting probability P of the multistage model, where the model parameters are adjusted to fit given epidemiologic data. In Appendix A we provide the parameters k_{ni}, k_{it}, M, and k_r obtained to fit the Morales *et al.*[16] bladder cancer data for Taiwan with a male life expectancy of 73 years.

The corresponding exposure-response curve is shown (plain line) in Fig. 7 for the Taiwan population. Note that if we use the same parameters k_{ni}, k_{it}, M, and k but decrease the life span to 58 years to account for the shorter life expectancy in Bangladesh (both male and female life expectancy; data currency: July 2002, see http://www.state.gov), we obtain for the Bangladesh population an exposure-response curve with lower lifetime probability of bladder cancer, as shown (dash-dotted line) again in Fig. 7. As exhibited in Fig. 7, the multistage model provides a carcinogenesis description of As on the natural DNA repair mechanism that results in a nonlinear As-bladder cancer relationship. Indeed, some scientists argue that the dose-response relationship between ingested As and cancer may not be linear and a threshold or sublinear response may exist.[19−21] The multistage carcinogenesis health effect $H(\boldsymbol{p})$ curve starts at zero As concentration with a small background value slightly over 0.006 μg/L for the Taiwanese population (Fig. 7, plain line), similar to that of the linear regression model (dashed line), and an even lower background value of about 0.003 μg/L for the Bangladesh population (dash-dotted line). At low As concentrations X (at about 0–400 μg/L) the $H(\boldsymbol{p})$ increases in a sublinear manner, that is, the curve moves below the tangent at any of its points in that range of X values; then, for X greater than about 400 μg/L, $H(\boldsymbol{p})$ increases supra-linearly, that is, the tangent at any of the curve's points lies below the curve.

Each of the models presented above provides an exposure-response relationship that can be used to predict the health risk corresponding to the BME map of As exposure in Bangladesh (Fig. 5). For illustration, Fig. 9 shows a map of the health effect $H(\boldsymbol{p})$ based on the linear model of Equation (6). While Fig. 5 refers to As exposure at shallow wells, Fig. 9 translates these exposure values (via the exposure-response model) into the associated health effects in terms of lifetime probability of bladder cancer incidence.

In the exposure assessment Stage 1 we examined the estimated As exposure errors derived from the BME probability model f_k. In the current stage we combine this source of uncertainty with that of the exposure-response models above (linear or nonlinear) to obtain uncertainty information regarding the health effect distribution across space. The PDF is a valuable tool for studying uncertainty as it allows us to study this very important issue of human exposure in several ways. For example, in Stage 1 we used the standard deviation as an exposure uncertainty measure, which is a common measure in statistical calcula-
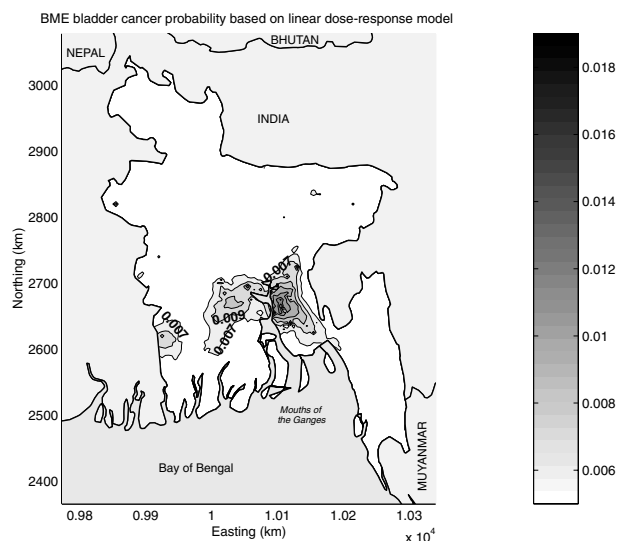

Fig. 9. Spatial map of the health effect $H(\boldsymbol{p})$ (expressed in lifetime bladder cancer probability) over Bangladesh based on the linear exposure-response model.

tions. The results of the present Stage 2, however, may be of interest to scientists from different disciplines, for example, epidemiologists or policymakers. A typical uncertainty measure used in these scientific areas is the confidence interval (CI). Again, we can use the integration PDF to derive, for example, the 68% CI of the As concentration estimates X at each location $\boldsymbol{p}$ of the mapping grid. These CI provide the size of the intervals within which X values are expected to vary with 68% probability. Given Equation (6), the combined uncertainty is generally the convolution of the uncertainties in X and k. Nevertheless, for simplicity in this example we did not assume any uncertainty in k and the background cancer probability parameters. In view of Equation (6), the CI bounds of As concentrations are processed in a straightforward manner to obtain the overall uncertainty for the health effect $H(\boldsymbol{p})$. The map in Fig. 10 displays the 68% CI of the $H(\boldsymbol{p})$ values across space, that is, under the above assumptions there is a probability of 68% that the $H(\boldsymbol{p})$ at each mapping location $\boldsymbol{p}$ will vary locally within intervals as wide as those shown in the map.

3.3. Population Health Impact of Bladder Cancer Due to Arsenic

In Stage 3 of holistochastic analysis we focus on the assessment of the adverse effects at the population level, which is of primary interest to public health management and administration. To this end, a
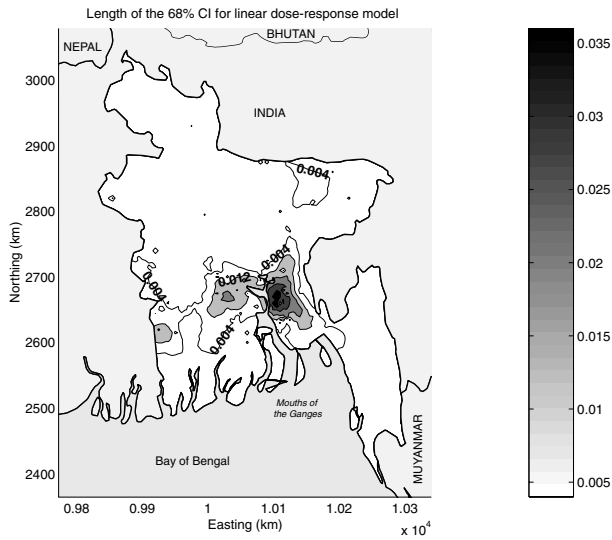
Fig. 10. Integrated uncertainty in holistochastic processing of Bangladesh *As* health risk designated by the 68% confidence interval of the health effect $H(\boldsymbol{p})$. The map shows that within a probability of 68%, the health effect $H(\boldsymbol{p})$ will vary within a range of values of the size shown at location $\boldsymbol{p}$.

variety of indicators—simple and sophisticated—have been proposed.[7] A simple indicator is, for example, the number of cancer cases observed per 100,000 people; another one is the number of people per km^2 expected to develop the health effect in their lifetime. In the following, the lifetime cancer probability obtained earlier is combined with population data in order to produce a useful measure of the population health effect. The indicator used is the number of people per km^2 expected to develop bladder cancer in their lifetime.

The lifetime probability $H(\boldsymbol{p})$ of bladder cancer is calculated from Equation (8); demographics information for Bangladesh was provided by the Center for International Earth Science Information Network (CIESIN). The Bangladesh population density distribution $D(\boldsymbol{p})$ data are displayed in Fig. 11. The results of the holistochastic analysis based on the earlier BME results as well as the input from health experts are shown in Fig. 12. In particular, Fig. 12 depicts a spatial map of the distribution of the population damage $L(\boldsymbol{p})$ representing the number of people per km^2 with a lifetime bladder cancer development expectancy. Under the conservative assumption that all the drinking water comes from shallow wells, the L-value at each point of the map of Fig. 12 is simply calculated as follows:

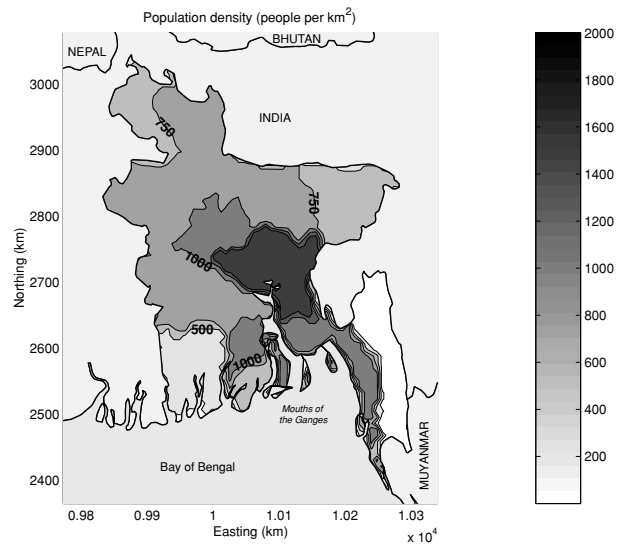$$L(\boldsymbol{p}) = D(\boldsymbol{p})H(\boldsymbol{p}). \qquad (9)$$

Fig. 11. Population density map for Bangladesh.

Note that as water supply shifts to deep wells, Equation (9) can easily be modified to account only for the fraction of the population getting its water from shallow wells. The map identifies areas in Bangladesh where it was estimated that as many as 45 lifetime cancer incidents are expected to occur per km^2 (roughly corresponding to three incidents every four years for a life span of 58 years). Such maps are very useful for the public health planner and policymaker
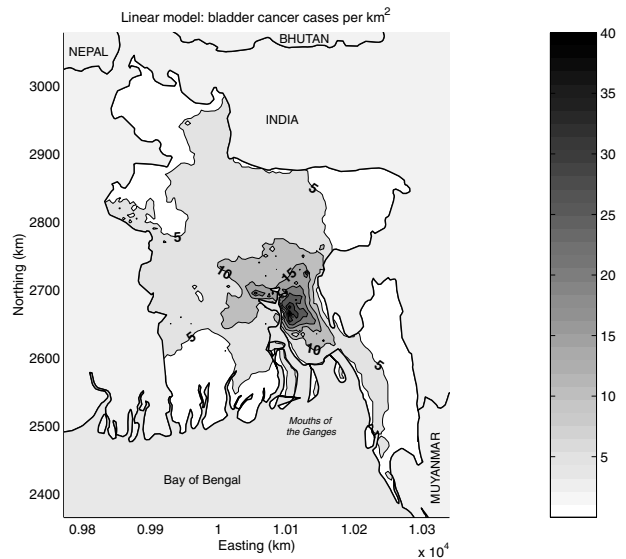

Fig. 12. Map of the population damage $L(\boldsymbol{p})$ (in people per km^2) for the linear exposure-response model.

because they allow, for example, estimation of the cost of treatment of population damages due to a particular health effect. This analysis offers an additional level of information by combining Equations (8) and (9). A more detailed expression for the population damage emerges then as follows:

$$L(\boldsymbol{p}) = L_B + L_{As}(\boldsymbol{p}), \qquad (10)$$

where

$$L_B(\boldsymbol{p}) = D(\boldsymbol{p})H_B \qquad (11)$$

and

$$L_{As}(\boldsymbol{p}) = D(\boldsymbol{p})H_{As}(\boldsymbol{p}), \qquad (12)$$

that is, we can distinguish between the portion $L_B(\boldsymbol{p})$ of the population damage due to background bladder cancer and the damage $L_{As}(\boldsymbol{p})$ exclusively due to the health effect caused by the presence of As in drinking water.

3.4. Total Population Health Impact

Accuracy in health effect prediction is crucial for subsequent stages of health planning and decision making. A meaningful analysis should, therefore, incorporate the uncertainty involved at all different stages of the predictions. In that manner, the results for the health planner should incorporate an accumulated variation that will reflect correctly the accruing uncertainties at the different processing levels. Often, this accumulated variation shows in the form of best-case and worst-case scenarios that, respectively, translate, for example, into the predicted lower and upper number of cancer cases. Best-case and worst-case scenarios are important indicators, based upon which health policies, regulations, and financial plans can be structured. Consequently, it is vital to derive these indicators using as many information resources as possible and methodically monitor their uncertainty throughout the process—both of which tasks are skillfully and scientifically performed by BME and the proposed holistochastic analysis.

In the preceding sections, we focused on the uncertainty sources and included the uncertainty analysis with the results. Unlike other currently used mapping methods, the BME approach allowed us to account for uncertainties at the very basic level of data inclusion. Its rigorous theoretical foundations make possible a full-scale uncertainty analysis at the end of the mapping process—a result of which is the integration PDF model and the estimation standard deviation map shown in Fig. 6. Then, given a particular

dose-response model (the linear model of Equation (6)), we defined the health effect $H(\boldsymbol{p})$ in Equation (8) and obtained the integrated uncertainties in Fig. 10 by using the 68% CI as an uncertainty measure. This brings us in a position to provide the uncertainty in the population health damage $L(\boldsymbol{p})$ level, as well, given its association with $H(\boldsymbol{p})$ through Equation (9). The 68% CI of the estimated values of $L(\boldsymbol{p})$ provides us with a probability of 68% that the number of people developing the health effect will range within the CI upper and lower bounds calculated at each location $\boldsymbol{p}$. In the remaining, we discuss possible ways in which this kind of information can be useful to health planners.

A very useful risk assessment indicator is the population global damage indicator Λ, defined as:[7]

$$\Lambda = \int_V L(\boldsymbol{p}) \, dV. \qquad (13)$$

According to Equation (13), Λ provides the total population damage (in number of affected people) over an area V where the population damage is distributed as $L(\boldsymbol{p})$ people per km^2. In our study, Λ refers to the total number of lifetime bladder cancer incidents expected for the Bangladesh population. By integrating the L-values of Fig. 12 over the given population density at each one of the estimation locations $\boldsymbol{p}$ we obtain Λ for all of Bangladesh. The uncertainty analysis for Λ is derived in this case directly from the uncertainty in L. Specifically, we use the 68% CI bounds of L to obtain the upper and lower bounds for Λ. In other words, we use the lower and upper values of population damage L that is expected with a probability of 68%, and Equation (13) provides us, respectively, with the best- and worst-case scenarios for the particular CI. Using the linear Equation (6), we estimated $\Lambda = 801,878$ cases to be lying within the bounds [698,795; 1,158,881] of bladder cancer cases.

Instead, if the background cancer probability is excluded from the calculations, then we can use Equation (13) to get the population global damage indicator:

$$\Lambda_{As} = \int_V L_{As}(\boldsymbol{p}) \, dV. \qquad (14)$$

In view of the previous definition, Λ_{As} is the total population damage (number of Bangladesh lifetime bladder cancer incidents, in our case) purely due to As over an area V where the population damage exclusively due to As is given by $L_{As}(\boldsymbol{p})$. Again, using the linear Equation (6) and the 68% CI bounds of L,

the net effect due to *As* is estimated as $\Lambda_{As} = 145{,}529$ cases lying within the bounds [42,445; 502,532].

It is noteworthy that the population damage estimates in the present study were obtained using the *As* dose-response data for bladder cancer incidence in Taiwan reported by Morales *et al.*[16] Based on the discussion by Christakos and Kolovos,[22] accurate estimation in the case of Bangladesh requires input from yet unavailable local dose-response models to account for the particular characteristics of the Bangladesh population. In that sense, the previously reported estimates regarding the linear model outcome are based on the conservative scenario that Bangladesh exhibits the same bladder cancer incidence as Taiwan. The estimates for that model can change considerably using the life expectancy adjustment for Bangladesh that was applied to the nonlinear model. Indeed, using the multistage carcinogenesis model (Equation (A1)) properly adjusted to the Bangladesh life expectancy of 58 years (dash-dotted line in Fig. 7), we find the total lifetime bladder cancer incidents to amount $\Lambda = 482{,}687$ cases (compared to an expected $\Lambda = 988{,}659$ cases for the Taiwanese male life expectancy of 73 years using the solid-line model in Fig. 7), whereas the net effect due to *As* is now estimated to $\Lambda_{As} = 61{,}897$ cases (down from $\Lambda_{As} = 157{,}842$ for the Taiwanese population).

As there are currently related epidemiologic studies for Bangladesh in the works, in the previous we have accommodated the illustration of BME and the holistochastic approach using available information for bladder cancer as an *As* health effect example, and some conservative estimates for the framework application. The scientist who wishes to make use of the holistochastic approach needs to apply suitable information and use proper KB for specialized case studies, for example, dose-response models for target populations or target groups, localized contaminant data, population habits and trends, and so forth. It is plausible that due to the relatively short life expectancy in Bangladesh and the possibly long latency of bladder cancer (which makes it affect the older population), exposed people in Bangladesh may die of other causes before developing bladder cancer, thus resulting in an even smaller bladder cancer incidence than that reported in this work. The input from different disciplines that contribute to studies like this should allow for adjustments so that, eventually, useful information is provided. For example, the mechanistic carcinogenesis model presented in this work depends on adjustable parameters such as the life expectancy parameter to fit its output to experimental data like the ones reported by Morales *et al.*[16]

4. CONCLUSIONS

The well-known holistochastic framework of human exposure was used to study lifetime population damage due to *As* in Bangladesh drinking water. Within this framework, the BME mapping concept offers rigorous theoretical tools to assimilate various kinds of knowledge bases (physical, epidemiologic, toxicokinetic, etc.) and produce useful results for a broad spectrum of end-users. For illustration, we presented a case study where knowledge bases from various scientific fields were integrated, leading to highly informative maps of the *As* concentrations in Bangladesh drinking water, as well as maps of the related health effect (bladder cancer) on the population.

We used the BME method to account for uncertain information in a strict scientific manner. The rigorous assimilation of important knowledge bases (uncertain or secondary data sources) can complement the scientific reasoning of human exposure, whereas formerly input was limited to hard data only. Dealing with additional, albeit uncertain, information resources is not a limiting factor. Instead, these resources take over at the point where hard data cease providing assistance and further channel research in the correct direction. Subsequently, more accurate results are obtained in space-time within adequate confidence levels. This is a very desirable feature in areas such as public health planning and policy making, where every bit of useful information is crucial.

A human exposure framework has been portrayed, where interdisciplinary interaction becomes the epicenter for the creation of mapping products of key importance to health research, management, and decision making. In the presence of adequate information the combination of the powerful BME stochastic mapping approach and interdisciplinary input can prove to be substantially fruitful.

APPENDIX

Description of the Multistage Carcinogenesis Model with Variable Repair Rate

The multistage carcinogenesis model is described by the following set of differential equations:

$$\frac{dN_n(t)}{dt} = -k_{ni} N_n(t) + k_r N_i(t), \tag{A1a}$$

$$\frac{dN_i(t)}{dt} = k_{ni} N_n(t) - k_r N_i(t) - k_{it} N_i(t) + M N_i(t),$$
(A1b)

and

$$\frac{dN_t(t)}{dt} = k_{it} N_i(t),$$
(A1c)

where $N_n(t)$ is the number of normal cells at instance t; $N_i(t)$ is the number of initiated cells at instance t; $N_t(t)$ is the number of tumor cells at instance t; k_{ni} is the rate of transition from normal to initiated cells (in probability $\cdot[T]^{-1}$ units); k_{it} is the rate of transition from initiated to tumor cells (in probability $\cdot[T]^{-1}$ units); k_r is the repair rate constant from initiated to normal cells (in probability $\cdot[T]^{-1}$ units); and M is the net growth rate for the pool of initiated cells (in probability $\cdot[T]^{-1}$ units). M refers to mitosis (cell division); therefore, the total number of cells naturally increases with time.

First, the parameters k_{ni}, k_{it}, M, and k_r are set to fit the experimental data of Morales *et al.*[16] using a male life expectancy of 73 years corresponding to the population in Taiwan. We start with one healthy cell, that is, $N_n(0) = 1$ and $N_i(0) = N_t(0) = 0$. In that case, as stated in the text, the lifetime bladder cancer probability is defined as $P[cancer] = N_t$ ($t = 73$ *years*). We used the following fitting constants: $k_{ni} = 0.00075$ *per day*, $k_{it} = 6 \cdot 10^{-7}$ *per day*, and the mitotic parameter background rate of $M = 0.002$ *per day*. We considered the repair rate k_r to be a function of the *As* concentration X, based on the following fitting equation:

$$k_r(C) = 0.015 - 7.10^{-4} X^{0.3},$$
(A2)

where X is expressed in $\mu g/L$. Starting at $t = 0$ *days* we advance using increments of $dt = 5$ *days* and solve the Equations (A1a), (A1b), and (A1c) for a total of 26645 *days* = 73 *years*. We found the results not to be significantly sensitive to perturbations in the increment size dt. We iterate repeating solutions of the differential equation system (A1) for a variety of concentrations in the range of 0–2,000 $\mu g/L$ (e.g., in increments of 100 $\mu g/L$) to obtain the $P[cancer]$ model values at these concentrations. A least squares fit of these model points results in a polynomial function that describes the dose-response process based on the multistage carcinogenesis model. This function is represented by the solid line in Fig. 7 that reasonably fits the experimental dose-response data by Morales *et al.*[16]

We subsequently adjusted the carcinogenesis model for the Bangladesh life expectancy of 58 years.

Again, we solved the Equations (A1a–c) using the same increment of $dt = 5$ *days*, and parameters k_{ni}, k_{it}, M, and k_r as above. In this case, however, we advanced in time from $t = 0$ *days* to a total of only 21,170 *days* = 58 *years* so as to account for the lower life expectancy in Bangladesh (compared to Taiwan). The multistage carcinogenesis dose-response curve obtained, which applies to the Bangladesh population, is shown as the dash-dotted curve in Fig. 7.

ACKNOWLEDGMENTS

REFERENCES

1. Augustinraj R., Serre, M. L., & Christakos, G. (2001). Modelling the spatiotemporal distribution of subsurface heavy metal contamination at the Cherry Point, North Carolina Superfund site. Presented at the Superfund Annual Research Symposium. Chapel Hill, NC.
2. British Geological Survey (BSG). (2000). *Groundwater studies of arsenic contamination in Bangladesh* (Final Report Summary). Available at http://www.bgs.ac.uk/arsenic/bangladesh/reports.htm.
3. National Research Council. (1999). *Arsenic in drinking water*. Washington, DC: National Academy Press.
4. Chen, C.-J., Chen, C. W., Wu, M.-M., & Kuo, T.-L. (1992). Cancer potential in liver, lung, bladder and kidney due to ingested inorganic arsenic in drinking water. *British Journal of Cancer*, *66*, 888–892.
5. Tsai, S.-M., Wang, T.-N., & Ko, Y.-C. (1999). Mortality for certain diseases in areas with high levels of arsenic in drinking water. *Archives of Environmental Health*, *54*(3), 186–193.
6. Christakos, G. (2000). *Modern spatiotemporal statistics*. New York: Oxford University Press.
7. Christakos, G., & Hristopulos, D. T. (1998). *Spatiotemporal environmental health modelling: A tractatus stochasticus*. Boston, MA: Kluwer Academic Publishers.
8. Chowdhury, U. K., Biswas, B. K., Dhar, R. K., Samanta, G., Mandal, B. K., Chowdhury, T. R., Chakraborti, D., Kabir, S., & Roy, S. (1999). Groundwater arsenic ccontamination and suffering people in Bangladesh. In W. R. Chappell, C. O. Abernathy, & R. L. Calderon (Eds.), *Arsenic exposure and health effects* (pp. 165–182). Amsterdam: Elsevier.
9. Rahman, M., & Axelson, O. (2001). Arsenic ingestion and health effects in Bangladesh: Epidemiological observations. In W. R. Chappell, C. O. Abernathy, & R. L. Calderon (Eds.), *Arsenic exposure and health effects* (pp. 193–200). Amsterdam: Elsevier.
10. Christakos, G., Bogaert, P., & Serre, M. L. (2002). *Temporal GIS* (CD-Rom included). New York: Springer-Verlag.
11. Christakos, G., & Serre, M. L. (2000). BME analysis of spatiotemporal particulate matter distributions in North Carolina. *Atmospheric Environment*, *34*, 3393–3406.

12. National Research Council. (2001). *Arsenic in drinking water*. Washington, DC: National Academy Press.

13. Smedley P. L., Kinniburgh, D. G., Huq, I., Zhen-dong, L., & Nicolli, H. B. (2001). International perspective on naturally occurring arsenic problems in groundwater. In W. R. Chappell, C. O. Abernathy, & R. L. Calderon (Eds.), *Arsenic exposure and health effects* (pp. 9–25). Amsterdam: Elsevier.

14. Smith A. H., Biggs, M. L., Moore, L., Haque, R., Steinmaus, C., Chung, J., Hernandez, A., & Lopipero, P. (1999). Cancer risks from arsenic in drinking water: Implications for drinking water standards. In W. R. Chappell, C. O. Abernathy, & R. L. Calderon (Eds.), *Arsenic exposure and health effects* (pp. 191–199). Amsterdam: Elsevier.

15. Steinmaus C., Moore, L., Hopenhayn-Rich, C., Biggs, M. L., & Smith, A. H. (2000). Arsenic in drinking water and bladder cancer. *Cancer Investigation*, *18*(2), 174–182.

16. Morales K. H., Ryan, L., Kuo, T.-L., Wu, M.-M., & Chen, C.-J. (2000). Risk of internal cancers from arsenic in drinking water. *Environmental Health Perspectives*, *108*(7), 655–661.

17. Moolgavkar S., Cross, F., Luebeck, G., & Dagle, G. (1990). A two-mutation model for radon-induced lung tumors in rats. *Radiation Research*, *121*, 118.

18. Crawford-Brown, D., & Hofmann, W. (1996). The testing of radiobiological models of radon carcinogenesis needed for *in vitro* to *in vivo* extrapolations. *Environment International*, *22*, S985.

19. Carlson-Lynch, H., Beck, B. D., & Boardman, P. D. (1994). Arsenic risk assessment. *Environmental Health Perspectives*, *102*, 354–356.

20. Abernathy C. O., Chappell, W. R., Meek, W. E., Gibb, H., & Guo, H. R. (1996). Is ingested inorganic arsenic a "threshold" carcinogen? *Fundamental and Applied Toxicology*, *29*(2), 168–175.

21. Slayton T. M., Beck, B. D., Reynolds, K. A., Chapnick, S. D., Valberg, P. A., Yost, L. J., Schoof, R. A., Gauthier, T. D., & Jones, L. (1996). Issues in arsenic cancer risk assessment. *Environmental Health Sciences*, *1104*, 2–4.

22. Christakos G., & Kolovos, A. (1999). A study of the spatiotemporal health impacts of ozone exposure. *Journal of Exposure Analysis and Environmental Epidemiology*, *9*, 322–335.