HHS Public Access

Author manuscript

Ann Appl Stat. Author manuscript; available in PMC 2018 August 10.

Published in final edited form as:

Ann Appl Stat. 2017 March ; 11(1): 139–160. doi:10.1214/16-AOAS995.

STATIC AND ROVING SENSOR DATA FUSION FOR SPATIO-TEMPORAL HAZARD MAPPING WITH APPLICATION TO OCCUPATIONAL EXPOSURE ASSESSMENT[1]

Guilherme Ludwig[*],
Department of Statistics, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA, gvludwig@stat.wisc.edu

Tingjin Chu[†],
Center for Applied Statistics and Institute of Statistics and Big Data, Renmin University of China, Beijing 100872, China, tingjin_chu@outlook.com

Jun Zhu[*],
Department of Statistics and Department of Entomology, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA, jzhu@stat.wisc.edu

Haonan Wang[‡], and
Department of Statistics, Colorado State University, Fort Collins, Colorado 80523, USA, wanghn@stat.colostate.edu

Kirsten Koehler[§]
Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205, USA, kkoehle1@jhu.edu

[*]University of Wisconsin-Madison

[†]Renmin University of China

[‡]Colorado State University

[§]Johns Hopkins University

Abstract

Rapid technological advances have drastically improved the data collection capacity in occupational exposure assessment. However, advanced statistical methods for analyzing such data

[1]Supported in part by the CAPES Foundation, Brazil, Grant 5588–10–3 (Ludwig), the National Natural Science Foundation of China, Grant 11301536 (Chu), USGS CESU Award G16AC00344 (Zhu), NSF Grants DMS-1106975 and DMS-1521746 (Wang) and NIOSH Grant R01 OH010533 (Koehler).

SUPPLEMENTARY MATERIAL

Appendix: Tuning parameter selection and simulation study
(DOI: 10.1214/16-AOAS995SUPPA;.pdf). The Appendix contains a description of the leave-one-sensor-out cross-validation procedure for MSPE evaluation and tuning parameter selection, a detailed approach for the choice of tuning parameters for the smoother terms and number of components for the data analysis in Section 5, and a simulation study comparing the static and roving sensor data fusion for the spatio-temporal mapping (STDF) method to fixed-time universal kriging, thin-plate spline smoothing and least squares regression.

Animation: Animated versions of Figures 1–2 and Figures 4–5 (DOI: 10.1214/16-AOAS995SUPPB;.zip). This supplemental material contains animated versions, dynamic in time, for the indicated figures.

and drawing proper inference remain limited. The objectives of this paper are (1) to provide new spatio-temporal methodology that combines data from both roving and static sensors for data processing and hazard mapping across space and over time in an indoor environment, and (2) to compare the new method with the current industry practice, demonstrating the distinct advantages of the new method and the impact on occupational hazard assessment and future policy making in environmental health as well as occupational health. A novel spatio-temporal model with a continuous index in both space and time is proposed, and a profile likelihood-based model fitting procedure is developed that allows fusion of the two types of data. To account for potential differences between the static and roving sensors, we extend the model to have nonhomogenous measurement error variances. Our methodology is applied to a case study conducted in an engine test facility, and dynamic hazard maps are drawn to show features in the data that would have been missed by existing approaches, but are captured by the new method.

Keywords

Geostatistics; kriging; semiparametric methods; spatial statistics; spatiotemporal statistics

1. Introduction.

Occupational exposure assessment refers to assessment of the level of contaminants an employee is exposed to during their work shift. The traditional method for occupational exposure assessment is personal monitoring using lightweight devices that can be worn by the workers. Personal exposure estimates are typically sought because they can be compared against regulatory standards to ensure compliance with existing laws. However, personal monitoring is generally expensive and requires workers to carry equipment with them during their work. As such, it is common for a small number of measurements, on asmall number of employees, to be collected [Tornero-Velez et al. (1997), Cherrie (2003)], resulting in small sample sizes that cannot accurately capture true levels of contamination. Additionally, without the ability to track worker location, there is little ability to apportion exposures to different areas or tasks.

Occupational hazard maps, contour plots of contaminant concentration over the two-dimensional floor plan of the workplace, have gained popularity as a method to overcome some of the limitations of the traditional personal sampling that is generally expensive with small sample sizes [Koehler and Peters (2013), Peters et al. (2006, 2012), Evans et al. (2008), Ologe, Akande and Olajide (2006)]. Hazard maps are commonly produced by industrial hygienists or researchers using direct-reading instruments (DRIs) to capture contaminant concentrations at high spatial resolution following a predetermined grid throughout the facility of interest (hereafter, roving sensors). Such maps are powerful tools to communicate risk in an easily understood format and to guide decisions on control strategies aimed at reducing worker exposures [O’Brien (2003)].

Hazard maps that rely on roving monitor data alone, while cost-effective to produce and conceptually simple, likely fail to represent the temporal variability in concentrations present in many occupational settings [Koehler and Volckens (2011), Lake et al. (2015)].

Augmenting the data with static sensors that collect time series data, but at a few locations, can allow practitioners to expand the temporal and spatial coverage of data collection [Lake et al. (2015)]. As DRIs become more affordable and accessible, these types of exposure data (from static and roving sensors with known spatial information) are expected to become more abundant, but rigorous statistical methods for analyzing data and drawing proper inference remain limited. The current hazard mapping approach to occupational exposure assessment, although novel, represents several challenges. Maps that are created from roving sensor data alone are often either collected over a short temporal interval or aggregated over time and neglect the temporal variability in the dataset. As such, temporal variability can be mistakenly displayed as spatial variability. In our previous work, we compared maps created using the roving sensor data and static sensor data separately [Lake et al. (2015)]. The method employed was somewhat ad hoc because a comprehensive statistical methodology was lacking to combine the datasets (static and roving) to provide a representation of exposures across space and time. The maps should give not only the most representative indication of the mean value, accounting for both data types, but also an indication of the variability in concentrations, as a function of both time and space.

Statistical methods for integrating different sources of data in space and/or time have been researched in the past. For example, Isaacson and Zimmerman (2000) developed methodology for combining environmental data that are temporally correlated and from two measurement systems. Their autoregressive moving-average models allowed a common time trend, system-specific measurement errors and missing data, for which the inference was conducted by both frequentist and Bayesian approaches. Cowles et al. (2002) extended Isaacson and Zimmerman (2000) to temporally correlated data from multiple measurement systems that are measured at distinct sites in space. A Bayesian approach was taken to estimate the long-term trend and evaluate differences among the measurement systems. Further, Smith and Cowles (2007) considered an integrated model for correlating point-referenced radon and areal uranium data for quantifying a common spatial process using also a Bayesian approach, whereas Sahu, Gelfand and Holland (2010) fused point-referenced and areal wet deposition data in space and time. Such prior research is illuminating, but none considered the possibility of roving sensors, and thus is not directly applicable for the hazard mapping under consideration. The objectives of this paper are (1) to develop new spatio-temporal methodology that combines data from both roving and static sensors for data processing and hazard mapping across space and over time in an indoor environment and (2) to compare the new method with the current industry practice, demonstrating the distinct advantages of the new method and the impact on occupational hazard assessment and future policy making in environmental health as well as occupational health.

Combining data from static and roving sensors in a statistically sound way is challenging. First of all, while the roving sensors expand the spatial coverage of data, the observations are sparse in time at any given location. This is in contrast to the static sensors that are at a smaller number of sampling locations, but observations are denser in time at each sampling location. An ad hoc approach would be to analyze the two types of data separately, but there is potential benefit to be gained by developing statistical methodology that pools the two data sources and takes full advantage of their respective strengths. In addition, inaccurate

and missing data can be a thorny issue in such data analysis due to different measurement systems, instrumentation failures, and uneven or asynchronous monitoring times, etc. To address these challenges, we propose a novel spatio-temporal process that has a continuous index in both space and time; that is, in the spatial domain of interest, the sampling locations can occur anywhere, in which sense the modeling is geostatistical [see, e.g., Cressie (1993)], whereas within the temporal window of interest, the sampling can occur at any time and thus the modeling may be viewed as functional [see, e.g., Ramsay and Silverman (2005)]. We then develop a model fitting procedure that allows the fusion of the two types of data based on profile likelihood accompanied by a fast computational algorithm. Further, to account for potential differences between the static and roving sensors, we extend the spatio-temporal model to allow for inhomogenous measurement error variances. Finally, we compare our new methodology with the current industrial standard/practice which does not model temporal variability and generates hazard maps from data averaged in time.

As we will demonstrate in a case study conducted in an engine test facility, the dynamic hazard maps that interpolate across space and over time are far more informative and representative of the evolution of hazard levels in space and time. This finding can impact the way occupational hazards are to be mapped in the future and move the industry and regulation forward to more accurate assessment of environmental hazards.

2. Application and data.

2.1. Case study.

A study was conducted in the spring of 2013 in an engine test facility located in Colorado to evaluate occupational exposure [Lake et al. (2015)]. The facility has two rooms, both rectangular in shape (14.8 m by 6.5 m and 14.8 m by 33.7 m, respectively), separated by a sliding door and encompassing a combined area of about 595 m^2. A floor plan is shown in Figure 1(a). In one experiment, for example, an active engine, located in the upper left corner of the facility [black square in Figure 1(a)], was operating between 10:00:00 am and 11:10:00 am, while the sliding door was open. Measurements of noise intensity were collected by 18 static sensors and 2 roving sensors. The locations of the static sensors are given in Figure 1(a), whereas the pathways of the roving sensors are shown in Figure 1(b). The static sensors started collecting data at 9:45:00 am and ended at 11:23:20 am when they were turned off. The operation of the first roving sensor started at 10:28:45 am and that of the second roving sensor started at 10:52:45 am. Both roving sensors were in operation until the end of the experiment, but not continuously. Static sensor measurements were collected at every minute, while roving sensors measured hazard levels with a resolution of 20 seconds. There are thus 100 sampled points for each of the 18 static sensors, 105 sampled points for the first roving sensor and 72 sampled points for the second roving sensor, for a total of 1977 observations.

The measurements sampled over time are plotted in Figure 2. For each static sensor, a time series of noise intensity is plotted. The static sensors near the active engine in the upper left corner of the facility had higher intensity (gray solid lines) than those further away (dashed lines). One static sensor (#18) was far away from the active engine but had high noise

intensity due to an unexpected noise source outside the facility (black solid line). The roving sensors are also displayed as filled circles (sensor 1) or open circles (sensor 2).

The static sensors all have dense sampling points in time, and thus relatively complete profiles of the temporal processes at the sampling locations (Figure 2). However, the spatial coverage by the static sensors is limited to the 18 sampling locations where they were installed [Figure 1(a)]. In contrast, each roving sensor has a wider spatial coverage [Figure 1(b)], but the information at any given location is sparse in time (Figure 2). The sampling time points for the roving sensors are also irregularly spaced, with occasional breaks that vary from about 20 seconds to about 5 minutes.

2.2. Current practice of hazard mapping.

In a recent review of hazard mapping approaches, Koehler and Peters (2013) noted that relatively simple methods have been used to construct maps, often averaging the sensors in time before employing a spatial interpolation technique, or interpolating across space for a fixed time and averaging the maps. We show static maps commonly obtained by industrial hygienists for the case study in Figure 3, after averaging over all data in time at each unique sampling location. The map labeled “Roving” is based on kriging estimates and uses the roving sensors exclusively. It corresponds to early practice with DRIs and still is the most common approach. The map labeled “Roving and Static” incorporates both sources of data. The aggregated map gives equal weights to observations available in the roving path (only one or two data points) and the ones from the static sensors (a hundred observations over the data collection period). Both maps tend to show local features that are not necessarily informative. In addition, the map labeled “Static” uses only static sensor data. It loses the local features provided by roving sensors due to a fairly limited spatial coverage. All these static maps naturally fail to capture the evolution of the hazard level in time and, as we will demonstrate, misrepresent the intensity of a secondary noise source in the southeastern part of the facility. Health effects of short duration but high-level exposures are unclear, and the static maps in current practice have limited capacity for studying these events.

3. Data and model specification.

3.1. Data specification.

We now specify the notation for the spatio-temporal process of a generic hazard. Let $D \subset \mathbb{R}^2$ denote a two-dimensional spatial domain of interest, and let $[0, T]$ denote a temporal window of interest where $T > 0$. Let $\mathbf{s}_1, \ldots, \mathbf{s}_{n_S}$ denote the locations of the static sensors, where $\mathbf{s}_i = (s_{i,x}, s_{i,y})'$ is the location of the ith static sensor and n_S is the number of static sensors. For the ith static sensor, let there be p_i sampling time points, denoted as $t_{k,i}$, for $k = 1, \ldots, p_i$, $i = 1, \ldots, n_S$. In contrast, let n_R denote the number of roving sensors. For the jth roving sensor, let there be q_j sampling time points. A roving sensor generally has different sampling locations at different sampling time points and, therefore, the sampling locations are denoted by $\mathbf{r}_{1,j}, \ldots, \mathbf{r}_{qj,j}$, where $\mathbf{r}_{l,j} = (r_{l,j,x}, r_{l,j,y})'$ is associated with sampling time $t_{l,j}$, for $l = 1, \ldots, q_j$, $j = 1, \ldots, n_R$. Let $y_{\mathbf{s}}(t)$ denote the intensity of a hazard at a given point in time t in $[0, T]$ and a given spatial location $\mathbf{s} = (s_x, s_y)' \in D$. We will denote the samples

collected by the static sensors as $y_{\mathbf{s}_i}(t_{k,i})$, for $k = 1, \ldots, p_i$, $i = 1, \ldots, n_S$, and the samples collected by the roving sensors as $y_{\mathbf{r}_{l,j}}(t_{l,j})$, for $l = 1, \ldots, q_j$, $j = 1, \ldots, n_R$.

3.2. Model specification.

To model the static sensor data $\{y_{\mathbf{s}_i}(t_{k,i})\}$ and roving sensor data $\{y_{\mathbf{r}_{l,j}}(t_{l,j})\}$, we consider a spatio-temporal model

$$y_{\mathbf{s}}(t) = \mu_{\mathbf{s}}(t) + \eta_{\mathbf{s}}(t) + \varepsilon_{\mathbf{s}}(t), \quad (1)$$

where $\mu_{\mathbf{s}}(t)$ is a deterministic mean function, $\eta_{\mathbf{s}}(t)$ is a random spatio-temporal process with $\mathbb{E}(\eta_{\mathbf{s}}(t)) = 0$ and covariance function $\gamma(t, \mathbf{s}, t', \mathbf{s}') = \mathrm{Cov}(\eta_{\mathbf{s}}(t), \eta_{\mathbf{s}'}(t'))$, and $\varepsilon_{\mathbf{s}}(t)$ is a measurement error process with $\mathbb{E}(\varepsilon_{\mathbf{s}}(t)) = 0$, constant variance $\sigma^2 = \mathrm{Var}(\varepsilon_{\mathbf{s}}(t))$ and zero correlation [Cressie and Wikle (2011)]. Further, we assume that the spatio-temporal process $\eta_{\mathbf{s}}(t)$ is square integrable and the spatio-temporal covariance function of $\eta_{\mathbf{s}}(t)$ at location $\mathbf{s} \in D$ satisfies

$$\gamma(t, \mathbf{s}, t', \mathbf{s}) = \gamma_0(t, t'), \quad (2)$$

where $\gamma_0(t, t') \gamma(t, \mathbf{s}_0, t', \mathbf{s}_0)$ is a temporal covariance function at any spatial location $\mathbf{s}_0 \in D$; that is, the temporal correlation function $\eta_{\mathbf{s}}(t)$ and $\eta_{\mathbf{s}}(t')$ is invariant in space.

The spatio-temporal process $\eta_{\mathbf{s}}(t)$ has a type of Karhunen–Loève decomposition [see, e.g., Gromenko and Kokoszka (2013)]:

$$\eta_{\mathbf{s}}(t) = \sum_{\ell=1}^{\infty} \xi_{\ell}(\mathbf{s}) \varphi_{\ell}(t), \quad (3)$$

where $\{\varphi_{\ell}(t)\}_{\ell=1}^{\infty}$ is a sequence of deterministic orthogonal temporal functions and $\{\xi_{\ell}(\mathbf{s})\}_{\ell=1}^{\infty}$ is a sequence of zero-mean random spatial processes that are uncorre lated with each other. The decomposition (3) represents the spatio-temporal process as a linear combination of the temporal basis functions $\varphi_{\ell}(t)$ (based on the temporal covariance function) with the random spatial processes $\xi_{\ell}(\mathbf{s})$ as coefficients.

We assume that the spatial covariance function of $\xi_{\ell}(\mathbf{s})$ takes on the form

$$\mathrm{Cov}(\xi_{\ell}(\mathbf{s}), \xi_{\ell}(\mathbf{s}')) = \lambda_{\ell} \rho_{\ell}(\|\mathbf{s} - \mathbf{s}'\|; \boldsymbol{\theta}_{\ell}),$$

where $\lambda_{\ell} = \mathrm{Var}(\xi_{\ell}(\mathbf{s}))$ is the variance of $\xi_{\ell}(\mathbf{s})$, $\rho_{\ell}(\cdot; \boldsymbol{\theta}_{\ell})$ is a correlation function parameterized by $\boldsymbol{\theta}_{\ell}$ and $\|\cdot\|$ denotes the Euclidean distance. From (3), we can write the spatio-temporal covariance function $\gamma(t, \mathbf{s}, t', \mathbf{s}')$ of $\eta_{\mathbf{s}}(t)$ as

$$\gamma(t, \mathbf{s}, t', \mathbf{s}') = \sum_{\ell=1}^{\infty} \mathrm{Cov}\left(\xi_\ell(\mathbf{s}), \xi_\ell(\mathbf{s}')\right)\varphi_\ell(t)\varphi_\ell(t'). \quad (4)$$

In the case $\mathbf{s} = \mathbf{s}'$, (4) is reduced to the temporal covariance function $\gamma_0 = (t, t') = \sum_{\ell=1}^{\infty} \lambda_\ell \varphi_\ell(t)\varphi_\ell(t')$ This makes clear that $\{\varphi_\ell(t)\}_{\ell=1}^{\infty}$ are analogous to the eigenfunctions of $\gamma_0(t, t')$ with the corresponding eigenvalues $\{\lambda_\ell\}_{\ell=1}^{\infty}$ It is based on (4) that we will devise a semiparametric likelihood approach to fitting the spatio-temporal model (1) to the static and roving sensor data, as well as mapping the true spatio-temporal process of the hazard $\mu_{\mathbf{s}}(t) + \eta_{\mathbf{s}}(t)$, while taking into account the spatial and temporal variability.

Our modeling approach is tailored toward the distinct features of static and roving sensor data. The spatial index is continuous in the spatial domain, and the temporal index is continuous within the time window. Thus, the sensors can be placed anywhere in the study area and do not need to be on a regular grid. Further, sampling can occur at any point in time and no regular time intervals are required. In addition, our modeling framework is semiparametric and flexible. The specification of the deterministic mean function $\mu_{\mathbf{s}}(t)$ is nonparametric, while the specification of the spatio-temporal process is semiparametric in the sense that the spatial covariance function $\mathrm{Cov}(\xi_\ell(\mathbf{s}), \xi_\ell(\mathbf{s}'))$ in (4) is parametric but the temporal covariance function $\gamma_0(t, t')$ in (2) is nonparametric. The nonparametric specification allows for capturing different sources for a hazard, some of which are unexpected to be present or are present at unexpected time intervals, such as the outside noise near sensor #18 in the case study.

The class of spatio-temporal covariance functions (4) is broad, encompassing processes that are nonstationary and nonseparable in space and time with the separable case corresponding to $\lambda_\ell = 0$ for $\ell \geq 2$ These properties can be contrasted to spatio-temporal kriging [Cressie and Wikle (2011), page 321], which in many practical scenarios requires the specification of the spatio-temporal covariance function, and the time series of the spatial process approach [Cressie and Wikle (2011), page 336], which requires the temporal coordinates to be sampled at regular intervals.

4. Statistical inference.

4.1. Profile likelihood estimation.

Parameter estimation by maximum likelihood can be challenging due the large number of parameters in the model (1). Thus, we use the idea of functional principal components and develop a profile likelihood approach to estimating the model parameters [Ramsay and Silverman (2005)]. Further, only finitely many eigenvalues are estimable from a sample covariance matrix in the nonparametric specification of the covariance function and, therefore, it is necessary that, for some $L \leq \min_i\{p_i\}$ [see Wahba (1990), page 5],

$$\eta_{\mathbf{s}}(t) \approx \sum_{\ell=1}^{L} \xi_{\ell}(\mathbf{s})\varphi_{\ell}(t). \quad (5)$$

Let $\mathbf{y}_{\mathbf{s}i} = (y_{\mathbf{s}i}(t_1), \ldots, y_{\mathbf{s}i}(t_{pi}))'$ and $\boldsymbol{\mu}_{\mathbf{s}i} = (\mu_{\mathbf{s}i}(t_1), \ldots, \mu_{\mathbf{s}i}(t_{pi}))'$ denote the vector of the data from the ith static sensor at time points t_k for $k = 1, \ldots, p_i$ and the corresponding mean vector for $i = 1, \ldots, n_S$. Let $\mathbf{y}_{\mathbf{r}_j} = \left(y_{\mathbf{r}_{1,j}}(t_1, j), \ldots, y_{\mathbf{r}_{q_j,j}}(t_{q_j}, j)\right)'$ and $\boldsymbol{\mu}_{\mathbf{r}_j} = \left(\mu_{\mathbf{r}_{1,j}}(t_{1,j}), \ldots, \mu_{\mathbf{r}_{q_j},j}(t_{q_j,j})\right)'$ denote the vector of the data from the jth roving sensor at spatial locations $\mathbf{r}_{l,j}$ and time points $t_{l,j}$ for $l = 1, \ldots, q_j$ and the corresponding mean vector for $j = 1, \ldots, n_R$. Also, let $\mathbf{y} = \left(\mathbf{y}'_{\mathbf{s}_1}, \ldots, \mathbf{y}'_{\mathbf{s}_{n_S}}, \mathbf{y}'_{\mathbf{r}_1}, \ldots, \mathbf{y}'_{\mathbf{r}_{n_R}}\right)'$ and $\boldsymbol{\mu} = \left(\boldsymbol{\mu}'_{\mathbf{s}_1}, \ldots, \boldsymbol{\mu}'_{\mathbf{s}_{n_S}}, \boldsymbol{\mu}'_{\mathbf{r}_1}, \ldots, \boldsymbol{\mu}'_{\mathbf{r}_{n_R}}\right)'$ denote the vector of all the observations and the corresponding mean vector.

Let $\boldsymbol{\Phi} = \operatorname{diag}\left\{\operatorname{diag}\left\{\mathbf{A}_i\right\}_{i=1}^{n_S}, \operatorname{diag}\left\{\operatorname{diag}\left\{\mathbf{b}_{l,j}\right\}_{l=1}^{q_j}\right\}_{j=1}^{n_R}\right\}$ denote a block diagonal matrix, where

$$\mathbf{A}_i = \begin{pmatrix} \varphi_1(t_{1,i}) & \cdots & \varphi_1(t_{p_i,i}) \\ \vdots & \ddots & \vdots \\ \varphi_L(t_{1,i}) & \cdots & \varphi_L(t_{p_i,i}) \end{pmatrix} \quad \text{and} \quad \mathbf{b}_{l,j} \begin{pmatrix} \varphi_1(t_{l,j}) \\ \vdots \\ \varphi_L(t_{l,j}) \end{pmatrix}.$$

Let

$$\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Lambda}_{S,S} \boldsymbol{\Lambda}_{S,R} \\ \boldsymbol{\Lambda}_{R,S} \boldsymbol{\Lambda}_{R,R} \end{pmatrix}$$

with $\boldsymbol{\Lambda}_{S,S} = (\operatorname{diag}\left\{\lambda_\ell \rho_\ell\left(\|\mathbf{s}_i - \mathbf{s}_{i'}\|; \boldsymbol{\theta}_\ell\right)\right\}_{\ell=1}^{L})_{i,i'=1}^{n_S}$, where $\lambda_\ell \rho_\ell(\|\mathbf{s} - \mathbf{s}'\|) = Cov(\xi_\ell(\mathbf{s}), \xi_\ell(\mathbf{s}'))$, $\lambda_\ell = Var(\xi_\ell(\mathbf{s}))$ and $\rho_\ell(\cdot)$ is a spatial correlation function that may be modeled by the Matérn class [Stein (1999)]. Note $\boldsymbol{\Lambda}_{S,S}$ is a block matrix with blocks corresponding to distinct spatial locations. The submatrices $\boldsymbol{\Lambda}_{S,R}$ and $\boldsymbol{\Lambda}_{R,R}$ are defined analogously; however, for a given roving sensor, each distinct spatial location corresponds to its own block. This illustrates that the covariance structure is more complex than a sampling scheme that involves only static sensors, showing that roving sensors play a role in both spatial and temporal dependence. The rank of $\boldsymbol{\Lambda}$ can be as large as $L(n_S + \sum_{j=1}^{n_R} q_j)$.

Suppose $\eta_{\mathbf{s}}(t)$ and $\epsilon_{\mathbf{s}}(t)$ are Gaussian processes. Then $\xi_\ell(\mathbf{s})$ are Gaussian processes, and $\mathbf{y}$ follows a multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance

$$\sum = \mathbf{\Phi}' \mathbf{\Lambda} \mathbf{\Phi} + \sigma^2 \mathbf{I}_N, \quad (6)$$

where $N = \sum_{i=1}^{n_S} p_i + \sum_{j=1}^{n_R} q_j$ is the total sample size combining static and roving sensor data and $\mathbf{I}_N$ is the N-dimensional identity matrix.

Estimation of all the components is not always possible depending on the choices made for the parameters θ_ℓ in the spatial correlation function $\rho_k(\cdot)$ and those made for the shape of the temporal process $\varphi_k(\cdot)$ We now develop a profile likelihood approach to parameter estimation. At initialization, we estimate the mean function $\mu_{\mathbf{s}}(t)$ by ordinary least squares (OLS) and denote the estimated mean function as $\hat{\mu}_{\mathbf{s}}(t)$. For example, a fitted mean function could comprise both a linear model spatially and a nonparametric model temporally,

$$\hat{\mu}_{\mathbf{s}}(t) = \hat{\beta}_0 + \hat{\beta}_x s_x + \hat{\beta}_y s_y + \sum_{k=1}^{K} \hat{\beta}_k \boldsymbol{B}_k(t) \quad (7)$$

where $\mathbf{s} = (s_x, s_y)'$, $B_k(\cdot)$ are cubic spline basis functions, K is the number of basis functions that controls the smoothness of $\hat{\mu}$ and $\hat{\beta}$ are the OLS estimates of the coefficients. Let $\hat{\boldsymbol{\mu}}_{\mathrm{OLS}}$ denote the vector of $\hat{\mu}_{\mathbf{s}}(t)$ at all sampling locations and time points. Let $\hat{\mathbf{y}} = \mathbf{y} - \hat{\boldsymbol{\mu}}_{\mathrm{OLS}}$ denote the detrended data comprising $\hat{\mathbf{y}}_{\mathbf{s}_i} = \left(\hat{\mathbf{y}}_{\mathbf{s}_i}(t_i, k) : k = 1, \ldots, p_i\right)'$ for $i = 1, \ldots, n_S$ $\hat{\mathbf{y}}_{\mathbf{r}_j}\left(\hat{\mathbf{y}}_{\mathbf{r}_{j,l}}(t_{j,l}) : l = 1, \ldots, q_j\right)'$ for $j = 1, \ldots, n_R$.

Next, we estimate λ_ℓ and $\varphi_k(t)$ by applying a functional principal component analysis to the data from the static sensors [Ramsay and Silverman (2005), pages 178–182]. In this step, we estimate the temporal covariance function $\hat{\gamma}_0(t, t')$ from vectors $\hat{\mathbf{y}}_{\mathbf{s}_i}$ expanded in B-spline basis functions, obtaining a functional estimate of $y_{\mathbf{s}_i}(t) - \mu_{\mathbf{s}_i}(t)$ denoted by $\hat{y}_{\mathbf{s}_i}(t)$. Thus,

$$\hat{\gamma}_0(t, t') = n_S^{-1} \sum_{i=1}^{n_S} \hat{y}_{\mathbf{s}_i}(t) \hat{y}_{\mathbf{s}_i}(t)' . \quad (8)$$

The estimate of the first temporal function $\hat{\varphi}_1(\cdot)$ is the maximizer of

$$\max_{\|f(t)\|_\zeta = 1} \int_0^T \int_0^T f(t)\, \hat{\gamma}_0(t, t') f(t')\, \mathrm{d}t\, \mathrm{d}t', \quad (9)$$

where $\|f(t)\|_\zeta = 1$ in the L^ζ norm $\langle f, g\rangle_\zeta = \int_0^T f(t)g(t)\mathrm{d}t + \zeta\int_0^T f''(t)g''(t)\,\mathrm{d}t$ is an inner product, and ζ is a tunning parameter that controls the smoothness of the estimates. The estimates of the subsequent temporal functions $\hat{\varphi}_\ell(\cdot)$ for $\ell \geq 2$ are the maximizer of (9) under the constraint of orthogonality of $\hat{\varphi}_\ell(\cdot)$ and $\hat{\varphi}_{\ell'}(\cdot)$ (in the $\langle\cdot,\cdot\rangle_\zeta$ sense) for all $\ell' < \ell$ Finally, λ_ℓ is estimated by

$$\hat{\lambda}_\ell = \int_0^T \int_0^T \hat{\varphi}_\ell(t)\,\hat{\gamma}_0(t,t')\hat{\varphi}_\ell(t')\,\mathrm{d}t\,\mathrm{d}t'$$

for $\ell = 1, \ldots, L$.

Unlike static sensors, it is not possible to obtain a full time series of data at a fixed location $\mathbf{r}_{j,l}$ for the roving sensor. Therefore, estimates of $\varphi_\ell(t)$ are based on static sensors only, and the values of $\hat{\varphi}_\ell(t)$ for the roving sensors are interpolated by evaluating the estimates $\hat{\varphi}_\ell(t)$ at time points $t_{l,j}$ for $l = 1, \ldots, q_j$, $j = 1, \ldots, n_R$. The smoothness of $\varphi_\ell(t)$ estimates is influenced by the number of basis functions K, the tuning parameter ζ and the number of functional principal components L.

Given $\hat{\boldsymbol{\mu}}$, $\hat{\lambda}_\ell$ and $\hat{\varphi}_\ell(t)$, we minimize the negative profile log-likelihood of $\boldsymbol{\theta}$ and σ^2 defined as

$$\begin{aligned} f\left(\boldsymbol{\theta}, \sigma^2\right) &= -2\ell\left(\boldsymbol{\theta}, \sigma^2\right) \\ &\propto \hat{\mathbf{y}}' \sum_1 \left(\boldsymbol{\theta}, \sigma^2\right)^{-1} \hat{\mathbf{y}} + \log\left[\det\left\{\sum_1 \left(\boldsymbol{\theta}, \sigma^2\right)^{-1}\right\}\right], \end{aligned} \tag{10}$$

where $\hat{\mathbf{y}} = y - \hat{\boldsymbol{\mu}}$ is the detrended data vector and $\sum_1\left(\boldsymbol{\theta}, \sigma^2\right) = \sum\left(\boldsymbol{\theta}, \sigma^2; \hat{\lambda}_\ell, \hat{\varphi}_\ell(t)\right)$ is the covariance (6) parameterized by $\boldsymbol{\theta}$ and σ^2 and evaluated at $\hat{\lambda}_\ell$ and $\hat{\varphi}_\ell(t)$ The solutions to (10) can be obtained by using box-constrained optimization to ensure that all estimated parameters are positive. Further, the coefficients in the mean function (7) are updated by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{X})^{-1}\mathbf{X}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{y},$$

where $\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}(\hat{\theta}, \hat{\sigma}^2; \hat{\lambda}_\ell\hat{\varphi}_\ell(t))$ is the covariance (6) evaluated at the parameter estimates and $\mathbf{X}$ is the design matrix for the covariates in (7).

4.2. Spatio-temporal kriging and prediction of spatial loadings.

To predict $y_{s0}(t)$ at an unsampled location $\mathbf{s}_0$ and time t_0, we use

$$\hat{y}_{\mathbf{s}_0}(t_0) = \hat{\mu}_{\mathbf{s}_0}(t_0) + \hat{\eta}_{\mathbf{s}_0}(t_0) = \hat{\mu}_{\mathbf{s}_0}(t_0) + \hat{\boldsymbol{\Sigma}}_{\mathbf{s}_0,t_0}\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y} - \hat{\boldsymbol{\mu}}), \tag{11}$$

where $\widehat{\mathbf{\Sigma}}_{\mathbf{s}_0, t_0} = \widehat{\mathbf{\Phi}}(t_0)' \widehat{\mathbf{\Lambda}}(\mathbf{s}_0) \widehat{\mathbf{\Phi}}(t_0)$, $\widehat{\mathbf{\Phi}}(t_0)$ is $\mathbf{\Phi}$ evaluated at t_0, and $\widehat{\mathbf{\Lambda}}(\mathbf{s}_0) = (\widehat{\mathbf{\Lambda}}_{\mathbf{s}_0, \mathrm{S}} \quad \widehat{\mathbf{\Lambda}}_{\mathbf{s}_0, \mathrm{R}})$, with $\mathbf{\Lambda}_{\mathbf{s}_0, \mathrm{S}} = \left\{ \mathrm{diag}\{\lambda_\ell \rho_\ell(\|\mathbf{s}_0 - \mathbf{s}_i\|; \theta_\ell\}_{\ell=1}^{L} \right\}_{i=1}^{n_\mathrm{S}}$, and $\mathbf{\Lambda}_{\mathbf{s}_0, \mathrm{R}} = \left\{ \mathrm{diag}\{\lambda_\ell \rho_\ell(\|\mathbf{s}_0 - \mathbf{r}_{l,j}\|; \theta_\ell\}_{\ell=1}^{L} \right\}_{l=1}^{q_\mathrm{j}}{}_{,j=1}^{n_\mathrm{R}}$. Equation (11) is used over a fine spatial grid to generate the hazard maps over time. The prediction standard error is given by $\hat{\sigma}_{\mathbf{s}_0}(t_0)$, where

$$\hat{\sigma}^2_{\mathbf{s}_0}(t_0) = \hat{\sigma}^2 + \sum_{\ell=1}^{L} \hat{\lambda}_\ell \hat{\varphi}^2_\ell(t_0) - \widehat{\mathbf{\Sigma}}'_{\mathbf{s}_0, t_0} \widehat{\mathbf{\Sigma}}^{-1} \widehat{\mathbf{\Sigma}}_{\mathbf{s}_0, t_0} + (\mathbf{x} - \mathbf{X}' \widehat{\mathbf{\Sigma}}^{-1} \widehat{\mathbf{\Sigma}}_{\mathbf{s}_0, t_0})' (\mathbf{X}' \widehat{\mathbf{\Sigma}}^{-1} \mathbf{X})^{-1} (\mathbf{x} - \mathbf{X}' \widehat{\mathbf{\Sigma}}^{-1} \widehat{\mathbf{\Sigma}}_{\mathbf{s}_0, t_0}), \tag{12}$$

and $\mathbf{x}$ are the covariates in (7) evaluated at $\mathbf{s}_0$ and t_0.

We can also predict the spatial loadings $\xi_\ell(\mathbf{s})$ using conditional expectations in an approach similar to Yao, Müller and Wang (2005) for nonspatial data. Let $\mathbf{s}_1^*, \ldots, \mathbf{s}_m^* \in D$ be the unsampled locations of interest. The ℓth loading evaluated at locations $\mathbf{s}_1^*, \ldots, \mathbf{s}_m^*$ is a linear predictor of $(\xi_\ell(\mathbf{s}_1^*), \ldots, \xi_\ell(\mathbf{s}_m^*))$. Assuming $\{\xi_\ell(\mathbf{s}_1^*), \ldots, \xi_\ell(\mathbf{s}_m^*)\}_{\ell=1}^{\infty}$ and $\varepsilon_\mathbf{s}(t)$ are jointly Gaussian, we have

$$\begin{aligned}(\hat{\xi}_\ell(\mathbf{s}_1^*), \ldots, \hat{\xi}_\ell(\mathbf{s}_m^*))' &= \widehat{\mathbb{E}}((\xi_\ell(\mathbf{s}_1^*), \ldots, \xi_\ell(\mathbf{s}_m^*))' \,|\, \mathbf{y}) \\ &= \widehat{\mathbf{\Sigma}}_{\mathbf{\xi}_\ell, \mathbf{y}} \widehat{\mathbf{\Sigma}}^{-1} (\mathbf{y} - \hat{\boldsymbol{\mu}}),\end{aligned}$$

where $\widehat{\mathbf{\Sigma}}_{\mathbf{\xi}_\ell, \mathbf{y}}$ is the plug-in estimate of $\mathbf{\Sigma}_{\mathbf{\xi}_\ell, \mathbf{y}}$ and $\mathbf{\Sigma}_{\mathbf{\xi}_\ell, \mathbf{y}}$ is the sample covariance of $(\boldsymbol{\xi}_\ell(\mathbf{s}_1^*), \ldots, \boldsymbol{\xi}_\ell(\mathbf{s}_m^*))'$ and $\mathbf{y}$, given by

$$\mathbf{\Sigma}_{\mathbf{\xi}_\ell, \mathbf{y}} = \lambda_\ell \begin{pmatrix} \varphi_\ell(t_{1,1}) \rho_\ell(\|\mathbf{s}_1^* - \mathbf{s}_1\|; \boldsymbol{\theta}_\ell) & \cdots & \varphi_\ell(t_{q_{n_\mathrm{R}}, n_\mathrm{R}}) \rho_\ell(\|\mathbf{s}_1^* - \mathbf{r}_{q_{n_\mathrm{R}}, n_\mathrm{R}}\|; \boldsymbol{\theta}_\ell) \\ \vdots & \ddots & \vdots \\ \varphi_\ell(t_{1,1}) \rho_\ell(\|\mathbf{s}_m^* - \mathbf{s}_1\|; \boldsymbol{\theta}_\ell) & \cdots & \varphi_\ell(t_{q_{n_\mathrm{R}}, n_\mathrm{R}}) \rho_\ell(\|\mathbf{s}_m^* - \mathbf{r}_{q_{n_\mathrm{R}}, n_\mathrm{R}}\|; \boldsymbol{\theta}_\ell) \end{pmatrix},$$

and $\widehat{\mathbf{\Sigma}}$, $\mathbf{y}$ and $\hat{\boldsymbol{\mu}}$ are as defined in Section 4.1.

4.3. Inhomogeneous variances.

The model given in (1) assumes that the measurement error variance is the same for the static and roving sensors. In practice, however, this assumption does not always hold. In the following we extend the data model to accommodate the situation that the measurement error variance for the static sensors is different from the roving sensors.

Consider model (1) again,

$$y_{\mathbf{s}}(t) = \mu_{\mathbf{s}}(t) + \eta_{\mathbf{s}}(t) + \varepsilon_{\mathbf{s}}(t), \quad (13)$$

but with measurement error variances $\mathrm{Var}(\varepsilon_{\mathbf{s}}(t)) = \sigma_{\mathrm{S}}^2$ for static sensors and $\mathrm{Var}(\varepsilon_{\mathbf{s}}(t)) = \sigma_{\mathrm{R}}^2$ for roving sensors. The model (13) will be referred to as the *inhomogeneous variance* case. The model (1) is a special case where $\sigma_{\mathrm{S}}^2 = \sigma_{\mathrm{R}}^2$ and will be referred to as the *homogeneous variance* case. In the inhomogeneous variance case, the covariance is a general case of (6), and the estimation algorithm is modified by using

$$\mathbf{\Sigma} = \mathbf{\Phi}' \mathbf{\Lambda} \mathbf{\Phi} + \sigma_{\mathrm{S}}^2 \mathbf{D}_{\mathrm{S}} + \sigma_{\mathrm{R}}^2 \mathbf{D}_{\mathrm{R}}, \quad (14)$$

where $\mathbf{D}_{\mathrm{S}}$, $\mathbf{D}_{\mathrm{R}}$ are diagonal matrices with diagonal entries equal to 1 for static and roving sensors, respectively, and 0 otherwise. The model (13) allows the fusion of the two sources of data with varying degrees of spatio-temporal resolutions and drawing inference about the true underlying process $\mu_{\mathbf{s}}(t) + \eta_{\mathbf{s}}(t)$. The profile-likelihood approach to parameter estimation is modified as follows. The data are detrended with initial estimates of $\boldsymbol{\mu}$ using (7), and the λ_ℓ and $\varphi_\ell(t)$ terms are estimated using functional principal component analysis over the static sensors as before. Given $\hat{\boldsymbol{\mu}}$, $\hat{\lambda}_\ell$ and $\hat{\varphi}(t)$, we minimize the negative profile log-likelihood of $\boldsymbol{\theta}$, σ_{S}^2 and σ_{R}^2 defined as

$$f(\boldsymbol{\theta}, \sigma_{\mathrm{S}}^2, \sigma_{\mathrm{R}}^2) = -2\ell(\boldsymbol{\theta}, \sigma_{\mathrm{S}}^2, \sigma_{\mathrm{R}}^2) = \hat{\mathbf{y}}' \mathbf{\Sigma}_2(\boldsymbol{\theta}, \sigma_{\mathrm{S}}^2, \sigma_{\mathrm{R}}^2)^{-1} \hat{\mathbf{y}} + \log[\det\{\mathbf{\Sigma}_2(\boldsymbol{\theta}, \sigma_{\mathrm{S}}^2, \sigma_{\mathrm{R}}^2)\}] + N \log(2\pi), \quad (15)$$

Where $\hat{\mathbf{y}} = \mathbf{y} - \hat{\boldsymbol{\mu}}$ and $\mathbf{\Sigma}_2(\boldsymbol{\theta}, \sigma_{\mathrm{S}}^2, \sigma_{\mathrm{R}}^2) = \mathbf{\Sigma}(\boldsymbol{\theta}, \sigma_{\mathrm{S}}^2, \sigma_{\mathrm{R}}^2; \hat{\lambda}_\ell, \hat{\varphi}_\ell(t))$ given in (14) is parameterized by $\boldsymbol{\theta}$, σ_{S}^2 and σ_{R}^2 evaluated at $\hat{\lambda}_\ell$ and $\hat{\varphi}_\ell(t)$. The coefficients in the mean function are updated by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}' \hat{\mathbf{\Sigma}}^{-1} \mathbf{X})^{-1} \mathbf{X} \hat{\mathbf{\Sigma}}^{-1} \mathbf{y},$$

where $\hat{\mathbf{\Sigma}} = \mathbf{\Sigma}(\hat{\boldsymbol{\theta}}, \hat{\sigma}_{\mathrm{S}}^2, \hat{\sigma}_{\mathrm{R}}^2; \hat{\lambda}_\ell, \hat{\varphi}_\ell(t))$ and $\mathbf{X}$ is the design matrix for the covariates in (7).

5. Case study: Spatio-temporal occupational hazard mapping.

5.1. Model fitting.

Before fitting an inhomogeneous variance model for the noise data, we selected the tuning parameters by a leave-one-sensor-out cross-validation approach. More specifically, we considered a grid of values for the number of deterministic spline basis functions K, the number of temporal principal components L and the principal components smoothing parameters ζ, and searched for a minimizer of the estimated mean squared prediction error (MSPE); see Appendix A.2 of the Supplemental Materials. The selected parameters are $L = 3$, $K = 12$ and $\zeta = 0$. The estimated deterministic component is

$$\hat{\mu}_{\mathbf{s}}(t) = 83.64 - 0.40 s_x + 0.31 s_y + \hat{S}(t),$$

where the spline term $\hat{S}(t)$ is shown in Figure 7(a). The estimates of the spatial process variances are $\hat{\lambda}_1 = 13.20$, $\hat{\lambda}_2 = 8.07$ and $\hat{\lambda}_3 = 0.13$. The estimates for spatial range parameters, assuming an exponential spatial covariance model $\rho_\ell(\|\mathbf{s}_i - \mathbf{s}_{i'}\|; \boldsymbol{\theta}_\ell) = \exp\{-\|\mathbf{s}_i - \mathbf{s}_{i'}\| / \theta_\ell\}$, for $\ell = 1, 2, 3$, are $\hat{\theta}_1 = 22.34$, $\hat{\theta}_2 = 10.831$ and $\hat{\theta}_3 = 40.34$. The nonparametric estimates of the temporal functions $\boldsymbol{\varphi}_\ell(\cdot)$ are shown in Figure 7(b). The estimated measurement error variances are $\hat{\sigma}^2_{\mathrm{S}} = 1.49$ and $\hat{\sigma}^2_{\mathrm{R}} = 1.05$. The parameter estimates for the homogeneous variance case and for the case in which only the static sensor data are used are given in Table 1. The inhomogeneous variance case is denoted by STDF (which stands for spatio-temporal data fusion), the homogeneous variance case by STDFh, and the scenario with static sensors only by STDF*.

A series of dynamic hazard maps for the predicted noise intensity using our STDF model in Figure 4 show overall low intensity levels at the beginning and near the end of the study (panels 09:50 am and 11:20 am). They identify peaks in most time transects around the spatial coordinates $s_x = 2.5$ and $s_y = 35$, the noise source located. The noise intensity decreases as the readings are made further away from the noise source. In addition, a secondary noise source located near sensor #18 is captured [see also Figure 1(a)]. The standard errors of the predicted noise intensity across time and space are plotted in Figure 5. As expected, lower standard errors are found near the static sensors and along the trajectories of the roving sensors, as shown in Figure 1. The prediction standard error is larger at sites further apart from the static sensors location. The locations visited by the roving sensors display smaller standard errors than unsampled sites but higher standard errors than the static sensors' sites. The prediction standard error levels at roving sensor locations are carried over time and do not change as a function of the roving sensor position. The prediction standard errors are also larger in general at the end of the experiment (starting at around 11:20 am). This corresponds to the time period in which the main source of noise was turned off, and the secondary noise source increased in intensity, as shown in Figure 2. Such an increase in the standard error levels is due to $\hat{\varphi}_2(t)$ in (12), which captures the temporally short, but high local noise intensity change around sensor 18 and the temporally short, abrupt decrease in noise intensity in the northernmost room.

Figure 6 maps the estimated $(\hat{\xi}_\ell(\mathbf{s}_1^*), \ldots, \hat{\xi}_\ell(\mathbf{s}_m^*))'$, $\ell = 1, 2, 3$ over a fine grid of spatial locations $\mathbf{s}_1^*, \ldots, \mathbf{s}_m^*$, with each loading standardized to have zero mean and standard deviation one. The temporal components $\hat{\varphi}_\ell(t)$, $\ell = 1, 2, 3$ are shown in Figure 7(b). The interpretation of the the temporal functions $\varphi_\ell(\cdot)$ can be made in light of Figure 2, as well as the Karhunen–Loève expansion based on (5). The first temporal component $\hat{\varphi}_1$ is nearly constant around 1 (Figure 7). The corresponding $\hat{\xi}_1(\mathbf{s})$ shown in Figure 6 shows higher noise intensity values in the northern room and near the secondary source around sensor #18. For the second temporal component, since $\hat{\varphi}_2$ is negative until approximately 11:10 am, it subtracts the $\hat{\xi}_2(\mathbf{s})$ effect from $\hat{\xi}_1(\mathbf{s})$, but after 11:10 am it adds the effect of the secondary noise source. Thus, in the dynamic hazard maps (Figure 4), the noise intensity is higher in the northern room before 11:10 am, and afterward the noise intensity is higher near sensor #18. The third temporal component $\hat{\varphi}_3$ is associated with a much variance ($\hat{\lambda}_3 = 0.13$), and does not impact the hazard map as much as the first two components. The roles of $\hat{\varphi}_3$ and $\hat{\xi}_3(\mathbf{s})$ seem to be compensating $\hat{\xi}_1(\mathbf{s})$ and $\hat{\xi}_2(\mathbf{s})$ at the very beginning and the very end of the experiment in order for the hazard map to be closer to background noise.

The standard errors for the parameter estimates are obtained via cross-validation by leaving one sensor out at each time, as detailed in Appendix A.1, and are shown in Table 1. While the main objective of Table 1 is to quantify the uncertainty regarding the parameter estimates, we note that the large range parameter estimate $\hat{\theta}_3$ reflects a relatively weak spatial dependence. Consequently, the standard error for $\hat{\theta}_3$ also might be inflated by static sensors within a certain radius of each other being removed during the cross-validation step. Table 1 also displays how the roving sensors affect the estimates of the spatial dependence coefficients. When only the static sensors are used, the range parameter estimates $\hat{\theta}_2$ and $\hat{\theta}_3$ are similar. However, when the roving sensors are included, the estimate $\hat{\theta}_2$ becomes quite a bit smaller than $\hat{\theta}_3$ and is more informative, in the sense that the spatial effect of the secondary noise source becomes more prominent owing to the roving sensors near this secondary source toward the end of the experiment. This illustrates that the inclusion of roving sensors adds information about the spatial dependence at finer scales.

5.2. Scientific implications.

Both static (Figure 3) and dynamic (Figure 4) hazard maps capture the high noise intensity in the northern room near the working engine. However, the dynamic hazard maps also show that the average intensity of hazard exposure in the southern room increases during the period between 10:00 am and 11:00 am when the engine in the northern room is turned on. Further, the outside noise source is detected in the southeastern corner after 11:10 am. In particular, the intensity levels exceed 85 dB, which is generally viewed as a harmful level for exposures longer than 8 hours. While the static maps do detect a small noise peak near the same location, the predicted intensity levels are understated.

The problem of interpolating hazard maps in time and space from discrete sampled observations was discussed in Koehler and Peters (2013). While kriging is well accepted for spatial interpolation in the occupational hygiene literature, often the observations or the hazard maps are averaged in time to produce estimates, and thus the temporal aspect of the data is ignored. To compare our method with the existing approaches, we consider kriging using the roving data only and kriging while incorporating roving and static sensor data, both averaging the observations in time. We also consider fixed-time universal kriging (UK), fixed-time thin-plate spline (TPS) and fixed-time simple linear regression on the spatial coordinates (LM). By fixed-time we mean that a time point is fixed and the spatial map is constructed for the data corresponding to the time transect selected, thus preserving some of the temporal structure from the data.

To compare the methods globally, the MSPE values for our spatio-temporal data fusion method are obtained using the leave-one-sensor-out cross-validation described in Appendix A.1. We generate prediction maps at every time point in which the static sensors were sampled, and averaged the values across space and over time. When including the roving sensors, the inhomogeneous variance case (STDF) has an MSPE of 11.82, and the homogeneous version (STDFh) has an MSPE of 11.66. When using only the static sensors (STDF*), the MSPE is 14.56. The MSPE for the static map using roving sensors only is 37.42, and the MSPE for the static map incorporating roving and static sensor data is 38.33. This shows a clear advantage of our method over the current practice of using static maps. For the other alternative approaches, the MSPE for fixed-time universal kriging (UK) is 24.4, thin-plate spline regression (TPS) is 14.99, and simple linear regression model (LM) is 25.14, all of which are outperformed by our method. We observe that the homogeneous variance case gives similar predictions to the inhomogeneous variance case.

In Figure 8 we focus on a small-scale example from the dataset to illustrate the possibility of detecting short duration hazard intensity peaks. The static sensors are sampled every minute (Section 2.1), and thus the measurements are available between 11:04 am and 11:05 am. For the UK and TPS generated maps, we use data from 11:04 am and 11:05 am, and interpolate linearly the values from the maps at times 11:04:20 and 11:04:40. We can see in Figure 8 that UK maps do not capture the secondary noise source in the southeastern part of the facility. This is because the method underestimates the range of spatial dependence, and produces predictions that resemble a plane except where the static sensors are located. On the other hand, the TPS method oversmooths the data, losing local features such as the sharper distinction between the northern and southern rooms in the facility. We omit the linear model estimates, which are planes only. In addition, Figure 9 shows the corresponding standard error maps for each method in Figure 8. We observe that the STDF maps generally have lower standard errors than those of UK. The TPS only has comparable standard errors when very near the static sensors, but otherwise much higher standard errors than either STDF or UK.

A final remark is that we can consider cases in which the spatial covariance function is not exponential. For example, we repeated the analysis using the Matérn class of covariance functions, with known smoothness parameter $\nu = 2.5$. The resulting hazard maps (not

shown) are similar except for slightly smoother contour lines. We anticipate that our method is robust to the choice of a covariance function in the Matérn class.

6. Discussion.

In this paper, we have developed a spatio-temporal static and roving data fusion model, with each data sensor having potentially different instrument variances. The approach to model fitting and statistical inference has been applied to produce hazard maps that capture dependence across space and over time in indoor environments. Modeling the spatio-temporal dependence structure allows the hazard maps to capture features that are missed by the current practice in occupational hazard assessment. Furthermore, our approach enables continuous-time prediction of hazard, which the existing approaches are unable to produce.

With the semiparametric model specification, our method is able to detect unexpected hazard sources that occur sporadically during a study. A sudden fluctuation of intensity, such as the secondary noise source in the southeastern corner of the facility, are undetected or underestimated when using current practices, but can be detected by our method. Moreover, health effects of short duration but high-level exposures are unclear, and our method provides a way to better capture such transient exposures.

Cross-validation shows that our methodology outperforms the traditional methods in the scientific application, a conclusion that is corroborated by the simulation study given in Appendix B of the Supplemental Materials. The simulations have also demonstrated that our method is robust to different instrument variances, while the traditional approaches tend to provide less accurate prediction.

While the height of the sensors is not accounted for directly, the model with heterogeneous measurement error variances may accommodate possibly different heights for different sensors. It would be interesting, however, to examine this third dimension more closely, as well as to consider three-dimensional hazard maps when data are collected at different heights [see, e.g., Tracey et al. (2014)].

Other covariance modeling allows for nonseparability, although stationarity in time is generally assumed [Gneiting (2002), Ma (2003), Quick, Banerjee and Carlin (2015)]. Stein (2005) proposed models that are asymmetric in time, allowing for different smoothness degrees in space over time. It may be of interest to extend such models and develop estimation methods for fusion of static and roving sensor data. Moreover, while the methodology developed here is geared toward spatio-temporal hazard mapping, we believe that other scientific disciplines might benefit from our approach for fusing data with very different spatial and/or temporal scales. For example, for data collected by individuals with personal devices versus data collected at stationary monitoring stations used to study exposure and averse health effects in environmental epidemiology [see, e.g., Hall and McMullen (2004)]. The applicability of our approach to other scientific studies is currently under investigation, such as integration of animal movement data with static field cameras in ecological monitoring. At this point, our preliminary studies show that when the sample size increases, the increase in the computational cost is higher for more roving sensor data than

static sensor data. For datasets of much larger sample size with more sampling locations and/or time points, the proposed profile likelihood approach would need to be improved for it to be more computationally feasible by utilizing some form of approximation, such as blocking [Caragea and Smith (2007), Chu, Wang and Zhu (2014)] or tapering [Furrer, Genton and Nychka (2006), Kaufman, Schervish and Nychka (2008), Du, Zhang and Mandrekar (2009), Chu, Zhu and Wang (2011)], in the covariance matrix inversion. We leave this for future research as well.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments.

The authors gratefully acknowledge the Editor, an Associate Editor and four anonymous reviewers for their excellent comments and constructive suggestions that helped to improve this manuscript in content and presentation.

REFERENCES

Caragea PC and Smith RL (2007). Asymptotic properties of computationally efficient alternative estimators for a class of multivariate normal models. J. Multivariate Anal 98 1417–1440. MR2364128

Cherrie JW (2003). Commentary: The beginning of the science underpinning occupational hygiene. Ann. Occup. Hyg 47 179–185.12639831

Chu T , Wang H and Zhu J (2014). On semiparametric inference of geostatistical models via local Karhunen–Loève expansion. J. R. Stat. Soc. Ser. B. Stat. Methodol 76 817–832. MR3248678

Chu T , Zhu J and Wang H (2011). Penalized maximum likelihood estimation and variable selection in geostatistics. Ann. Statist 39 2607–2625. MR2906880

Cowles MK , Zimmerman DL , Christ A and McGinnis DL (2002). Combining snow water equivalent data from multiple sources to estimate spatio-temporal trends and compare measurement systems. J. Agric. Biol. Environ. Stat 7 536–557.

Cressie NAC (1993). Statistics for Spatial Data. Wiley, New York MR1239641

Cressie N and Wikle CK (2011). Statistics for Spatio-Temporal Data. Wiley, Hoboken, NJ MR2848400

Du J , Zhang H and Mandrekar VS (2009). Fixed-domain asymptotic properties of tapered maximum likelihood estimators. Ann. Statist 37 3330–3361. MR2549562

Evans DE , Heitbrink WA , Slavin TJ and Peters TM (2008). Ultrafine and respirable particles in an automotive grey iron foundry. Ann. Occup. Hyg 52 9–21.18056626

Furrer R , Genton MG and Nychka D (2006). Covariance tapering for interpolation of large spatial datasets. J. Comput. Graph. Statist 15 502–523. MR2291261

Gneiting T (2002). Nonseparable, stationary covariance functions for space-time data. J. Amer. Statist. Assoc 97 590–600. MR1941475

Gromenko O and Kokoszka P (2013). Nonparametric inference in small data sets of spatially indexed curves with application to ionospheric trend determination. Comput. Statist. Data Anal 59 82–94. MR3000043

Hall DL and Mcmullen SA (2004). Mathematical Techniques in Multisensor Data Fusion. Artech House, Boston.

Isaacson JD and Zimmerman DL (2000). Combining temporally correlated environmental data from two measurement systems. J. Agric. Biol. Environ. Stat 5 398–416. MR1812083

Kaufman CG , Schervish MJ and Nychka DW (2008). Covariance tapering for likelihood-based estimation in large spatial data sets. J. Amer. Statist. Assoc 103 1545–1555. MR2504203

Koehler KA and Peters TM (2013). Influence of analysis methods on interpretation of hazard maps. Ann. Occup. Hyg 57 558–570.23258453

Koehler KA and Volckens J (2011). Prospects and pitfalls of occupational hazard mapping: 'between these lines there be dragons'. Ann. Occup. Hyg 55 829–840.21917819

Lake K , Zhu J , Wang H , Volckens J and Koehler KA (2015). Effects of data sparsity and spatiotemporal variability on hazard maps of workplace noise. J. Occup. Environ. Hyg 12 256–265.25437137

Ludwig G , Chu T , Koehler K , Wang H and Zhu J (2017a). Supplement to "Static and Roving Sensor Data Fusion for Spatio-Temporal Hazard Mapping with Application to Occupational Exposure Assessment." DOI:10.1214/16-AOAS995SUPPA.

Ludwig G , Chu T , Koehler K , Wang H and Zhu J (2017b). Supplement to "Static and Roving Sensor Data Fusion for Spatio-Temporal Hazard Mapping with Application to Occupational Exposure Assessment." DOI:10.1214/16-AOAS995SUPPB.

Ma C (2003). Families of spatio-temporal stationary covariance models. J. Statist. Plann. Inference 116 489–501. MR2000096

O'Brien DM (2003). Aerosol mapping of a facility with multiple cases of hypersensitivity pneumonitis: Demonstration of mist reduction and a possible dose/response relationship. Appl. Occup. Environ. Hyg 18 947–952.14555448

Ologe FE , Akande TM and Olajide TG (2006). Occupational noise exposure and sensorineural hearing loss among workers of a steel rolling Mill. Eur. Arch. Oto-Rhino-Laryngol 263 618–621.

Peters TM , Heitbrink WA , Evans DE , Slavin TJ and Maynard AD (2006). The mapping of fine and ultrafine particle concentrations in an engine machining and assembly facility. Ann. Occup. Hyg 50 249–257.16361396

Peters TM , Anthony TR , Taylor C , Altmaier R , Anderson K and T O'shaughnessy P (2012). Distribution of particle and gas concentrations in swine gestation confined animal feeding operations. Ann. Occup. Hyg 56 1080–1090.22904211

Quick H , Banerjee S and Carlin BP (2015). Bayesian modeling and analysis for gradients in spatiotemporal processes. Biometrics 71 575–584. MR340259325898989

Ramsay JO and Silverman BW (2005). Functional Data Analysis, 2nd ed. Springer, New York MR2168993

Sahu SK , Gelfand AE and Holland DM (2010). Fusing point and areal level space-time data with application to wet deposition. J. R. Stat. Soc. Ser. C. Appl. Stat 59 77–103. MR2750133

Smith BJ and Cowles MK (2007). Correlating point-referenced Radon and areal uranium data arising from a common spatial process. J. Roy. Statist. Soc. Ser. C 56 313–326. MR2370992

Stein ML (1999). Interpolation of Spatial Data: Some Theory for Kriging. Springer, New York MR1697409

Stein ML (2005). Space-time covariance functions. J. Amer. Statist. Assoc 100 310–321.MR2156840

Tornero-Velez R , Symanski E , Kromhout H , Yu RC and Rappaport SM (1997). Compliance versus risk in assessing occupational exposures. Risk Anal 17 279–292.9232013

Tracey JA , Sheppard J , Zhu J , Wei F , Swaisgood RR and Fisher RN (2014). Movement-based estimation and visualization of space use in 3D for wildlife ecology and conservation. PLoS ONE 9 e101205.24988114

Wahba G (1990). Spline Models for Observational Data CBMS-NSF Regional Conference Series in Applied Mathematics 59 SIAM, Philadelphia, PA MR1045442

Yao F , Müller H-G and Wang J-L (2005). Functional linear regression analysis for longitudinal data. Ann. Statist 33 2873–2903. MR2253106

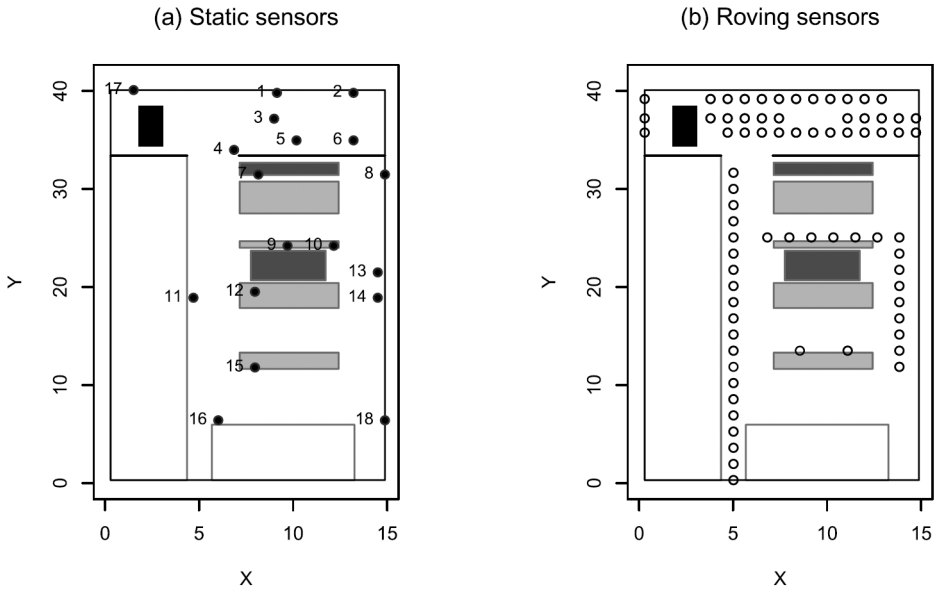

Fig. 1.

(a) The floor plan of the engine test facility. The black rectangle in the upper left corner is the source of noise, white rectangles are offices, gray rectangles are inactive engines, and dark gray rectangles are floor openings. The locations of static sensors are numbered from 1 to 18. (b) The pathway of the first roving sensor is drawn in open circles. The pathway of the second roving sensor is similar, and thus omitted.

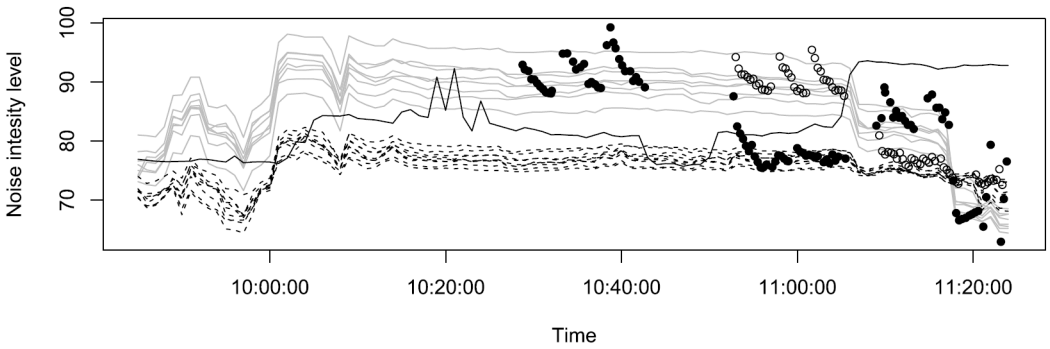

Fig. 2.

Observed noise intensity over time from 9:45:00 am to 11:23:20 am. Gray solid lines are time series for static sensors near one noise source (#1 through 7, and 17). The black solid line is near the secondary noise source (#18). Dashed lines are for the remainder static sensors. Filled and open circles are samples taken by the first roving sensor that started at 10:28:45 am and the second roving sensor that started at 10:52:20 am, respectively.

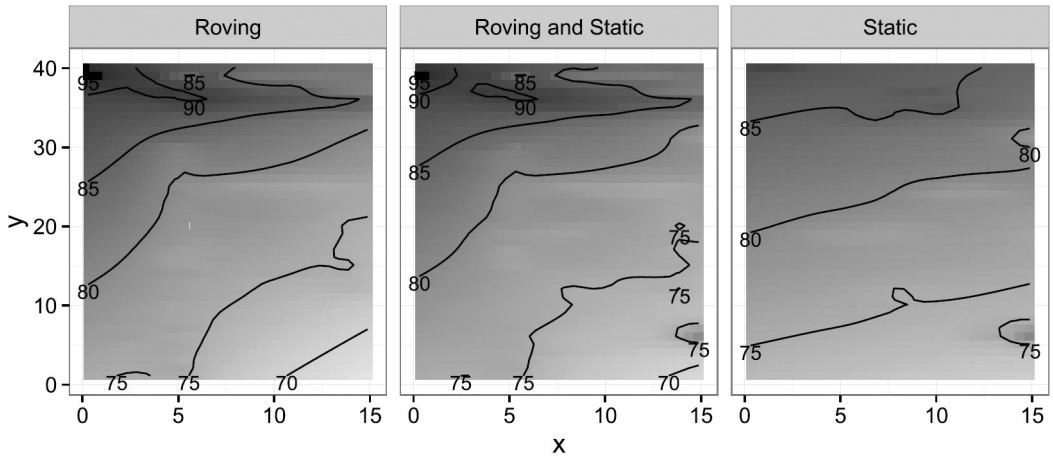

Fig. 3.

Static maps of the noise intensity obtained by kriging using the roving sensor data only (left), the roving and static sensor data (center) and the static sensor data only (right), averaging data at the same location in time.

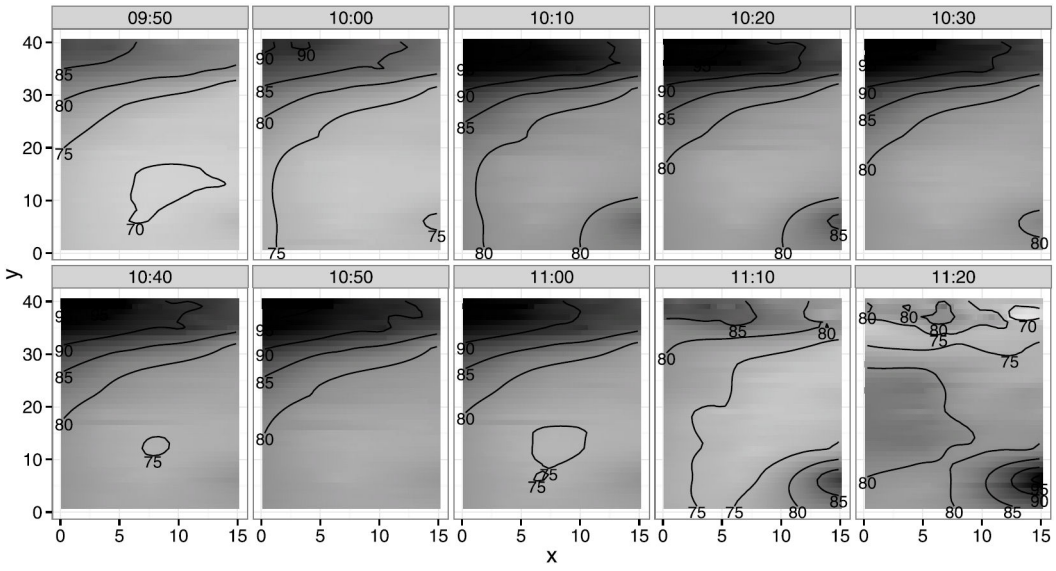

Fig. 4.

Dynamic hazard maps with contour lines obtained from the spatio-temporal data fusion (STDF) method; each panel corresponds to a point in time from 9:50 am to 11:20 am at 10-minute intervals.

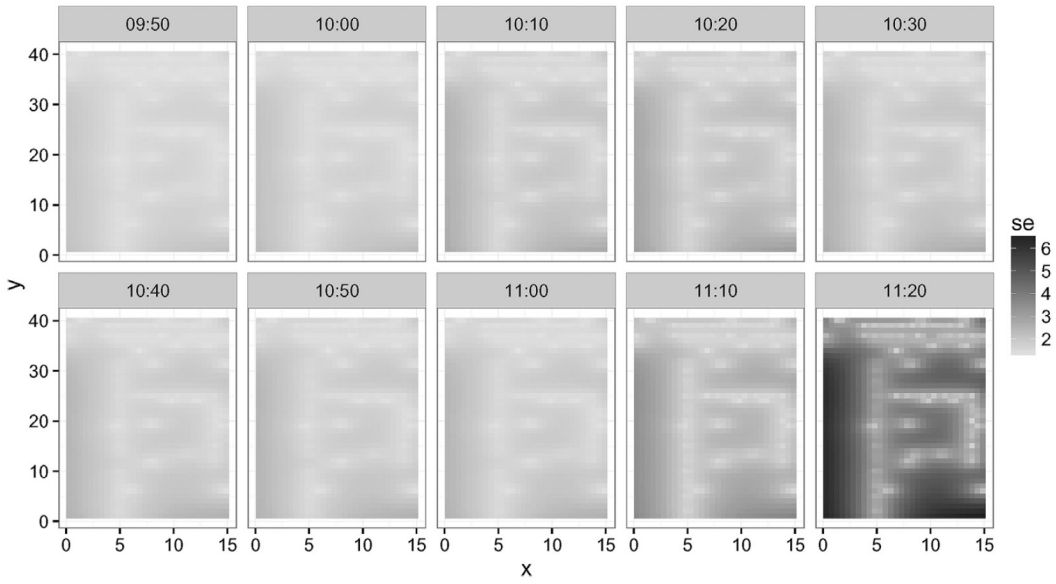

Fig. 5.

Prediction standard error maps for the dynamic hazard maps given in Figure 4; each panel corresponds to a point in time from 9:50 am to 11:20 am at 10-minute intervals.

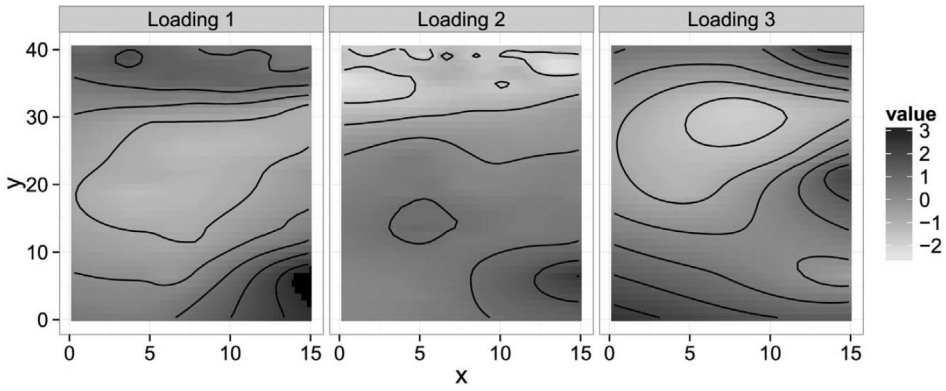

Fig. 6.

Estimated $(\hat{\xi}_{\ell}(\mathbf{s}_1^*), \ldots, \hat{\xi}_{\ell}(\mathbf{s}_m^*))'$, $\ell = 1, 2, 3$. Each ℓ loading is standardized to have zero mean and standard deviation one. The associated variability coefficients are $\hat{\lambda}_1 = 13.20$, $\hat{\lambda}_2 = 8.07$ and $\hat{\lambda}_3 = 0.13$.

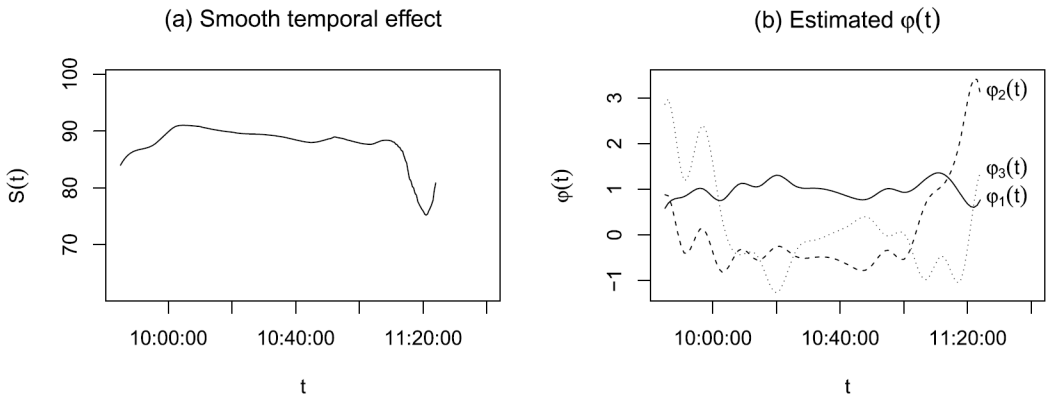

Fig. 7.

(a) *Deterministic effect for the temporal component estimated with a spline function*; (b) $\hat{\varphi}_{\ell}(t)$, $\ell = 1, 2, 3$ *components, corresponding to a random effect in the temporal component.*

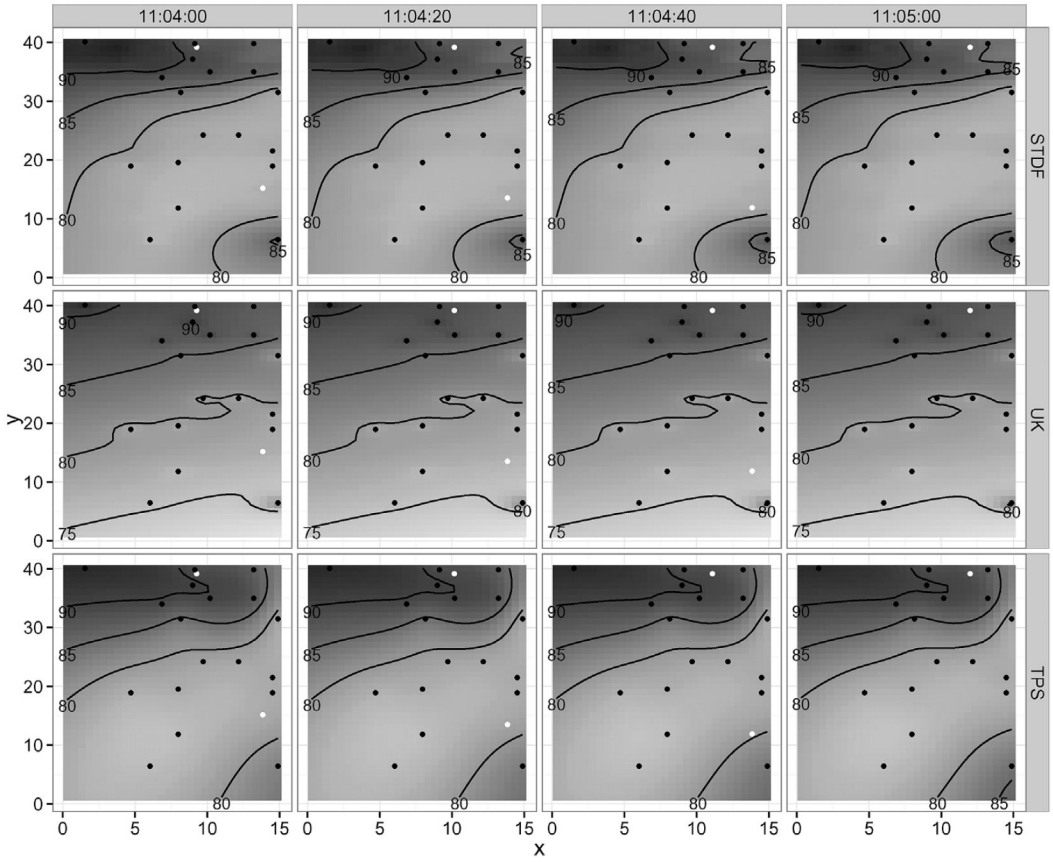

Fig. 8.

Comparison of STDF hazard map with maps created by universal kriging (UK) and thin–plate splines (TPS) methods. For the UK and TPS, data are from 11:04:00 and 11:05:00, and interpolated linearly between 11:04:20 and 11:04:40. Black points mark the static sensor locations, while white points mark the roving sensor locations at the corresponding time.

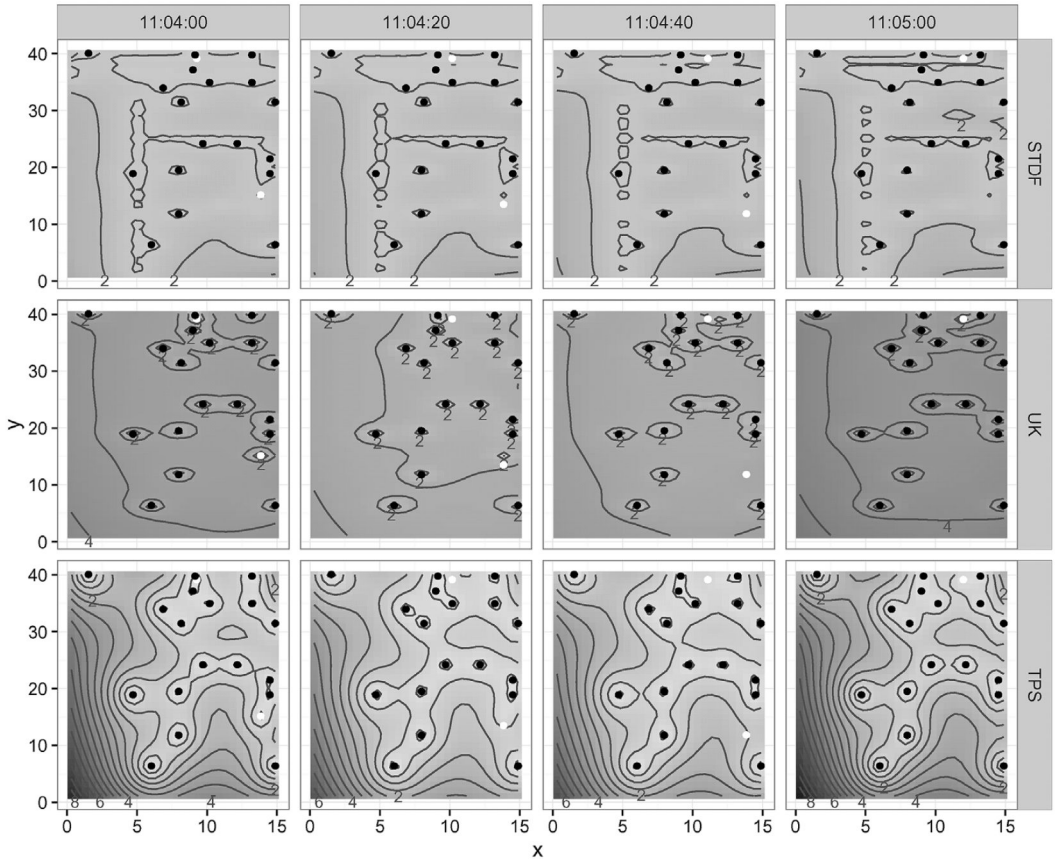

Fig. 9.

Comparison of standard errors of STDF hazard maps with the standard error of universal kriging (UK) and thin-plate splines (TPS) methods. For UK and TPS, data are from 11:04:00 and 11:05:00, and interpolated linearly between 11:04:20 and 11:04:40. Black points mark the static sensor locations, while white points mark the roving sensor locations at the corresponding time.

Table 1

Parameter estimates and cross-validated standard errors (in parenthesis). STDF denotes spatio-temporal data fusion for the inhomogeneous variance case, STDFh for the homogeneous variance case, and STDF* for only the static sensors

Coefficient	STDF	STDFh	STDF*
β_0	83.64 (4.52)	83.53 (4.49)	82.23 (3.59)
β_x	−0.40 (0.08)	−0.40 (0.08)	−0.38 (0.11)
β_y	0.31 (0.04)	0.31 (0.04)	0.31 (0.04)
θ_1	22.34 (2.93)	22.20 (2.95)	13.17 (1.83)
θ_2	10.93 (2.13)	12.08 (1.88)	30.99 (2.57)
θ_3	40.34 (11.15)	40.34 (11.10)	32.56 (8.42)
σ^2_{S}	1.49 (0.16)	1.48 (0.15)	1.46 (0.12)
σ^2_{R}	1.05 (0.29)	–	–