HHS Public Access

Author manuscript

Ann Work Expo Health. Author manuscript; available in PMC 2023 August 29.

Published in final edited form as:

Ann Work Expo Health. 2017 March 01; 61(2): 183–194. doi:10.1093/annweh/wxw025.

Sampling Strategies for Accurate Hazard Mapping of Noise and Other Hazards Using Short-Duration Measurements

Kirsten A. Koehler[1,*], **Jun Zhu**[2,3], **Haonan Wang**[4], **Thomas M. Peters**[5]

[1]Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, Baltimore, MD 21205, USA

[2]Department of Statistics, University of Wisconsin, 1300 University Avenue, Madison, WI 53706, USA

[3]Department of Entomology, University of Wisconsin, 1450 Linden Drive, Madison, WI 53706, USA

[4]Department of Statistics, Colorado State University, 1877 Campus Delivery, Fort Collins, CO 80523, USA

[5]Department of Occupational and Environmental Health, University of Iowa, 145 N. Riverside Drive, Iowa City, IA 52242, USA

Abstract

Hazard mapping is an effective way to depict spatial variability in hazard intensity obtained with direct-reading instruments on a facility floor plan. However, the extent to which temporal variability affects map accuracy is unknown, and guidance on sampling strategies to minimize map bias is lacking. In this study, we evaluated the accuracy of hazard maps produced for simulated sources and sampling strategies in a hypothetical facility. Hazard maps were produced from sampled data at high, mid, and low spatial resolution and with and without replicates and compared to a reference time-weighted average hazard map using several map comparison metrics. In agreement with 'real-world' mapping datasets, the simulation showed that increasing the number of replicates improved the overall comparability of the hazard map produced from the sampled data with the time-weighted average hazard map more efficiently than increasing the sampling spatial resolution. However, if accurately capturing peak exposures near sources is of interest, increasing the spatial resolution of the measurements, particularly near sources, is needed. From these results, we formulated guidelines to use the preliminary assessment of the temporal variability of large and intermittent sources to inform the spatial resolution and need for replicate measures to minimize the bias in hazard maps.

Keywords

hazard mapping; spatiotemporal variability

[*]Author to whom correspondence should be addressed. Tel: +1-410-955-7706; fax: +1-410-955-9334; kirsten.koehler@jhu.edu.

Supplementary Data

Supplementary data are available at *Annals of Work Exposures and Health* online.

Introduction

Occupational hazard maps—depictions of hazard concentration or intensity (hereafter, 'intensity') displayed on a two-dimensional floor plan of a facility—can inform decision making in the control of exposures to ultimately improve worker health (O'Brien, 2003). Often, the data used to produce these maps are obtained by measuring the intensity of the hazard over short durations (several minutes or less) with a direct-reading instrument (DRI) by an operator who follows a predetermined sampling path through the facility, referred to as 'roving monitoring' (Lake *et al.*, 2015). To mitigate uncertainty introduced from temporal variability, an entire set of mapping measurements are typically completed within a short time period (e.g. 1–2 h) with a nominally 1-min sample obtained at each point and ~1 min to move to the next sampling location. Thus, ~60 points on the path can be measured in 2 h.

Originally developed to document the changes in mass concentration of oil mist over time in the automotive industry (O'Brien, 2003), mapping has emerged as a way to better understand hazards in a variety of industries, such as engine machining (Dasch *et al.*, 2005; Peters *et al.*, 2006), foundries (Heitbrink *et al.*, 2007), restaurants (Park *et al.*, 2010), apparel manufacturing (Vosburgh *et al.*, 2011), and plastics manufacturing (Lake *et al.*, 2015). Moreover, different types of DRIs can be used simultaneously to provide information on multiple hazards, such as the simultaneous mapping of ventilation parameters (carbon dioxide, carbon monoxide, and temperature) with dust concentrations (Peters *et al.*, 2012).

Hazard maps are only as reliable as the data that are used to produce them. Koehler and Volckens (2011) used a small simulation study to show that exposures to workplace hazards can be mischaracterized due to temporal variability in hazard intensity. Lake *et al.* (2015) attempted to validate the results of Koehler and Volckens (2011) using noise data collected in two facilities. In each facility, they measured the spatiotemporal variability of noise intensity with 15 DRIs at fixed locations referred to as 'static monitoring', as well as roving monitoring by two operators. In agreement with simulations of Koehler and Volckens (2011), they found that noise intensity maps based on short-duration, roving measurements provided a poor estimate of time-weighted average (TWA) intensity at facilities with sources exhibiting high temporal variability. They also found that the temporal variability in hazard intensity could be mistakenly portrayed as spatial variability, particularly when few measurements were used for the generation of a map.

Specific guidelines for selecting a sampling strategy for conducting roving monitoring campaigns that reduce uncertainties in hazard maps are unavailable for the practicing industrial hygienist (IH). Specifically, guidelines are needed for the number of sampling locations within a facility and the number of repeated measurements that should be taken at each sampling location to result in the most representative hazard maps that are aligned with the goals of the monitoring. Such goals may be to ascertain TWA hazard intensity or peaks in intensity to target specific areas for the focus of control measures. Ideally, such guidelines would be tailored for a given workplace based on preliminary measurements of temporal variability in source intensity that the practicing IH can easily make. Without this type of guidance, the reliability and comparability of hazard maps is questionable.

Thus, the objective of this work was to investigate the extent to which the temporal variability of sources affects the accuracy of hazard maps. We conclude with recommendations for sampling strategies to reduce the impact of temporal variability on maps to estimate the TWA intensity.

Methods

Simulated hazard maps

Simulated hazard intensity in a facility and TWA hazard map generation— Simulated hazard intensities were modeled for a theoretical facility following methods similar to those described in Koehler and Volckens (2011). The simulation consisted of three cases wherein one, three, or five sources were modeled in a facility—a square room (100 × 100 arbitrary units that can be thought of as meters) with no interior structures. For each case, 500 iterations were run where each iteration represents intensities modeled over time and space for an 8-h period (i.e. one work day). For each of the 500 iterations, the source location(s) were positioned randomly (with one, three, or five sources modeled, depending on the case). The hazards emanating from these sources (and their relative magnitudes) were arbitrary, but can be thought of as noise here. The emission rates from each source were randomly determined, but included temporal correlation. The hazard intensity was randomly assigned an initial magnitude between 0 and 100 and modeled over time with an autoregressive AR(1) model (i.e. the intensity at one time step is equal to the intensity at the previous time step multiplied by a correlation coefficient plus a random fluctuation). Example hazard intensity time series for an iteration with one source (left panels) and five sources (right panels) are shown in the top panels of Fig. 1.

Hazard intensity was calculated at each non-source location in the facility (100 × 100 = 10 000 gridded locations spread evenly throughout the facility) during a full workday (a single iteration encompasses 8 h at 10-s resolution yielding 2880 time steps). We assumed that the intensity of the hazard was inversely related to the square of distance, which would result in a drop of 6 intensity units (if the hazard is thought of as noise, dB) as the distance doubled. Intensities from different sources were combined using the following equation (Peterson, 1980):

$$\text{Total Intensity} = 10 \times \log_{10}\left[10\text{^}\left(\frac{\text{intensity}_1}{10} + \frac{\text{intensity}_2}{10} + \ldots\right)\right] \quad (1)$$

where $\text{intensity}_{1,2,\ldots}$ are the location-specific intensities from source 1, 2, …, and Total Intensity is the resulting summed intensity. We also assumed that the intensity at any location was in instantaneous steady state with the emissions from the sources. This assumption is reasonable for noise and ionizing radiation, but less so for other agents (e.g. dust) that take time to transport or diffuse through the space and can be dramatically influenced by air currents. To further simplify the modeling effort, we assumed that the hazard propagated equally in all directions and neglected convection, reflection, or refraction of the hazard.

The TWA hazard map was constructed by averaging hazard intensities over the 8-h time period at each of the 10 000 simulated locations in the facility (lower panels, Fig. 1). Since intensity was calculated at all locations, spatial interpolation was not needed to determine the TWA hazard map. The TWA hazard map served as the 'truth', and all estimated hazard maps generated using subsets of the intensities were compared against it for both overall comparability and peak exposures.

Modeling of roving monitoring and estimated hazard maps—The simulated intensity data were used to develop 'sampled' hazard maps assuming that operators followed different sampling strategies for roving monitoring. We implemented a regular grid of sampling locations as being easy to implement by a practicing IH. Sampling locations were selected to provide good coverage throughout the facility and were held constant for all iterations, without regard for the location of the sources.

For each iteration, roving monitoring of the hazard at predefined locations was simulated along a path followed by up to three operators, each with a DRI. The first operator began in one corner of the room (at a randomly selected time during the first 3 h of the day), and the 10-s intensity at the first location was sampled. For this analysis, it was assumed that all hazard intensities were measured without error. This assumption was relaxed for sensitivity studies described in the Supplementary Material, available at *Annals of Work Exposures and Health* online. Then, the operator moved to the next predetermined location, recording the hazard intensity at the next time step (10 s later), and so on. A second operator began in the opposite corner of the facility at a randomly selected time during the first 4.5 h of the day and a third operator began in the same corner as the first operator, but at a randomly selected time during the first 6.5 h of the day. The times during which each operator could begin sampling were selected arbitrarily to encourage data collection throughout the day, but more than one operator could be collecting data simultaneously. The operators sampled at high spatial resolution for these sampling campaigns (441 locations evenly spread throughout the facility), so each operator takes 73.8 min to complete the sampling of the entire facility.

From the data sampled by the three operators, we defined six sampling strategies using subsets of the sampled data: (i) high spatial resolution using data from all three operators resulting in three repeated measurements at each of 441 sampling locations; (ii) high spatial resolution with data from the first operator resulting in a single measurement at each of 441 sampling locations; (iii) mid spatial resolution data from three operators resulting in three repeated measurements at each of 121 sampling locations; (iv) mid spatial resolution using data from the first operator resulting in a single measurement at each of 121 sampling locations; (v) low spatial resolution data from three operators resulting in three repeated measurements at each of 36 sampling locations; and (vi) low spatial resolution using data from the first operator resulting in a single measurement at each of 36 sampling locations. The 121 and 36 sampling locations were a subset of the 441 sampled locations evenly spaced throughout the facility. At the 121 and 36 sampling locations for the mid and low spatial resolution cases, measurements at those locations were collected at the same time steps as for the high spatial resolution case, and the temporal lag between locations is presumed to be the time it would take to move between locations. These strategies allowed us to evaluate the importance of sampling spatial resolution and replicate measurements

on sampled map accuracy. The importance of averaging over multiple measurements per location was evaluated by comparing estimated hazard map accuracy using data only from the first operator versus averaging over data collected by all three operators for each location. Sampling strategy pairs 2 and 3 and 4 and 5 result in approximately the same number of measurements used for mapping.

Estimated hazard maps were produced from the sampled intensities using ordinary kriging. Kriging is a geostatistical method for optimal interpolation and has been discussed for the purpose of hazard mapping previously (Koehler and Volckens, 2011; Koehler and Peters, 2013; Lake *et al.*, 2015). Because it would be impractical to manually examine the variogram used for kriging at every iteration, we used 10 bins (as determined by evaluating a subset of maps) and an exponential variogram, as recommended by Koehler and Peters (2013). Additionally, the R^2 value of the variogram model fit to the experimental variogram was monitored for each iteration and sampling strategy. If the R^2 value was <0.3 for Sampling Strategy 1 with 441 locations and 3 replicate measurements per location, then that iteration was discarded and the iteration was repeated. For each iteration, an estimated hazard map was generated using data from the six sampling strategies via kriging, and intensities were estimated at all 100 × 100 locations for comparison with the TWA hazard intensities.

Comparison of estimated hazard maps to the TWA hazard map—Differences between the estimated hazard maps (for each of the four sampling strategies) and the TWA hazard map were quantified using three metrics. First, a Pseudo – R^2 was calculated as one minus the mean absolute difference in hazard intensity between the estimated hazard map and the TWA hazard map at all 100 × 100 locations divided by the mean absolute difference between the TWA hazard map values and the mean TWA intensity over all locations:

$$\text{Pseudo} - R^2 = 1 - \frac{\left(\frac{\sum_{i=1}^{N} \left| \text{Intensity}_{\text{krige},i} - \text{Intensity}_{\text{TWA},i} \right|}{N}\right)}{V_{\text{TWA}}} \tag{2}$$

where $\text{Intensity}_{\text{krige},i}$ is the intensity estimated from the kriged map at location i, $\text{Intensity}_{\text{TWA},i}$ is the intensity from the TWA hazard map at location i, and N is the total number of locations. V_{TWA} is a measure of the mean variation in the TWA hazard map:

$$V_{\text{TWA}} = \frac{\sum_{i=1}^{N} \left| \text{Intensity}_{\text{TWA},i} - \overline{\text{Intensity}} \right|}{N} \tag{3}$$

where $\overline{\text{Intensity}}$ is the mean intensity over all locations and other variables are as defined for equation (2). The pseudo – R^2 is an estimator of overall difference between maps, and a larger value indicates a better agreement between the sampled and TWA hazard maps. The normalization by the mean absolute difference in the denominator allows comparison between cases with one, three, or five sources. Second, the error rate was calculated as the fraction of locations for which the values from the estimated hazard map differed from the TWA hazard map by more than a threshold percentage. In this simulation, we considered 5, 10, 20, and 50% threshold values. The error rate is an indication of the proportion of the

facility for which the short-duration sampled data provided a poor estimate of the TWA. Third, the percent error between the estimated hazard map estimate of intensity and the TWA intensity at each source location was calculated. This metric provides an indicator of how well the estimated hazard map represents the peak TWA intensities near source locations.

We also wish to evaluate how map accuracy varies at facilities with different levels of temporal variability in hazard intensity. For each case (one, three, or five sources), the temporal variability was different for each source and each iteration. Following Lake *et al.* (2015), the temporal variability was characterized by the coefficient of variation of the intensity for the source with the highest temporal variability (hereafter, CV_s). Although temporal variability can be expressed in a variety of ways, ultimately, we chose CV_s because a practicing IH can potentially measure it directly to inform the sampling strategy prior to roving monitoring (see Discussion for more details). The CV_s was calculated for each iteration (for cases with three or five sources, the maximum CV_s value was used), and CV_s values for each case and sampling strategy were sorted into seven bins with an approximately equal number of iterations per bin.

Comparison of accuracy in simulated maps to real-world maps—We compared the error rates from the simulations described above to those measured in two real-world facilities, an engine testing facility (ETF) and a plastics-manufacturing facility (PMF). As reported by Lake *et al.* (2015), each facility had one predominant noise source. Noise intensities were measured at high spatial (~0.1 measurements m^{-2} at the ETF and ~0.2 measurements m^{-2} at the PMF) and temporal resolution (noise intensities averaged to 10-s intervals to match the simulation) in each facility by static and roving monitoring. The noise intensity data from the 18 static monitors were used to generate TWA hazard maps that we used as 'truth', although we acknowledge that the 'true TWA' cannot be known in any real-world setting due to limited resources. The CV_s for the ETF and PMF were estimated as the largest CV value among the static monitors using the noise intensity data averaged to 10-s resolution.

Hazard maps were estimated first using the full spatial resolution of the roving monitoring data (28 or 74 locations at the ETF; 388–394 locations at the PMF, depending on sampling day). To evaluate the influence of spatial sampling resolution on the error rate, subsets of the roving monitoring data from a single operator were randomly selected (varying from full resolution of 0.1 or 0.2 measurements m^{-2} at the ETF and PMF, respectively, down to 0.01 measurements m^{-2}; 5, 10, 20, 35, or 50 locations at the ETF on Day 1; 5, 10, or 20 locations at the ETF on Day 2; and 20, 50, 100, 200, or 300 locations at the PMF on Days 1–3). For each spatial resolution and at each facility, estimated hazard maps were produced from randomly selected subsets of data 100 times and the error rates between the sampled and TWA hazard maps were calculated. To evaluate the influence of replicate measurements per location, estimated hazard maps were generated using data from a single operator or using data from two or three replicate measurements per roving monitoring location, as available. The median error rate was reported among the 100 replicate error rates.

Given the high spatial resolutions available from the PMF and ETF datasets, higher spatial resolution sampling strategies were also added to the previously described simulation. Hazard concentrations were generated as described above, but the sampling strategies were altered to produce 2601, 1301, 651, 434, 326, 121, 61, or 36 sampling locations in the 100 × 100 facility. We only considered the case with one source and sampling conducted with one or three operators, as described for the simulation above and to match the conditions in the field sites. Because increased spatial resolution requires additional time to conduct simulated roving monitoring throughout the facility, the three passes through the facility were each started at a random time in the first 45 min of the day. At the highest spatial resolution, 433.5 min were required to complete roving sampling and all reduced spatial resolutions are subsets of highest resolution case, as described above. To compare to the real-world data, we assumed that the square facility from the simulation was 100 × 100 m to determine the number of locations per square meter.

All simulations and calculations were completed in Matlab version R2013a (Mathworks, Natick, MA, USA). Code used to produce the simulation is included in the Supplementary Material, available at *Annals of Work Exposures and Health* online.

Results

Simulation

Map comparability metrics were compiled for the cases with one, three, and five sources and for the six sampling strategies in Table 1 (median and 25th to 75th percentile for all 500 iterations, without regard for temporal variability, are shown). The differences in values of the map comparison metrics among the different iterations were substantial for all sampling strategies, map comparison metrics, and number of sources.

Map comparability was first evaluated using a pseudo – R^2, which is analogous to the traditional R^2. The pseudo – R^2 can be 1 or smaller (including negative values) with values closer to 1 indicating good agreement between the estimated hazard map and the TWA hazard map and smaller values indicating poorer agreement. The pseudo – R^2 did not change substantially as the number of sources increased (Table 1). There was little difference between high-, mid- and low-resolution sampling strategies, but the median pseudo – R^2 was increased by approximately a factor of 3–4 when using data from three operators compared to just the data from the first operator. The interquartile range of pseudo – R^2 was also larger when only data from a single operator was used for the mapping compared to three operators.

Similar to the pseudo – R^2, the 20% error rates (i.e. the fraction of locations where the intensity predicted with the estimated hazard map was at least 20% different than the TWA hazard map; smaller values indicate better agreement between the sampled and TWA hazard maps) were similar for high and low spatial resolution sampling strategies, but the use of data from three operators, as opposed to a single operator, substantially reduced the error rate. When using data from three operators, the median 20% error rate was <2 for high, mid, and low sampling spatial resolutions. Additionally, the interquartile ranges

were higher when only data from a single operator was used compared to using data from three operators. However, in contrast to pseudo – R^2, the error rates were generally highest for the case with only one source in the facility and the interquartile ranges were larger, compared to the case with five sources when sampling with one operator at high or low spatial resolution. This finding suggests that for a facility with more sources, although the overall map accuracy throughout the facility is similar, fewer locations have errors exceeding the 20% threshold.

Lastly, we assessed map similarity by calculating the underestimation of the peak as the percentage error of the intensity at each source location from the estimated hazard map and the TWA hazard map. It was more likely that hazard intensities at the source locations were underestimated and percent underestimations are shown in Table 1. To reduce underestimation of source (i.e. peak) intensity, higher sampling spatial resolution was more important than using data from three operators. This was likely because using a higher sampling spatial resolution meant that a sample was more likely to be taken close to the source location. The underestimation was approximately a factor of 3 worse when low spatial resolution sampling strategies were used compared to the high spatial resolution sampling strategies. Underestimation was also larger for cases with one source compared to three or five sources. Minimal differences in peak intensity were observed when using data from one or three operators.

We also found that the map accuracy was strongly impacted by the temporal variability within the facility. Temporal variability will depend on the processes at a facility, and likely by hazard type. The pseudo R^2, as a function of the CV_S, for four of the sampling strategies (high and low spatial resolution) are shown in Fig. 2. When viewed in this way, it is clear that generating estimated hazard maps using the data from three operators compared to a single operator was far more effective at increasing the pseudo – R^2 than increasing the sampling spatial resolution. In fact it is hard to distinguish the difference in median pseudo – R^2, as a function of CV_S, between the low and high spatial resolution cases (dashed and solid lines fall nearly on top of one another, especially at higher CV_S values). The influence of the temporal variability in the facility resulted in decreasing pseudo – R^2 with increasing CV_S. The variability in pseudo – R^2 among different iterations generally increased as CV_S increased, was larger when only data from a single operator was used for mapping (gray lines), and was higher with a smaller number of sources in the facility. Interestingly, with more sources, the impact of the CV_S is smaller than for only a single source in the facility. The reduced impact of CV_S on the pseudo – R^2 may result from smoother map surfaces when more sources are present. For each of the 500 iterations within the cases with one or five sources, we calculated a normalized TWA concentration values in which the TWA value at each location was divided by the mean TWA concentration over the entire map (i.e. a normalized TWA concentration value of 1 indicates that the concentration was equal to the map mean value; values that diverge from 1 indicate that higher map variability). Compiled normalized TWA concentrations over all 10 000 locations for 500 iterations resulted in a broader distribution of values for the case with one source than the case with five sources. This suggests that smoother maps result in improved map accuracy.

Sampling strategy pairs 2 and 3 and 4 and 5 result in approximately the same number of measurements used for mapping and allow a more direct comparison of whether resources are better utilized when sampling a single time at more locations or three times at fewer locations. The pseudo – R^2, as a function of CV_s, is shown for Sampling Strategies 4 and 5 in Fig. 3. At facilities with low temporal variability, the difference between the two sampling strategies is minimal; however, as the temporal variability increased, the quality of the maps based on data collected at the mid spatial resolution with only a single measurement per location deteriorated more quickly than maps made with data collected at low spatial resolution with three replicate measurements per location. Additionally, the variability among the iterations was higher for the mid spatial resolution sampling strategy with only one measurement than the low spatial resolution sampling strategy with three replicate measurements per location.

Results for the 20% error rate are qualitatively similar as for the pseudo – R^2; lower spatial resolution with replicate measures was a better use of resources than increased spatial resolution without replicate measures. Intensity underestimation at source locations was only a weak function of CV_s, but the variability in the underestimation of the peak tended to increase with temporal variability and was stronger for the case with one source. Details on the 20% error rate and peak underestimation as a function of CV_s can be found in the Supplementary Material, available at *Annals of Work Exposures and Health* online.

Comparison of accuracy in simulated maps to real-world maps

Next, we compared the simulation results to map accuracy from two real-world hazard mapping data sets (Lake *et al.*, 2015). From the simulation, the median 10% error rate (i.e. the percentage of locations with predicted values from the estimated hazard map that were >10% different than the TWA hazard map) at seven levels of spatial resolution and as a function of CV_s was plotted as contours in Fig. 4, with darker colors indicating higher error rates. Using data from a single operator resulted in error rates plotted in Fig. 4a and using data from three operators resulted in error rates plotted in Fig. 4b. The stronger influence of CV_s compared to sampling resolution on the error rate is apparent in both Fig. 4a and Fig. 4b. Furthermore, the values in Fig. 4a are larger than in Fig. 4b, showing the reduction in error rate when replicate measurements are used to create the hazard map.

Overlain on Fig. 4 are points corresponding to the error rates at the two facilities described by Lake *et al.* (2015). The color of the circle represents the 10% error rate on the same scale as used for the contours of the error rates from the simulation. Error rates from the ETF are shown in circles, and error rates calculated using data from the PMF are shown in squares. Only one measurement per location was used in Fig. 4a. To compare the importance of sampling spatial resolution, subsets of the roving monitoring data were used (in each case, the CV_s remains the same, resulting in vertically aligned points on the plot). Averaging over two to three replicate measurements per location (as available) was compared to simulations in Fig. 4b, and the influence of sampling resolution was estimated again by using subsets of the roving monitoring data. Although the absolute values between the 10% error rates estimated from the real-world data sets and the simulated data are not the same, generally the trends are similar. As CV_s increases, the error rate generally increases. As the spatial

resolution is reduced, the error rate increases somewhat. Replicate measurements were also found to generally reduce the 10% error rate when hazard maps were produced from both real data and simulated data.

Discussion

A priori we expected that the sampling strategy with high spatial resolution and that used data from three operators would provide the best estimate of the TWA hazard map. Conversely, we expected the sampling strategy using low spatial resolution and that used data from a single operator would give the worse comparison with the TWA hazard map. The simulation study generally confirmed these results, yet showed that for overall map accuracy (as represented by the pseudo $-R^2$ or the error rate), the influence of replicate measurements were more important than an increase in sampling spatial resolution. Sampling strategies 1 and 2 are at very high spatial resolution requiring substantial resources and may represent a number of measurements not feasible in some environments.

Although overall map accuracy should be a goal, it may also be important to practicing IH staff to know whether the hazard maps accurately represented ‘hot spots’, locations where exposures may be highest. Underestimating peak TWA exposures may lead to a failure to act on high exposures. On the other hand, overestimating peak TWA exposures can lead to implementation of control strategies that do not result in desired reductions in personal exposures. If sampling locations were not near the source location, interpolation methods are generally unable to estimate peaks higher than measured intensities. Thus, to accurately capture the peak hazard intensities, measurements must be taken in close proximity to sources.

Temporal variability, represented in this study by CV_s, was a strong determinant of overall map accuracy. Estimates of temporal variability, sampling spatial resolutions, and number of replicates used for hazard mapping data collection at different facilities were compiled from the available literature in Table 2. Temporal CV_s values ranged from 0.02 to 0.96 when static monitoring was conducted (which was rare), indicating that high levels of temporal variability impacting the representativeness of hazard maps are likely encountered in occupational environments. Sampling spatial resolutions used in previous studies tended to be in the range considered in the simulation study (0.004–0.04 sampling locations per square meter, assuming a 100 × 100 m facility). One notable exception was Lake *et al.* (2015), where higher spatial resolutions were used for roving monitoring. An important limitation of many existing studies is the lack of replicate measurements. Four of the nine studies only collected measurements once per location per day.

Recommendations for IHs

The results from this simulation provide important insight that can guide IHs on how to design a sampling strategy to improve hazard mapping (Fig. 5). First, we recommend that the CV_s be determined for the facility. An IH can use a DRI for the hazard of interest and leave it for at least 2 h, but preferably for a full work shift, near sources likely to have variable emissions. It may be necessary to evaluate the temporal variability in hazard intensity near several sources to get a good estimate of the CV_s. For $CV_s > 0.5$ (e.g.

evaluated in the case with five sources), hazard mapping may not be a good strategy to evaluate TWA exposures and personal monitoring should be considered more appropriate for exposure assessment. Our simulation suggests that for facilities with temporal variability $CV_S > 0.5$, 20% or larger biases in the TWA estimates may be observed throughout the facility (Supplementary Fig. S1, available at *Annals of Work Exposures and Health* online) and that pseudo – R^2 values tend to be negative, regardless of the number of sources or sampling strategy (Fig. 2).

If the CV_S is <0.5, hazard mapping using short-duration measurements is a reasonable approach and an IH should consider the resources available. Important questions to consider are as follows: how many DRIs are available; how many people can contribute to roving monitoring data collection; how much time is available; and how large is the facility? At facilities with low temporal variability ($CV_S < 0.1$), we observed that differences in sampling strategies had little impact on pseudo – R^2 or the error rate. However, care should be taken to provide adequate spatial resolution around known sources to accurately capture the peak intensities. Since sound intensity, like many contaminants, decays following an inverse-square law, measures can be substantially underestimated as distance from the source increases. If the CV_S is >0.1, but <0.5, IHs should ensure that more replicate measurements can be collected to counteract increased uncertainty from the high temporal variability of sources. Our simulation suggests that for facilities with $0.1 < CV_S < 0.5$, error rates increase sharply and pseudo – R^2 values begin to drop, particularly in more complex facilities with more than one source and for sampling strategies that only use a single measurement per location. Thus, priority should be placed on replicate measurements over increased spatial resolution of measurements. To ensure that hazard intensity near sources are also representative, we recommend irregular sampling strategies that are less resolved away from known sources and more highly resolved near sources. Such strategies are likely to maximize resources, ensuring overall map accuracy, as well as information on peak intensities, provided that there are resources available to obtain repeated measures at each sampling location.

Limitations and future work

By not including internal structures or reflection off surfaces, the spatial dependence in these simulations was likely higher than in real facilities and may amplify the conclusion that relatively few measurements are required in space. The requirement that the variogram model fit to the experimental variogram had an R^2 value > 0.3 for Sampling Strategy 1 may also have led to simulations with stronger spatial correlation than may be encountered in real facilities. Complex facilities with many sources and/or internal structures may require more spatial resolution to adequately represent the spatial correlation in the data. We summed intensities from different sources according to equation (1), which is appropriate for noise, but not applicable for other hazards. Furthermore, we assumed that the intensity at any location was in instantaneous steady state with the emissions from the sources, which may be poor for agents that are affected by ventilation, such as gases or particles. In the absence of ventilation, however, gases and fine-mode particles (with negligible settling and Brownian diffusion), like sound intensity, should follow an inverse-square law.

Our ‘real-world’ data are insufficient to validate simulations for facilities with high temporal variability ($CV_s > 0.5$). It is also not clear whether the agreement between the simulation and ‘real-world’ data may depend on the hazard type. Additionally, we did not investigate how different operator motions might impact our results. The prescribed motion of the operators in the simulation may help explain some of the observed differences between the real-world and simulated hazard map accuracies. The simulation also did not account for ventilation rates in buildings, which can lead to localized hazard intensity gradients away from known sources. Future work should address these limitations.

Finally, future work should consider how such simulations or real-world hazard mapping data could be used to determine efficient site placement for longer duration sampling strategies using static monitors. With the decreasing cost of sensors for various hazards available on the market (Koehler and Peters, 2015), long-term monitoring of hazards with semi-permanent static monitors will allow for comprehensive exposure assessment. Such a comprehensive assessment would account for the spatial and temporal variability in exposures over long periods of time, minimizing or eliminating the need for time-intensive roving monitoring strategies.

Conclusions

A simulation study was conducted to evaluate the most efficient sampling methodologies when using short-duration measurements to produce estimated hazard maps that will best represent the TWA hazard map. In general, an increase in the number of replicates is preferable to an increase in sampling spatial resolution to improve overall accuracy. However, to accurately capture peak exposures, higher spatial resolution near sources was needed. We recommend that IHs conduct a preliminary assessment of the temporal variability in their workplace and identify variable sources of the hazard and design a sampling strategy that offers enough spatial resolution and replicate measures to minimize bias between short-duration measurements and the TWA. An irregular grid may provide the best solution, giving resolution near sources, and allowing for replicate measurements to be completed in a relatively short period of time.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Declaration

Funding for this project was provided in part by grant #R01OH010533 from the National Institute for Occupational Safety and Health (NIOSH). The authors declare no conflict of interest relating to the material presented in this article. Its contents, including any opinions and/or conclusions expressed, are solely those of the authors.

References

Dasch J, D’Arcy J, Gundrum A et al. (2005) Characterization of fine particles from machining in automotive plants. J Occup Environ Hyg; 2: 609–25. [PubMed: 16282165]

Evans DE, Heitbrink WA, Slavin TJ et al. (2008) Ultrafine and respirable particles in an automotive grey iron foundry. Ann Occup Hyg; 52: 9–21. [PubMed: 18056626]

Heitbrink WA, Evans DE, Peters TM et al. (2007) Characterization and mapping of very fine particles in an engine machining and assembly facility. J Occup Environ Hyg; 4: 341–51. [PubMed: 17454502]

Koehler KA, Peters TM. (2013) Influence of analysis methods on interpretation of hazard maps. Ann Occup Hyg; 57: 558–70. [PubMed: 23258453]

Koehler KA, Peters TM. (2015) New methods for personal exposure monitoring for airborne particles. Curr Environ Health Rep; 2: 399–411. [PubMed: 26385477]

Koehler KA, Volckens J. (2011) Prospects and pitfalls of occupational hazard mapping: 'between these lines there be dragons'. Ann Occup Hyg; 55: 829–40. [PubMed: 21917819]

Lake K, Zhu J, Wang H et al. (2015) Effects of data sparsity and spatiotemporal variability on hazard maps of workplace noise. J Occup Environ Hyg; 12: 256–65. [PubMed: 25437137]

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: 947–52. [PubMed: 14555448]

Park JY, Ramachandran G, Raynor PC et al. (2010) Determination of particle concentration rankings by spatial mapping of particle surface area, number, and mass concentrations in a restaurant and a die casting plant. J Occup Environ Hyg; 7: 466–76. [PubMed: 20526949]

Peters TM, Anthony TR, Taylor C et al. (2012) Distribution of particle and gas concentrations in swine gestation confined animal feeding operations. Ann Occup Hyg; 56: 1080–90. [PubMed: 22904211]

Peters TM, Heitbrink WA, Evans DE et al. (2006) The mapping of fine and ultrafine particle concentrations in an engine machining and assembly facility. Ann Occup Hyg; 50: 249–57. [PubMed: 16361396]

Peterson APG. (1980) *Handbook of noise measurement*. 9th edn. Concord, MA: GenRad Inc.

Vosburgh DJ, Boysen DA, Oleson JJ et al. (2011) Airborne nanoparticle concentrations in the manufacturing of polytetrafluoroethylene (PTFE) apparel. J Occup Environ Hyg; 8: 139–46. [PubMed: 21347955]

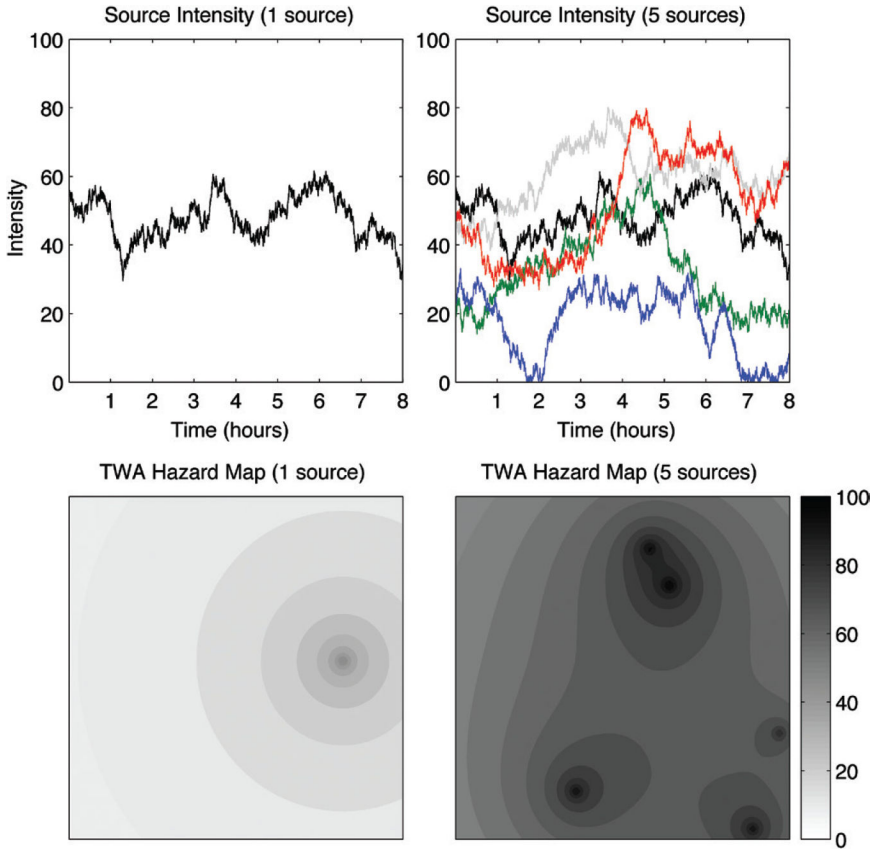

Figure 1.

Example simulated hazard intensities at a theoretical facility with one source (left panels) or five sources (right panels). Top panels show the time series of hazard intensities for one or five sources over an 8-h work shift. Bottom panels show the TWA hazard maps for the facility.

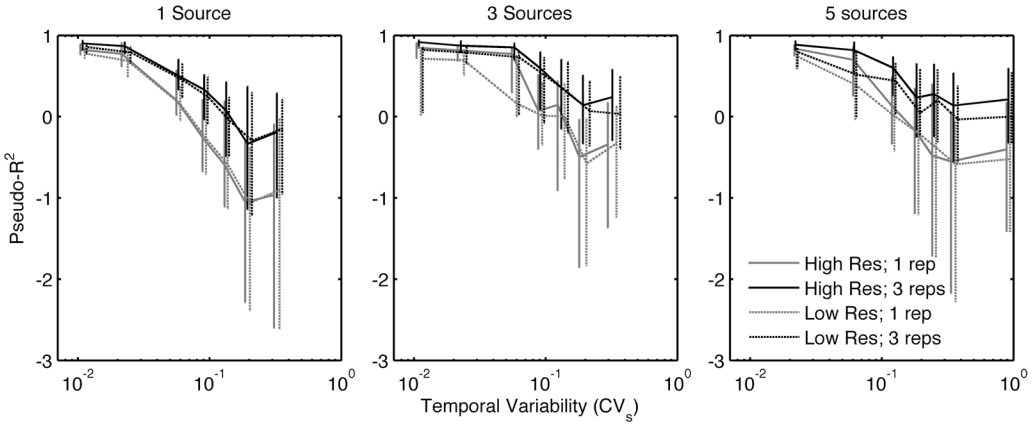

Figure 2.

The median pseudo – R^2 for the estimated hazard map compared to the TWA hazard map for four different sampling strategies: high or low spatial resolution with one or three replicate measurements per location, as labeled. The left panel shows all iterations for the case with one source, as a function of CV_s. The middle panel for the case with three sources, and the right panel for the case with five sources. Error bars represent the 25th and 75th percentile of pseudo-R^2 values and are slightly jittered for clarity.

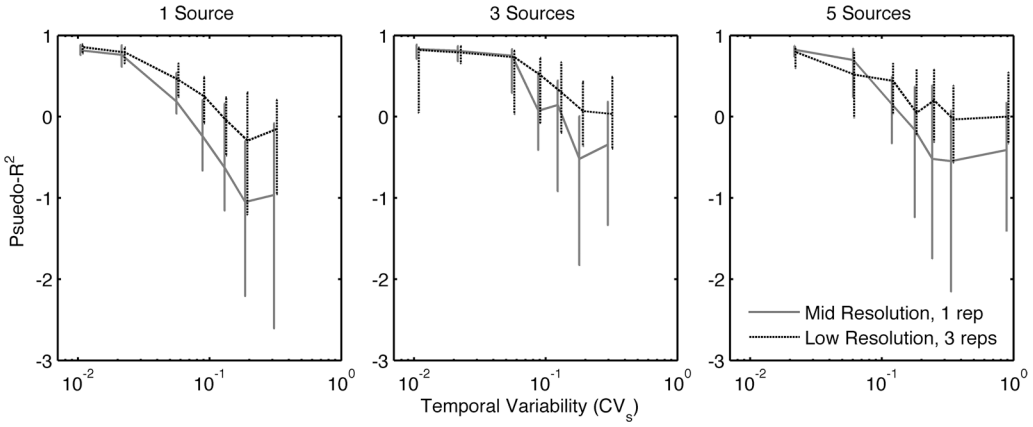

Figure 3.

The median pseudo – R^2 for the estimated hazard map and the TWA hazard map for two sampling strategies with approximately the same number of measurements, as labeled. The left panel shows all iterations for the case with one source, as a function of CV_S. The middle panel for the case with three sources, and the right panel for the case with five sources. Error bars represent the 25th and 75th percentile of pseudo-R^2 values and are slightly jittered for clarity.

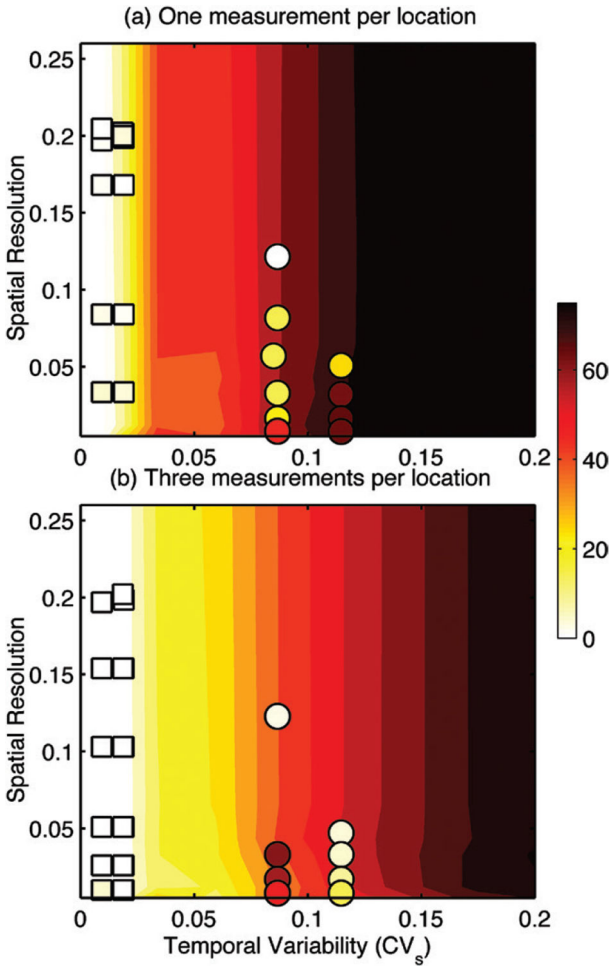

Figure 4.

Contour plot of the simulated median 10% error rate using data from a single operator (a) or using the data from three operators (b). Overlain are 10% error rates from ETF (circles) and PMF (squares) using one measurement per location (a) or two to three measurements per location (b).

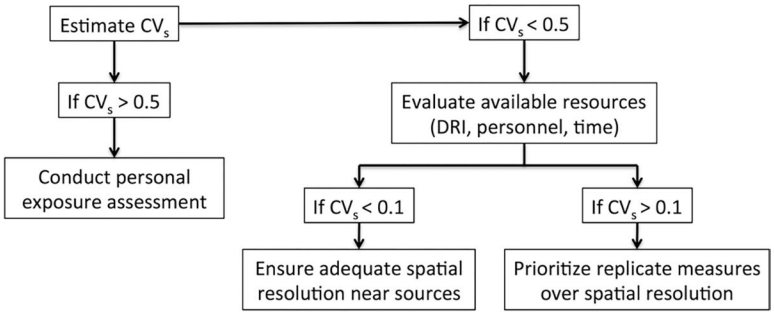

Figure 5.
Flow diagram for conducting hazard mapping.

Table 1.

Map comparison metrics [median (25th, 75th) quantiles] by sampling strategy averaged over all CV_s values for the three cases with one, three, or five sources (500 iterations per case).

	Sampling strategy					
	1	2	3	4	5	6
Number of sources	**High resolution[a], 3 operators (*n* = 1323)**	**High resolution[a], 1 operator {*n* = 441)**	**Mid resolution[b], 3 operator (*n* = 363)**	**Mid resolution[b],1 operator (*n* = 121)**	**Low resolution[c] operators (*n* = 108)**	**Low resolution[c] operator (*n* = 36)**
			Pseudo-R^2			
1 source	0.43 (−0.12,0.81)	0.12 (−0.75,0.68)	0.42 (−0.14, 0.79)	0.13 (−0.79, 0.67)	0.31 (−0.15,0.70)	0.02 (−0.79,0.58)
3 sources	0.52 (−0.01,0.83)	0.13 (−0.70,0.71)	0.53 (0.00,0.81)	0.16 (−0.73, 0.70)	0.36 (−0.10,0.73)	0.00 (−0.66,0.56)
5 sources	0.53 (0.01,0.81)	0.14 (−0.86,0.70)	0.53 (−0.01,0.80)	0.15 (−0.87, 0.69)	0.31 (−0.14,0.71)	−0.04 (−0.83,0.59)
			Error rate (percentage of locations with error > 20%)			
1 source	0 (0, 72)	30 (0, 88)	0 (0, 73)	30 (0, 87)	2 (0, 74)	30 (0, 88)
3 sources	0 (0, 3)	1 (0, 47)	0 (0,3)	0 (0,48)	0 (0,12)	4 (0, 50)
5 sources	0 (0, 0)	0 (0,15)	0 (0,0)	0 (0,16)	0 (0,1)	0 (0,18)
			Underestimation of peak intensity at source locations (percent)			
1 source	8 (3,14)	8 (1,16)	16 (11,26)	17 (10,28)	27 (20, 43)	30 (20, 45)
3 sources	5 (1,11)	5 (0,13)	10 (6,18)	11 (5,21)	19 (12, 28)	19 (11, 32)
5 sources	4 (0, 8)	4 (0,11)	7 (3,12)	8 (3,15)	13 (7, 20)	14 (7, 23)

The total number of measurements included in the sampling strategy is included in each column heading.

[a] 441 sampled locations in the facility; 0.04 points m^{-2} assuming the facility is 100 × 100 m.

[b] 121 sampled locations in the facility; 0.01 points m^{-2} assuming the facility is 100 × 100 m.

[c] 36 sampled locations in the facility; 0.004 points m^{-2} assuming the facility is 100 × 100 m.

Table 2.

Spatial resolution, number of replicate measurements per sampling location, and temporal variability at different facilities cited in the literature.

Study	Facility type	Spatial resolution (m^{-2})	Number of replicates (number per location per day/ number of days of sampling)	Maximum temporal CV (contaminant)
O'Brien (2003)	Machining plant	— (65 locations, facility size unknown)	1/2	—
Dasch *et al.* (2005)	Automotive plants	0.002–0.03	1/1 (per facility)	—
Peters *et al.* (2006)	Engine machining and assembly	0.0005–0.002	1/5	—
Heitbrink *et al.* (2007)	Engine machining and assembly	0.0005–0.002	2–3/5	—
Evans *et al.* (2008)	Automotive gray iron foundry	0.001–0.005	1–3/5	—
Park *et al.* (2010)	Restaurant	0.08	1/3	—
	Die casting plant	0.02	1/3	—
Vosburgh *et al.* (2011)	Apparel manufacturing	0.01	2/1	—
Peters *et al.* (2012)	Swine gestation	0.03	4/3	0.96 (particle mass concentration)
Lake *et al.* (2015)	PMF	0.2	3–4/3	0.02 (noise intensity)
	ETF	0.1	2–3/2	0.11 (noise intensity)