



OPEN Optimizing radiation monitoring networks to improve emergency response strategies during nuclear power plant accidents

Allan Gellenoncourt^{1✉}, Ali Ayoub² & Haruko M. Wainwright²

This paper presents a new strategy to optimize radiation monitoring networks for effectively predicting contaminated areas and radiation levels during nuclear power plant accidents in order to improve emergency response efforts. Our strategy addresses variable metrological fields by generating ensemble simulations of wind fields and radionuclide migration in the atmosphere using the WSPEEDI (Worldwide version of System for Prediction of Environmental Emergency Dose Information) simulator. GPCAM (Gaussian Process for Continuous-time Acquisition of Measurements) is then used to capture the heterogeneity of radiation levels by sparse monitoring points, and to optimize their locations. We consider three different scenarios: (a) a single static spatial distribution of the radiation levels, (b) the temporal evolution of the distribution within a single release scenario for mobile sensor deployment, and (c) ensemble optimization with variable metrological conditions for assessing risks and emergency responses at a particular site a priori. The results are compared with the homogeneously-distributed network. Our results show that GPCAM is able to identify effective monitoring locations for each of these scenarios, except that a prevailing wind direction is required for the ensemble case. In addition, we compare the effect of different acquisition functions, kernel functions, and hyperparameters in GPCAM on the sensor locations.

Keywords Sensors network, Dose rate, Network optimization, Machine learning, Gaussian process

Nuclear energy is now considered one of the key technologies for mitigating climate change. According to the International Atomic Energy Agency (IAEA)¹, the world nuclear electricity production is expected to increase by about 29% from the 2021 level by 2030 in the low case projection, and by more than 2.5-fold by 2050 in the high case. However, the revival of the nuclear industry comes with its share of concerns. Significant environmental and health concerns are generally associated with nuclear power plant accidents^{2,3}. There are indeed some high-profile accidents that are still remembered, such as the Chernobyl and Fukushima Daiichi nuclear power plant (NPP) accidents⁴.

Most countries have comprehensive emergency response strategies in case of an NPP accident. In the US, the Nuclear Regulatory Commission (NRC) provides guidance on emergency responses. The NRC report NUREG-0396⁵ introduces the concept of an Emergency Planning Zone (EPZ) as an area within which planning is recommended to enforce swift emergency and evacuation actions in order to protect the public during an NPP accident. EPZ is typically fixed as a 10-mile radius area around the NPP, considering the possible external gamma radiation exposure and inhalation exposure from the cloud of released radioactive gases and deposited material.

During the Fukushima NPP accident, however, evacuation was done successively from a 3 km radius, a 10 km radius, and then a 20 km radius⁴. Cell phone data confirmed that the majority of the people were evacuated before the biggest release and land contamination on Day 5⁶. At the same time, since Japan's radioactive release and dose prediction software, SPEEDI, was not properly used to forecast dose distributions, and the monitoring network was not functioning, the evacuation process resulted in tremendous confusion during the accident, as well as generated distrust toward the government and nuclear industry after the accident⁴. Ayoub et al.⁴ suggested that emergency responses, such as delineating the evacuation zones, should be dynamic and based on radionuclide transport forecasts, which themselves should rely on efficient monitoring systems.

¹Institut National des Sciences et Techniques Nucléaires, Cadarache, France. ²Nuclear Science and Engineering, Massachusetts Institute of Technology, Cambridge, USA. ✉email: hmwainw@mit.edu

To enable such adaptive evacuation strategies, a radiation-monitoring network should be able to effectively capture the extent of contamination and radiation levels during the accident. There have been a number of approaches to the design and optimization of an environmental monitoring network, such as genetic algorithms (GA), spatial simulated annealing (SSA), particle swarm optimization (PSO), ant colony optimization (ACO), or surrogate-based approaches like the Gaussian process (GP) model. (The literature review of these approaches is included in Sun et al.⁹.) In particular, GP has several advantages compared to others for optimizing a monitoring network with built-in uncertainty quantification capabilities. In the past, GP has been used for interpolating sparse measurements or for integrating multiscale, multitype radiation datasets^{7,8}. Sun et al.⁹ has developed the radiation-sensor-placement optimization methodology for capturing the existing contamination using GP. Their approach is based on the “greedy” algorithm, which involves adding one sensor at a time into the region that has the highest interpolation error compared to the reference map based on all the sensors. More recently, GP has been extended from the physical space (i.e., locations) to general parameter spaces to autonomously select parameter values for physics and chemistry experiments^{10,11}.

Despite these advances, there has not been specific development of radiation-monitoring optimization for responding to future NPP accidents. In particular, the approach by Sun et al.⁹ is only for fixed, nondynamic contamination, such as in the aftermath of the Fukushima NPP accident. (i.e., the spatial heterogeneity of the air dose rate has been relatively stable over time due to the limited mobility of Cs-137 in the terrestrial system). There is a need to consider the potential radioactive contamination extent in case of a future nuclear accident and to develop a solution that can support more effective emergency responses. Since the conditions of the accident and the prevailing meteorology cannot be precisely predicted ahead of time, the meteorological dynamic properties and all potential radioactive contaminations should be considered, in order to propose an optimized monitoring network capable of capturing the radionuclide propagation in all situations.

This paper proposes a methodology for optimizing a radiation-sensor network that can effectively capture the spatial distribution of radiation levels during a potential NPP accident. Using this methodology, an ensemble of potential radiation-level distributions is generated using the WSPEEDI (Worldwide version of System for Prediction of Environmental Emergency Dose Information) simulator and the analytical Gaussian plume solutions¹². We then use a machine-learning suite, called GPCAM, that automatically selects measurement parameters iteratively, by exploiting Gaussian process regressions to construct the interpolated distribution of the radiation level (gpcam.lbl.gov). We evaluate different kernel functions that define spatial correlation structures¹³ as well as acquisition functions (which in turn define the criteria for choosing the next sensor location to add in an iterative process) with respect to the interpolation performance.

We demonstrate our approach, considering a hypothetical location and scenarios, within three different settings (Fig. 1): (a) a single static spatial distribution of the radiation levels, (b) the temporal evolution of the distribution within a single release scenario, and (c) ensemble optimization with variable meteorological conditions. In (a), the static distribution is representative of the situation in which the radionuclides from the atmospheric release are settled on the ground and the radiation levels do not change over time (steady state after a release). This static case is first used to understand the effect of different parameters in GPCAM on the sensor locations. In (b), we consider the temporal evolution of the radiation-level distribution such that the spatial and temporal evolution must be monitored. Lastly, the ensemble case (c) represents the creation of an optimized network before the accident, considering the uncertainty in meteorological conditions.

In the static and temporal evolution case (Fig. 1a), the optimization of monitoring networks is a two-step process: interpolation (Step 1.1 within the figure) and new sensor placement (Step 1.2 within the figure). In the interpolation step, an interpolated map of air radiation dose rates is obtained through a GP model. In the second step, a new sensor is placed at the point with the highest merit score considering the results of the interpolation step. These two steps will be repeated until the interpolated map is considered to be sufficient to capture the regional heterogeneity. These steps are automated within GPCAM.

In the ensemble optimization case (Fig. 1b), the network optimization is done for each reference map produced by model simulations, and the overall optimization is done across the ensembles of the reference maps. Once the reference maps are created, the optimization of monitoring networks is a two-step process: (1) interpolation and creation of an optimized network for each training map in similar fashion as for the 2D static cases, and (2) post-processing of the step-1 networks to create one single optimized network able to capture all potential radioactive releases from the NPP.

Results

2D static case studies

First, WSPEEDI was used to generate the reference dose map of the radioactive plume—i.e., an airborne cloud of radioactive gases, particles, and/or vapors released from a nuclear plant—due to the external exposure to the deposited Iodine-131 and Cesium-137, at ground level. We assumed a source term similar to the release from the Fukushima nuclear power plant accident defined as 2.0E19 Bq/h for the Iodine-131 and 4.0E17 Bq/h for the Cesium-137 over the release time of one hour, according to the study by Ayoub et al.¹⁴. Using this map, we analyzed the impact of various kernel and acquisition functions on the performance of the GPCAM.

GPCAM was then used to optimize a series of static sensor locations for the reference map (Fig. 2a), starting from the initial radial distribution of four sensors around the source with a fixed radius, and one on the source location. The reference map is created using the WSPEEDI software based on the meteorological and other factors (e.g., land, surface) at the Massachusetts Institute of Technology (MIT, Cambridge, MA, USA). This follows the same approach as the study by Ayoub et al.¹⁵ (see Sect. 4.2.1). The simulations include the advection and dispersion of radionuclides, radioactive decay, and wet, dry, and fog water deposition on the ground. The study area covers a 600 km × 600 km region, with MIT positioned at its center. The numerical grid size is 2 km, which is considerably smaller than the average distance between sensors for the existing national networks

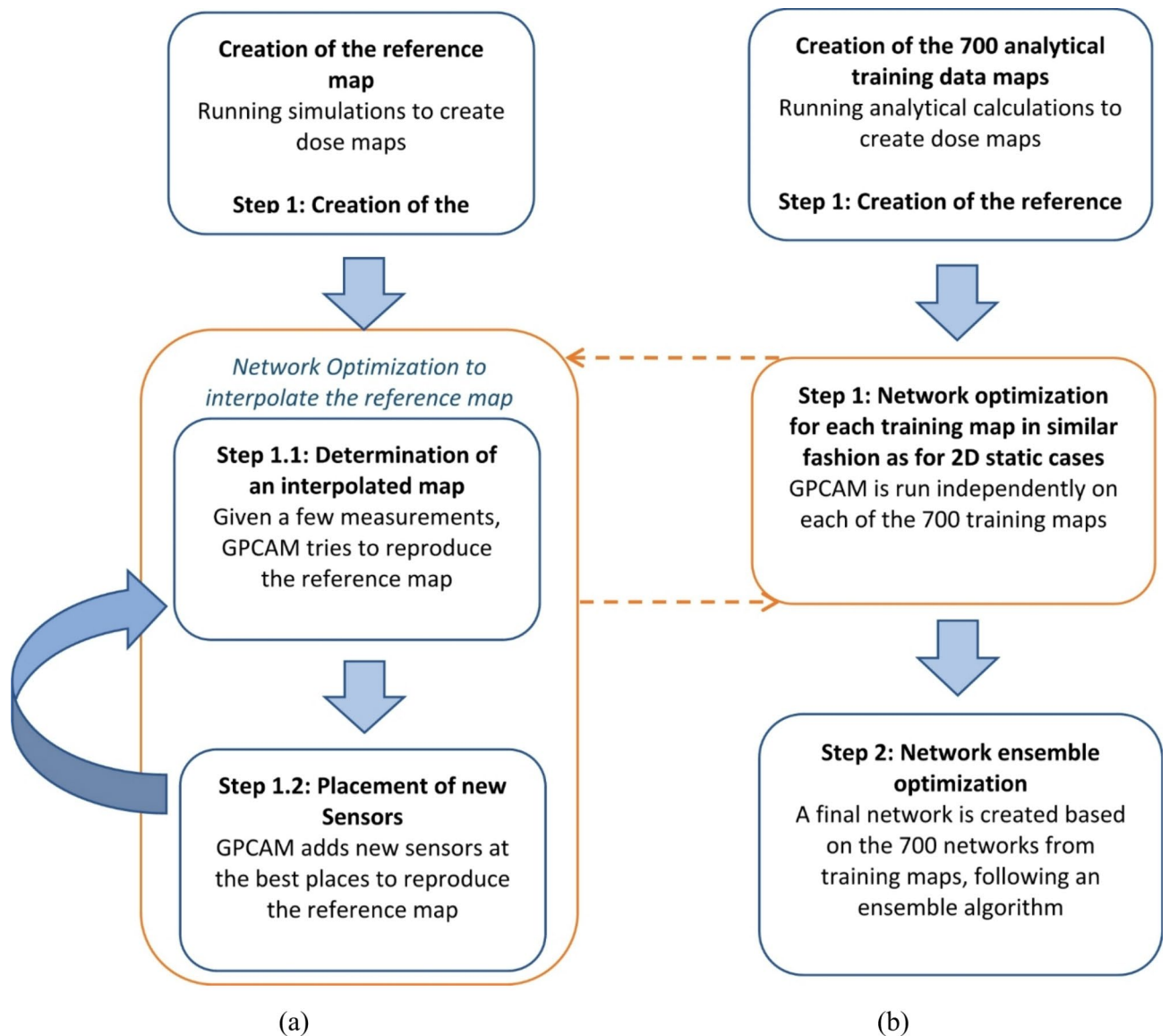


Fig. 1. Optimization network methodology for (a) 2D-static cases and temporal evolution scenarios, (b) ensemble optimization.

such as the Austrian radiation network, and the French OPERA network¹⁶. The resulting sensor locations are distributed mostly within the plume (Fig. 2a/left). The interpolated map based on these sensors (Fig. 2a/middle) shows that the sensors are capturing the plume shape by interpolating the sparse points. The variance field (Fig. 2a/right) shows that the uncertainty in the estimation is smaller close to the measuring points, and lower within the plume than the other region, due to the increased number of sensors within the plume.

As a contrast, we also tested the homogeneously distributed sensor network with the same number of sensors (Fig. 2b/left). The interpolated map (Fig. 2b/middle) appears smoother, particularly near the source, than the GPCAM-based map (Fig. 2a/middle). In addition, the dose rate is underestimated near the source. The estimation variance (Fig. 2b/right) is fairly uniform across the domain, except for the lower variance near the sensor locations.

Different acquisition functions in GPCAM create the distinct patterns of sensor locations (Fig. 2c). The error acquisition function (i.e., placing the sensor at the highest interpolation error location) $g_{error}(x)$ (Fig. 2c / middle) results in the sensors mostly placed near the edge of the plume, where the relative error is large. On the other hand, the Upper Confidence Bound, or UCB function (i.e., placing the sensor at the highest sum of the estimated field and its standard deviation) $g_{UCB}(x)$ (Fig. 2c/right) places the locations primarily within the plume and high dose region, because the UCB function considers the mean value of the estimated field, thereby directing sensors toward the regions with the high radiation intensity. When the mixed acquisition function (i.e., the mix of error and UCB) $g_{mixed}(x)$ (Fig. 2c/left) is used, the sensors are distributed both on the edge and on top of the plume.

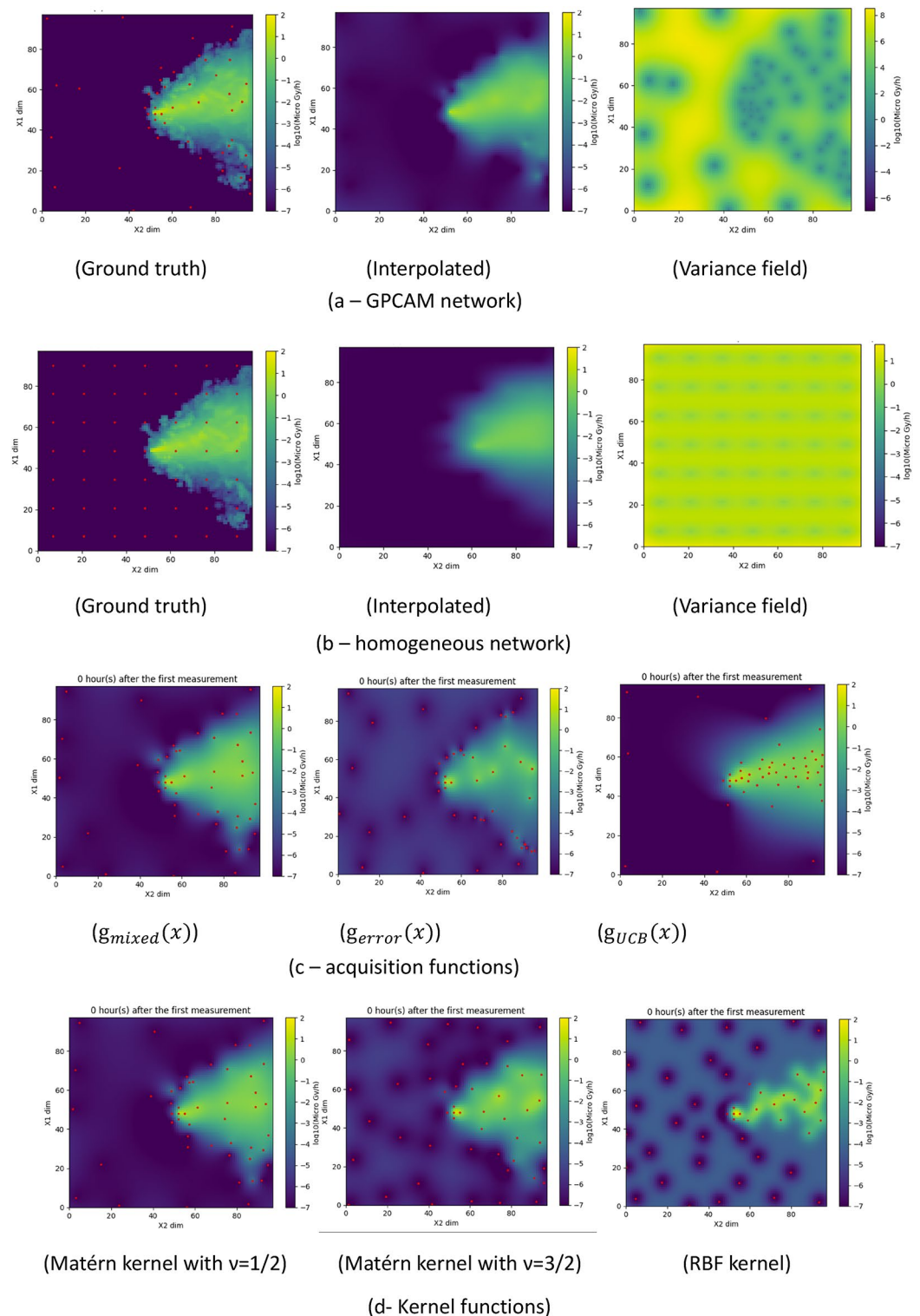


Fig. 2. (a) Air-dose-rate map at 0 m above ground, with the network of sensors created by GPCAM, extrapolated map and variance field, (b) Air-dose-rate map at 0 m above ground with a homogeneous network, extrapolated map, and variance field, (c) sensors network for $g_{mixed}(x)$, $g_{error}(x)$ and $g_{UCB}(x)$ acquisition functions, (d) sensors network for a Matérn kernel with $v = 1/2$, Matérn kernel with $v = 3/2$ and RBF kernel. The hyperparameters bound were fixed to [0.001 km, 20 km] for the variance and [0.1 $\log_{10}(\mu\text{Gy/h})$, 100 $\log_{10}(\mu\text{Gy/h})$] for the spatial scale. Hyperparameters were retrained for every new sensor. A background value of $1e-7 \mu\text{Gy/h}$ was set for all the maps rather than 0 $\mu\text{Gy/h}$, as the logarithmic value of the radiation is taken for the regression, so there cannot be any null value.

The impact of different kernel functions is then evaluated (Fig. 2d). We used the Matérn kernel with varying smoothness parameters (ν) with the mixed acquisition function. These smoothness parameters impact the spatial correlation between data points, such that a smaller ν makes the interpolation rely more heavily on local data values, resulting in the interpolated map that are more sensitive to the local variations. However, the impact of different kernels on the sensor placements is less apparent than the one of different acquisition functions. The exponential kernel ($\nu = 1/2$; Fig. 2d/left) results in more sensors around the edge of the plume, while the RBF kernel (Fig. 2d/right) appears to place more sensors within the plume as well as more uniformly over the domain. The network from the kernel of $\nu = 3/2$ (Fig. 2d/middle) is between these two, having more sensors on the edge of the plume than the RBF kernel and having sensors more uniformly distributed than $\nu = 1/2$.

We also compare the performance based on the overall R^2 score and the RMSE (the Root Mean Squared Error) evolution as a function of the number of sensors. The GPCAM-based sensor network performs better than the *homogeneous network* constantly across the number of sensors: the R^2 score is up to 20% higher and RMSE score up to 50% lower (Fig. 3a, b). Similarly, the mixed acquisition function has a consistently higher R^2 score up to 30%, and lower RMSE up to 40%, than the UCB and error acquisition functions (Fig. 3c, d). GPCAM's performance tends to converge around 25 to 30 sensors. This result suggests a minimum of 25 sensors

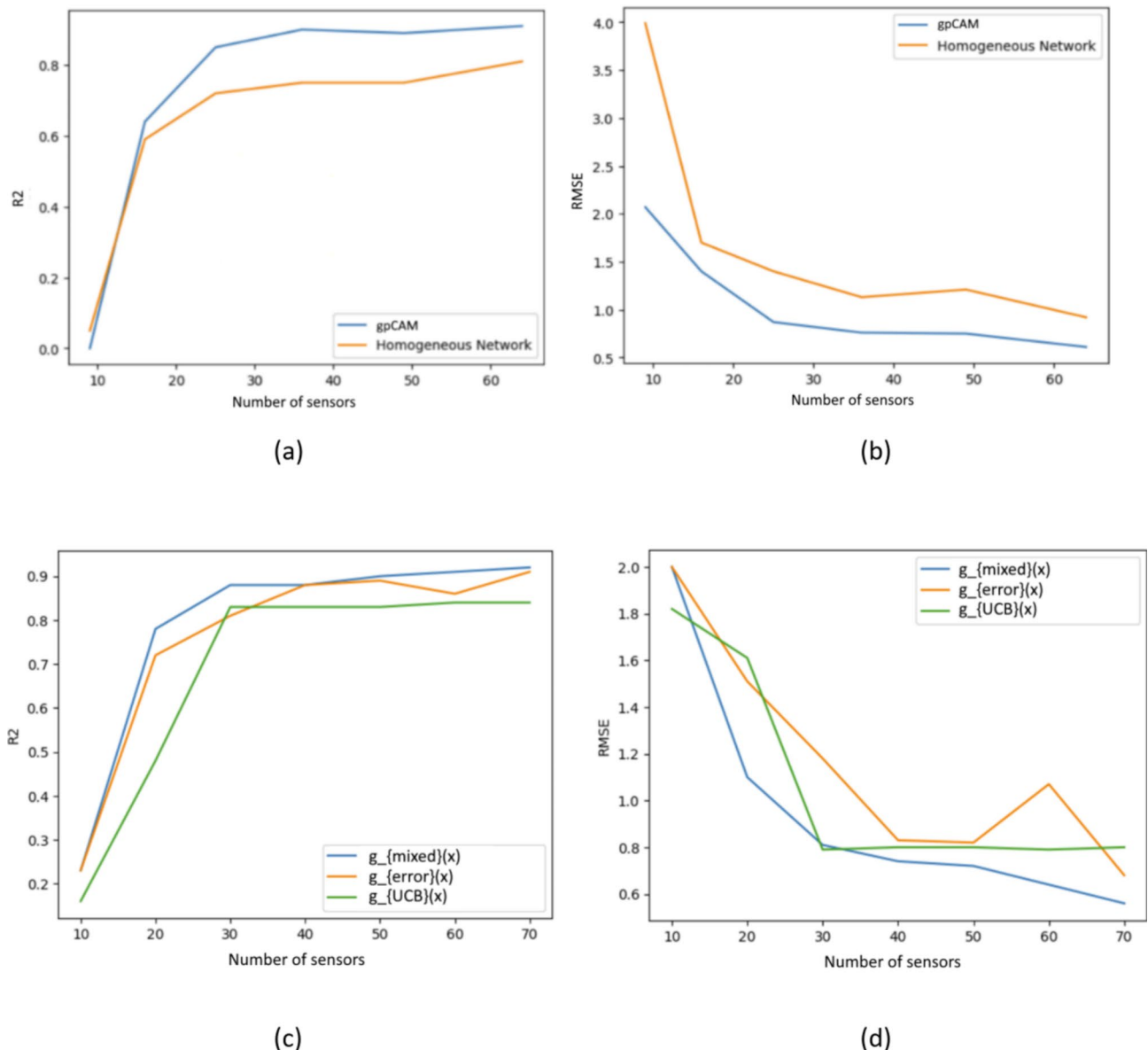


Fig. 3. Evolution of RMSE and R^2 scores with the number of sensors for (a,b) homogeneous vs. GPCAM and (c,d) different acquisition functions. The figures (c,d) were obtained with GPCAM, for a kernel function with $\nu = 1/2$, for all three acquisition functions. The R^2 score increases with the number of sensors, and the RMSE decreases in all figures, showing that the extrapolations of the radioactive plume are improved when there are more sensors.

for a small network. In the subsequent analysis, we use this number to compare the performance of various combinations of kernel and acquisition functions for such a network (Table 1).

The combined impacts of the different kernel and acquisition functions on GPCAM performance are also studied, so as to choose the best combination. The combination of the Matérn kernel with $\nu=1/2$ and the $g_{mixed}(x)$ acquisition function results in the best performance on average over 10 WSPEEDI reference maps, specifically, with R^2 scores 25–75% higher and RMSE 15–40% lower than the other combinations (Table 1). This combination is therefore selected for the rest of this study. The Matérn Kernel with $\nu=1/2$ yields the best results compared to the other kernels in combination with any acquisition function. Comparatively, the mixed acquisition function tends to result in the best results. It was only slightly outperformed in terms of RMSE by the UCB acquisition functions when combined with the RBF kernel.

Temporal evolution

Here we analyze the placement of mobile sensors for non-static plumes. Specifically, GPCAM was trained on a 4-hour long time series of plume forecasts in 1-hour intervals that represent the hourly plume propagation (Fig. 4). Essentially, GPCAM treats the three-dimensional dataset as a function of the 2D space and time axis, so that the interpolation between the points can capture the evolution of the plume with a small number of sensors in time and space. This translates to the situation in which mobile sensors are taking point measurements during the accident.

The number of measurements performed within each hour is 5 to 10, chosen by GPCAM with the mixed acquisition function. Indeed, since time is a part of the parameter space in GPCAM, the acquisition function not only determines the best locations for measurements, but also the best point in time to perform them. In the results, GPCAM reaches an R^2 score of 0.55 and an RMSE of $1.58 \log_{10}(\mu\text{Gy/h})$. It appears that GPCAM tends to place more sensors in the overall direction of the plume. As time passed, the increasing number of sensors was placed within the plume, while there were few sensors in the low air-dose-rate regions.

Ensemble optimization

In this section, we focus on designing a radiation monitoring network in advance, considering that metrological conditions are variable. These sensors are therefore assumed to be static, the locations of which are optimized to capture at best all potential plumes. We assume that we can simulate a potential plume migration at a particular location if we have sufficient historical metrological data. We use ensemble simulation results to optimize the sensor locations. Since we need a large number of reference maps for the ensemble-based optimization algorithm, we use reference maps based on an analytical Gaussian plume model (see Sect. 4.2.2). In the model, the source term, σ_y , σ_z , emission height, wind speed, and stability classes were based on several years of meteorological data simulations around MIT, which were established using actual WRF simulations in the study by Ayoub et al.¹⁴. The mean radiation map of the reference maps generated with the analytical plume generator was plotted (Fig. 5b) and proved to be similar to the expected wind rose (i.e., diagram showing the relative frequency of wind directions at a place) at MIT (Fig. 5a).

The four ensemble optimization algorithms are compared to the *homogeneous network*: (1) the *sequential network* approach places sensors based on the first chosen reference map, and more sensors are added in the following reference maps; (2) the *grid majority-voting network* is a majority voting approach such that the maps are divided into 2 km mesh grids, and sensors are taken from the grids with the most sensors and added to the final network; (3) the *adjustable grid majority-voting network* is another majority voting approach similar to the grid majority-voting approach, but the radius—or minimal distance between sensors—directly defines the grid size and can therefore be adjusted; (4) the *combined network* is a mix of the *homogeneous network* and *adjustable grid majority-voting network*.

Results show (Fig. 6) that the *homogeneous network* performs better than all the other networks, on average, when using less than 40 sensors. Above 40 sensors, the *sequential*, *adjustable grid majority-voting network* and *combined network* seemed to perform better. On closer look, it turns out that the *adjustable grid majority-voting network* performs better than the *homogeneous network* in many instances, as shown in the distribution of R^2 scores for both approaches on the different plume realizations (see Supplementary Fig. S3). However, the average performance of the *adjustable grid majority-voting network* is dragged down by a few cases in which the sensors are not covering the plume properly, resulting in R^2 scores between 0 and 0.4. On the other hand,

MEAN OF 10 MAPS		Matérn Kernel with $\nu=1/2$	Matérn Kernel with $\nu=3/2$	Kernel RBF
R ² score	$g_{mixed}(x)$	0.67	0.50	0.30
	$g_{error}(x)$	0.45	0.37	0.22
	$g_{UCB}(x)$	0.45	0.32	0.18
RMSE	$g_{mixed}(x)$	1.16	1.42	1.73
	$g_{error}(x)$	1.53	1.73	1.99
	$g_{UCB}(x)$	1.35	1.53	1.68

Table 1. Comparison of the mean performance of R^2 scores and RMSE for all combination of kernel functions (Matérn with $\nu=1/2$, Matérn with $\nu=3/2$, RBF) and acquisition functions ($g_{mixed}(x)$, $g_{error}(x)$, $g_{UCB}(x)$). GPCAM was run for each combination on 10 WSPEEDI reference maps, with 25 sensors.

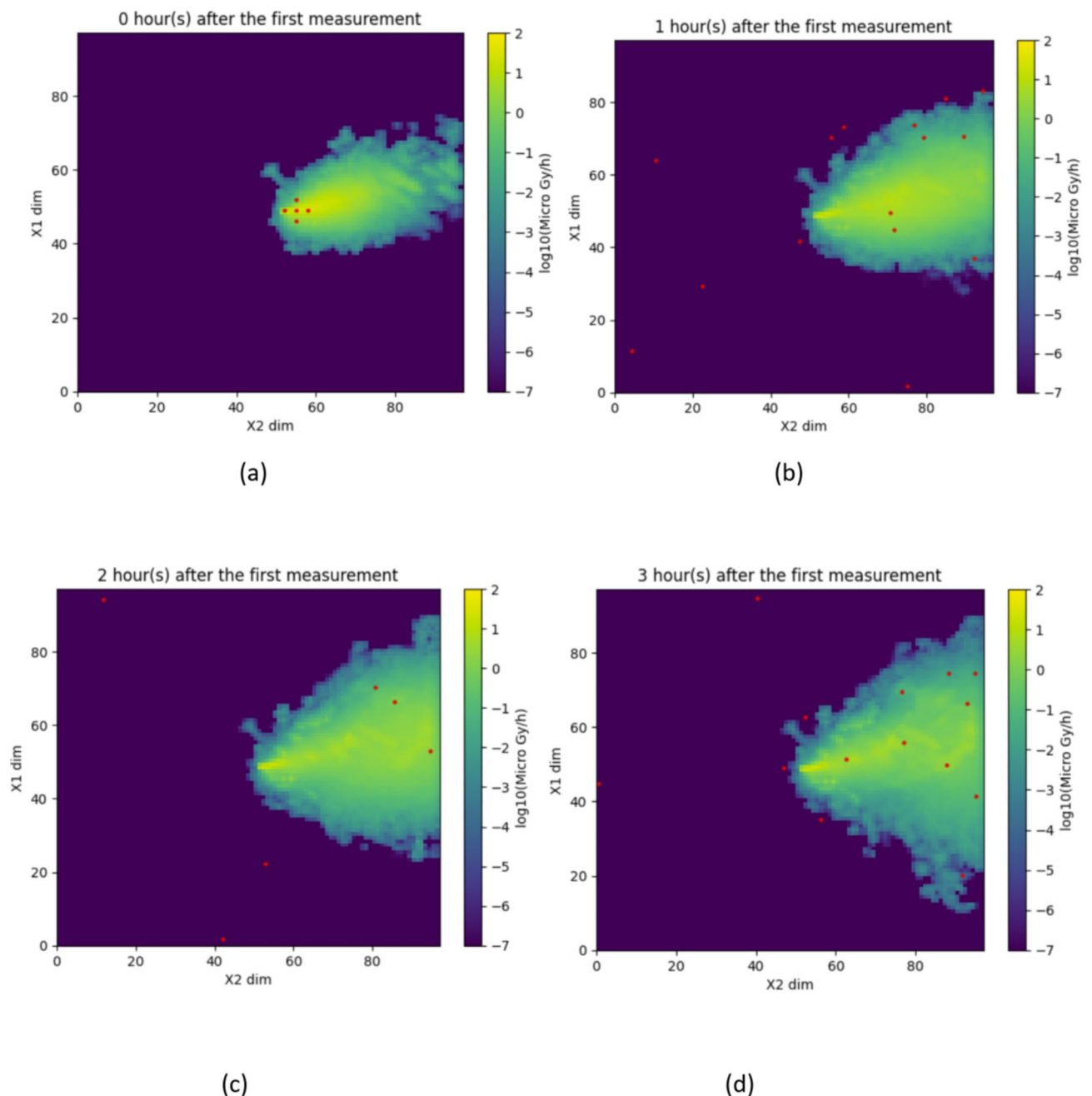


Fig. 4. Temporal evolution of the reference radioactive plume, at the moment of the first performed measurement (a), 1 to 3 h after the first measurement (b–d). The red dots are the measurements performed at the given moment in time. The initial radial distribution of sensors around the source appears clearly (a), then all the measurements performed within the next few hours are plotted on the following maps, at one-hour intervals, for a total of 40 measurements.

the *homogeneous network* manages to interpolate all the plumes with R^2 scores between 0.76 and 0.96. The histogram of the R^2 scores of the *combined network* shows that this approach manages to get rid of the very low scores found with the *adjustable grid majority-voting network*, while still maintaining a good distribution on the high score end.

This behavior can be understood by investigating the distribution of the sensors in the networks (Fig. 7). The *adjustable grid majority-voting network* (Fig. 7a), tends to capture the most likely plume scenarios, which can be shown in the rose map (Fig. 5), since it places more sensors in the prevailing wind directions, hence the very high scores compared to the *homogeneous network*. Alternatively, it fails to capture the rare occasions where the wind is not blowing in the prevailing directions, hence the very low scores. Figure 7c shows an example of the resulting network from the *combined* approach. A few more sensors are still placed towards the east, although the blank spots towards the west are here more occupied.

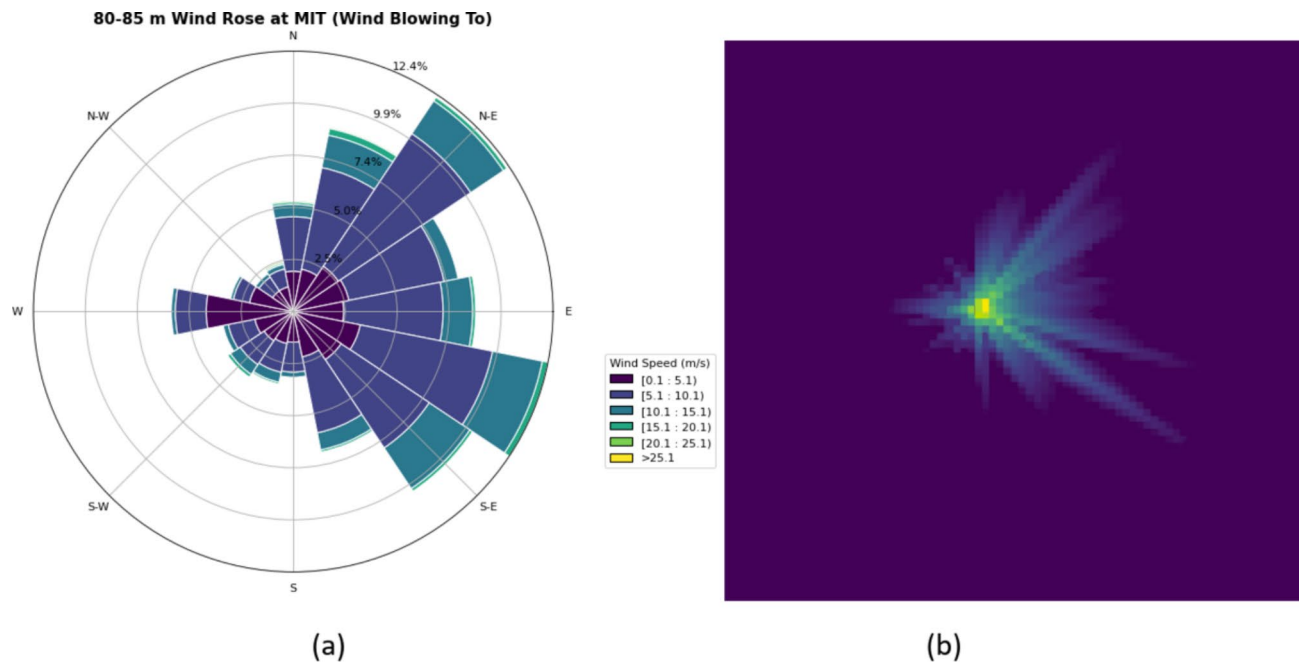


Fig. 5. Expected wind rose at the MIT nuclear reactor (a) Average plume over 1000 maps generated by the analytical plume generator (b). The shape of the mean plume resembles the shape of the MIT wind rose, with the plume profiles following the prevalent wind directions, toward the northeast and southeast.

The case in which there was a more distinctive prevailing wind direction was studied (Table 2) to see whether it had an impact on the performance of the *combined* and *homogeneous networks*. In this new case, it was assumed that the wind would blow within a 60° cone. Such situations show the advantage of the network optimizations over the homogenous case. The *combined network* has a higher R^2 score and RMSE score that is 30% lower than the *homogeneous* one. This represents a significant increase in the network performance that could be even bigger if the cone in which the wind blows were smaller.

Discussion

Utilizing the GPCAM software, this study has developed an approach to designing an optimized radiation-monitoring network for capturing the distribution of radiation levels and radioactive plumes with minimal sensors during nuclear power plant accidents. Specifically, the optimized network improves the estimation of the radiation dose-rate distribution based on sparse sensor measurements. To the authors' knowledge, GPCAM has not been used in the application of designing a radiation monitoring network. We explored different parameters and options within GPCAM as well as different scenarios—such as 2D static distributions, temporal evolution, and ensemble optimization algorithms—so that we can tailor it to the specific application of monitoring network optimization.

Results show that the selection of the kernel function has a significant impact on estimation performance. The Matérn kernel with $\nu = 1/2$ (i.e., the exponential kernel) leads to a better performance than the RBF kernel. In the past, the RBF kernel has been widely used for mapping air pollution, since RBF is designed to encode distance-based smoothness by Patel et al.¹⁷. The fact that the exponential performs better in our case may be explained by the fact that the radioactive plumes have the sharp outline of the advection-dominant atmospheric plume from a point source, making it harder to interpolate with the RBF kernel. Our results show that the exponential kernel performs better than RBF. This is due to the fact that the radioactive plumes have a sharp plume boundary due to the advection-dominant dispersion-limited transport. Since RBF has the smoothness parameter being infinity, it has difficulty capturing such a sharp boundary.

In addition, we have investigated the impact of different acquisition functions. The initial application of GPCAM was based on the UCB function. The error acquisition function was also commonly used in Sun et al.⁹ and Araki et al.¹⁸ to minimize the estimation error. Our results have shown the best performance using the linear combination of the UCB acquisition function and the error acquisition function. In our results, the UCB function tends to capture the center of the plume and the error function its periphery, while the mixed function puts stress on both of them, ultimately resulting in a better interpolation.

Our approach is similar to Araki et al.¹⁸, which used kriging (essentially GP) to interpolate the radiation levels between sparse points for regional air-quality monitoring. They also used the same performance metrics as this study, such as the R^2 score and RMSE. However, they used a genetic algorithm (GA) and simulated annealing (SA) for optimizing the network. Our approach includes the optimization steps embedded within GPCAM for improving the interpolation between the points. In addition, Araki et al.¹⁸ did not compare their algorithm

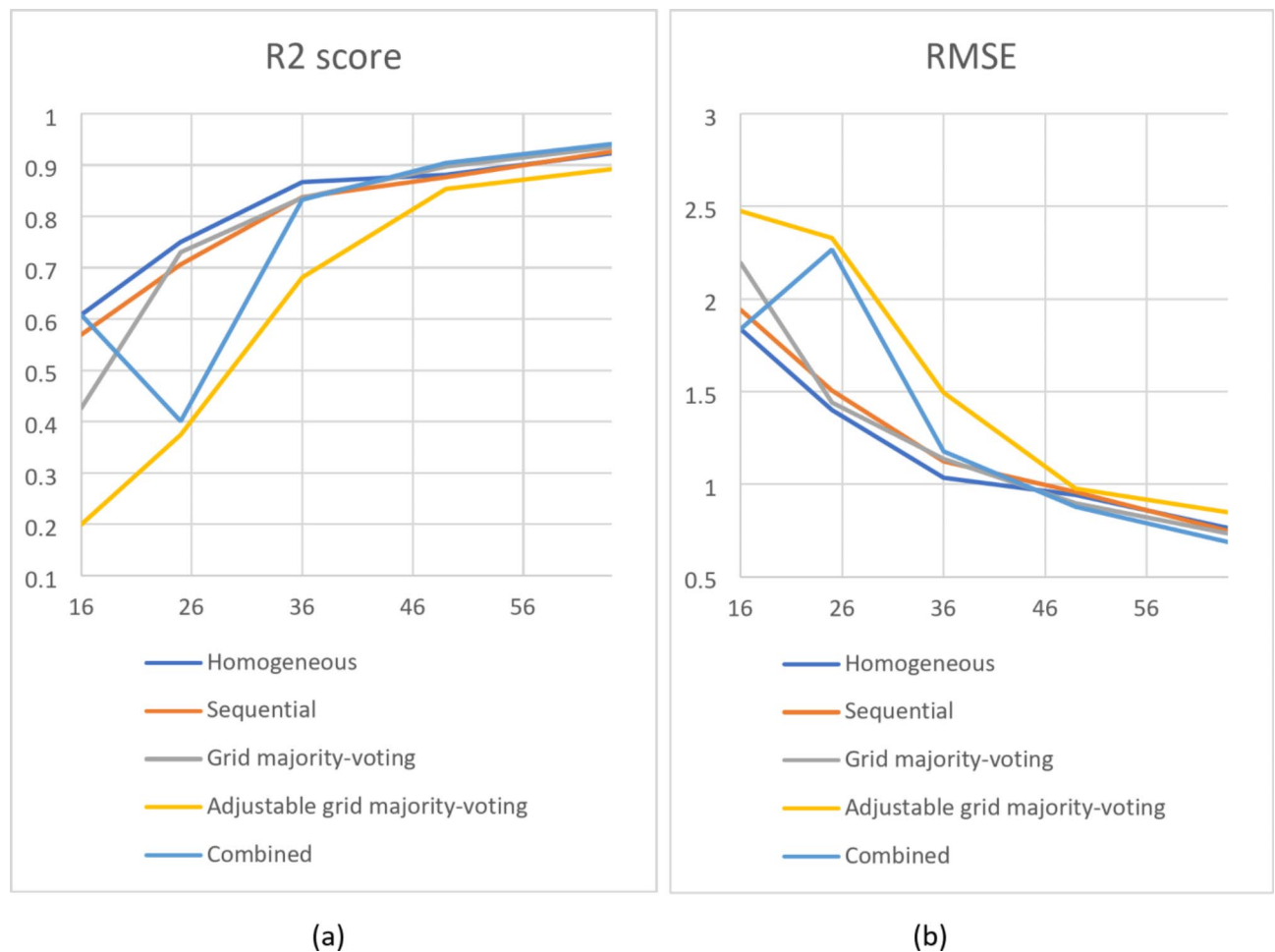


Fig. 6. A training dataset of 700 random gaussian plumes was created (see Sect. 4.2.2). Multiple approaches (homogeneous, sequential, grid majority-voting network, adjustable grid majority-voting network, and combined approach) described in the “Ensemble Optimization Networks” section, were tested and compared to propose various optimal sensors networks that can capture these different plume realizations. These networks were then tested on a dataset of 300 plumes, similarly generated by the Gaussian plume generator, and the R^2 score (a) and RMSE (b) evolution with the number of sensors was plotted. The acquisition function $g_{mixed}(x)$ and Matérn kernel with $\nu = 1/2$ function were chosen. Five sensors were placed in a radial distribution around the source. The plumes generated with the Gaussian plume model were tuned (by increasing the dispersion scales due to the increased mixing in the region) so as to replicate a WSPEEDI plumes (see Supplementary Figure S2).

performance to a homogeneously placed network, nor did they take into account the impact of uncertainty in the wind and metrological conditions on the results.

In this study, we have primarily focused on temporally fixed (static) contamination maps after radionuclides are deposited on the ground in a short time frame. This is motivated by the fact that radionuclide releases often happen as pulses associated with explosions or pressure drops in the containment. In fact, the major contamination area from the Fukushima NPP accident was created within several hours after the Unit 2 containment failed⁴. However, releases could happen multiple times or continuously over a long-term period. In such cases, GPCAM can include the temporal dimension in addition to the spatial dimensions such that the sensor measurements are optimized in both time and space. We have interpreted such spatiotemporal sensor optimization as the use of mobile sensors, and developed the potential application of GPCAM to optimize mobile sensor deployment—where and when to take measurements during the emergency response. Mobile sensors have rapidly advanced after the Fukushima Daiichi NPP accident, and they are still being advanced, as shown by Hochbaum et al.¹⁹ or Flanagan et al.²⁰.

In the ensemble optimization of the sensor network, the sensors were fixed (not mobile). Doing this is an effort to design the fixed monitoring network for locations with variable wind and other metrological conditions. Our results show that the sensors homogeneously scattered across the map generally perform well. This may be explained by the application to a case study where the wind blows in various directions throughout the year. However, our results show that for the sites with a clear prevailing wind direction, the GP-based algorithms perform the best, and show superior performance compared to the *homogeneous network*, even with few sensors.

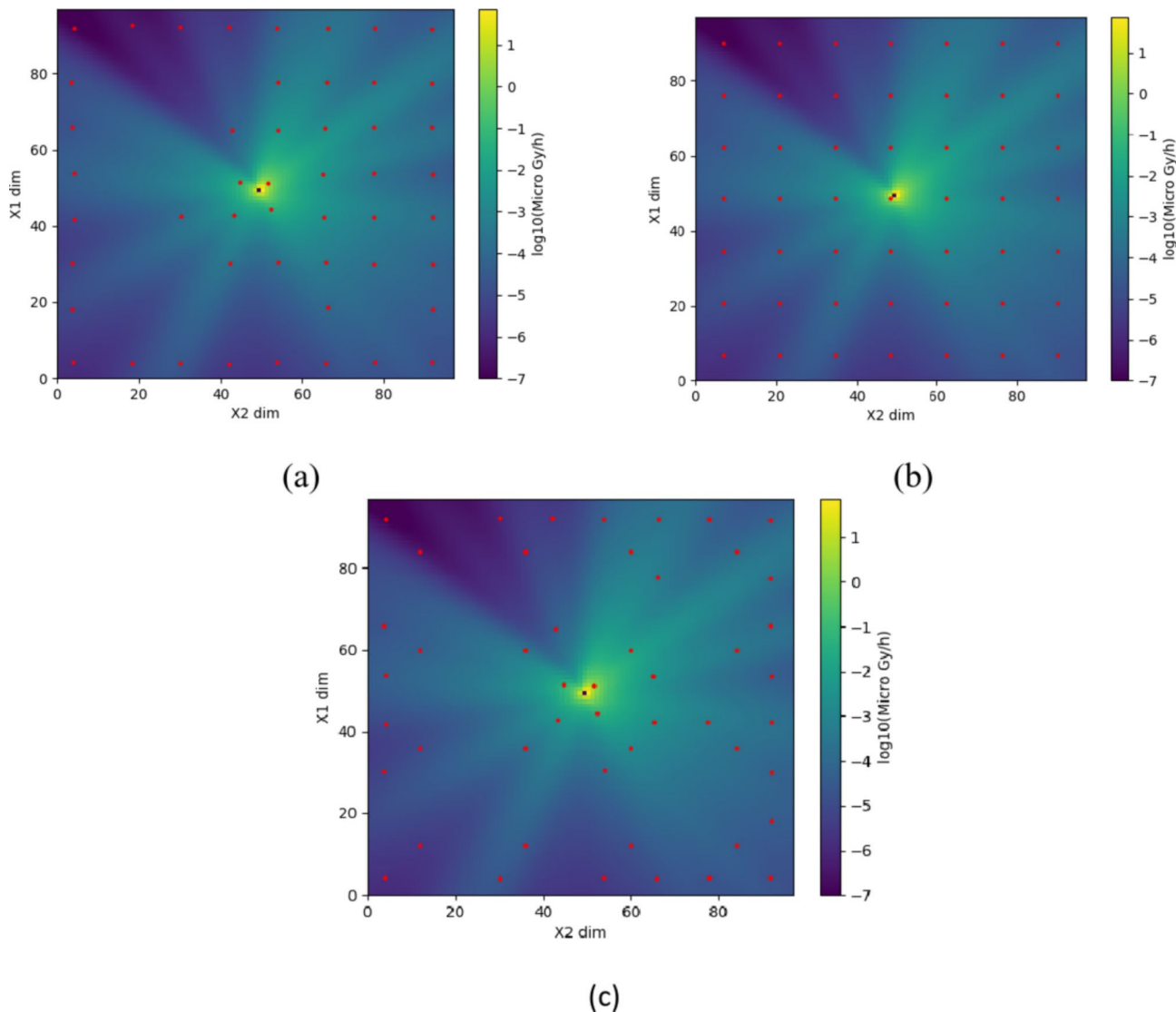


Fig. 7. Adjustable grid majority-voting network: (a) Homogeneous network and (b) combined network (c), on top of the average plume map. The combined approach shows a standard deviation similar to the homogeneous one, and a higher mean score. Indeed, the homogeneous network presents a mean value of 0.88 for a standard deviation of 0.07, the adjustable grid majority-voting network a mean value of 0.85 for a standard deviation of 0.23, and the combined approach a mean value of 0.90 for a standard deviation of 0.08. The significantly lower standard deviation of the combined vs. adjustable grid majority-voting network confirms that the combined approach manages to get rid of the adjustable grid majority-voting network low R^2 scores, which are far away from the mean.

	Homogeneous	Combined
R^2 SCORE	0.85	0.95
RMSE ($\log_{10}(\mu\text{Gy/h})$)	1.04	0.71

Table 2. R^2 score and RMSE of homogeneous network and combined approach for a 60° cone direction for the wind. 700 hundred analytical reference maps were generated for the training dataset, and 300 maps for the testing dataset. These results were obtained for 49 sensor networks.

We acknowledge that our approach still has some limitations. Other factors could impact the algorithm, such as the dataset itself, the shape of the map, or the placement of the first sensors. In our analysis, we found that the first several sensors need to be placed radially around the NPP location to minimize the number of sensors. Otherwise, GPCAM struggled to converge with a low number of sensors without these initial placements. We assume that this setup helps GPCAM identify the overall direction of the plume early on. In addition, the use

of numerical atmospheric transport simulations is constrained by their high computational costs, which forced us to use analytical plume models in this study. In the future, we may use the machine learning-based surrogate model approach developed by Ayoub et al.¹⁵ to generate ensemble simulations based on numerical simulations such as WSPEEDI. The numerical simulations allow us to consider other key parameters like land surface, boundary layer, microphysics, cumulus clouds, and radiation. In addition, the source term was assumed to be fixed in our study, although it should ideally be coupled with accident analysis. We consider this paper as a proof-of-concept study, which will enable more research on radiation-monitoring optimizations in the future. We also acknowledge that our study focuses only on the number and locations of sensor points and optimization algorithms. For real-world applications, further development will be necessary to consider the cost of radiation monitoring posts, including data transmission/management and system maintenance. At the same time, field deployment logistics should be included such as the accessibility of the locations and restricted areas. Our algorithms are expandable to include additional spatial data layers in a similar manner as Wainwright et al.²¹.

We acknowledge that our methods have been demonstrated solely using hypothetical scenarios and datasets. Further testing can be conducted on available experimental data (e.g., Gowardhan et al.²²) in the future. This paper is focused on documenting the developed algorithms and reporting their performance, with a particular emphasis on improving the understanding of algorithms such as the impact of kernel parameters and acquisition functions.

Methods

Gaussian process models

A Gaussian Process (GP) model uses a stochastic process to represent the spatial heterogeneity of natural variables or can be used as a regression method for describing target variables as a function of multiple parameters²³. Specifically, we assume that the target value f , which represents the contaminant distribution or dose rate, is a function of multivariate parameter vectors x , which are positions in space. The data points are defined as the set of pairs of f_i and parameter vector x_i , where i is the pair index. One of the required assumptions for using a Gaussian process is that the target variable or the target vector follows a multivariate Gaussian distribution with a mean vector (μ) and a covariance matrix (K).

$$f \sim MVN(\mu, K) \quad (1)$$

Using the data points, a Gaussian process model is trained to predict the value f^* at unseen points x^* based on observed points x . Training the Gaussian process involves the estimation of the mean and covariance parameters. The covariance matrix, or kernel function $K(x_1, x_2)$, quantifies the correlation between the point values x_1 and x_2 . $K(x_1, x_2)$ is a function of the distance between the locations of parameter vectors, often based on the square distance $\|x_1 - x_2\|^2$. If we note the predictive values with an asterisk, the relation between training and predictive values is the following:

$$\begin{bmatrix} f \\ f^* \end{bmatrix} \sim N\left(\mu, \begin{bmatrix} K(x, x) + \tau^2 I & K(x, x^*) \\ K(x^*, x) & K(x^*, x^*) \end{bmatrix}\right) \quad (2)$$

where I is the identity matrix, and $\tau^2 I$ models the assumption that the training labels present a Gaussian noise.

Considering the properties of a Gaussian distribution, one can determine the mean value of a predicted point at position x_2 and the associated error:

$$\begin{aligned} \mu^* &= K(x^*, x) (K(x, x) + \tau^2 I)^{-1} f \\ V^* &= K(x^*, x^*) - K(x^*, x) (K(x, x) + \tau^2 I)^{-1} K(x, x^*) + \tau^2 I \end{aligned} \quad (3)$$

GPCAM

GPCAM is a software tool developed by the Center for Advanced Mathematics for Energy Research Application (CAMERA)²⁴, designed to facilitate the acquisition and analysis of data. It is possible to control and change almost every part of the software, so as to adapt the algorithm to a wide range of domains.

Sensor network optimizations: kernels

There are several kernel functions to model the correlation between data points. The most commonly used one is the radial basis function (RBF), using the exponent of the squared distance between the data locations divided by the spatial scale (h). Kernel RBF is therefore defined as:

$$K(x_1, x_2, h, \sigma) = \sigma^2 e^{-\frac{\|x_1 - x_2\|^2}{h^2}} \quad (4)$$

GPCAM, for its part, uses by default an anisotropic stationary Matérn kernel function. The general expression of the Matérn kernel is as follows, where K_ν is the modified Bessel function of the second kind, σ^2 controls the variability of the GP, h defines dependence between nearby data points, and ν influences the mean-square differentiability of the GP

$$K(x_1, x_2, \nu, \sigma, h) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\|x_1 - x_2\|}{h} \right)^\nu K_\nu \left(\sqrt{2\nu} \frac{\|x_1 - x_2\|}{h} \right) \quad (5)$$

GPCAM uses an anisotropic stationary Matérn kernel function, where $\nu = \frac{3}{2}$, defined as:

$$K(x_1, x_2, 3/2, \sigma, \mathbf{h}) = \sigma^2 (1 + \sqrt{3}d) e^{-\sqrt{3}d}$$

$$d = \sqrt{\sum_{i=0}^n \left(\frac{(x_{1i} - x_{2i})}{h_i} \right)^2} \quad (6)$$

where σ is the variance hyperparameter of the Gaussian process, x_{1i}, x_{2i} are the parameters value at the points x_1, x_2 , and h_i is their associated scale hyperparameter.

Finally, this study considers another Matérn Kernel, with $\nu = \frac{1}{2}$, defined as the following:

$$K(x_1, x_2, 1/2, \sigma, \mathbf{h}) = \sigma^2 e^{-d} \quad (7)$$

The hyperparameters of the Kernel function must be properly determined through an optimization process. GPCAM allows the user to automatically train the process on a dataset in order to tune the hyperparameters and fit the training data. This is done through minimizing the negative log likelihood of the hyperparameters. Once the Gaussian process is trained and the hyperparameters are optimized, the kernel matrix can be inserted into Eq. (3) to determine the covariance and mean value of unobserved points.

Sensor network optimization: acquisition functions

In a Gaussian process model, the acquisition function is a criterion used to choose the next data point to measure in order to optimize the model. The default acquisition function implemented in GPCAM is the UCB function, and tends to place a new point where the variance is the most significant, so as to maximize the chance to improve the overall model.

Another acquisition function used is the absolute error between the predictive data and the reference data. This acquisition function has been chosen in the Sun et al.⁹ study. A new sensor is placed where the error is the highest.

It is worth noting that the choice of the acquisition function depends on the available information and the optimization goals. With a different acquisition function, a balance can be considered between the exploration of the search space and the exploitation of the known promising areas, like the area around the radiation source. For instance, the Upper Confidence Bound (UCB) acquisition function enables us to easily tune the weight associated to exploitation and exploration.

In this study, an approach mixing an Upper Confidence Bound (UCB) as well as the absolute error between the predictive and reference map was selected. The Bayesian objective function is given by:

$$g_{UCB}(x) = \mu^*(x) + 3 V^*(x) \quad (8)$$

where x is a point in space, $\mu^*(x)$, and $V^*(x)$ are the value and associated variance predicted at this point by the Gaussian process.

The other objective function that we opt to minimize is the error between the predicted map and the reference map, defined as:

$$g_{error}(x) = \mu^*(x) - \mu(x) \quad (9)$$

where $\mu(x)$ is the reference (true) value at the point x .

The acquisition function we used in this study, referred to as “mixed acquisition” in the rest of this report, is defined as:

$$g_{mixed}(x) = g_{UCB}(x) + \lambda g_{error}(x) \quad (10)$$

where λ is a weight coefficient that was empirically determined to be equal to 4 in this study (see Supplementary Figure S4).

Plume generation

In this study, both numerical and analytical plume models were used. As mentioned in the Results section above, this study was divided into two main parts. Numerical plumes generated with WSPEEDI were used to analyze the impact of the various kernel and acquisition functions on GPCAM performance. Then, analytical plumes were used to reduce computational time for the Ensemble optimization.

WSPEEDI plume simulations

We used the same model and setting as the study from Ayoub et al.¹⁴. We briefly describe it as follows for completeness. To select the optimal kernel and acquisition function to be used in GPCAM, we generate the radiation concentration and dose maps using the 1.1.4 version of WSPEEDI (Worldwide version of System for Prediction of Environmental Emergency Dose Information)¹². WSPEEDI uses the Weather Forecasting Model (WRF) to simulate meteorological fields required in the dispersion and transport calculations, and employs GEARN to model the latter.

WRF, developed by the National Center for Atmospheric Research²⁵, is a regional numerical forecasting system, with a compressible nonhydrostatic approach. The WRF model forecasts three-dimensional meteorological data

by addressing several key equations related to atmospheric dynamics. This model incorporates a range of physics configurations to account for aspects such as land surface, boundary layer, microphysics, cumulus clouds, and radiation.

GEARN models the dispersion of radionuclides released into the atmosphere by using the three-dimensional motion of marker particles of the meteorological conditions²⁶. It simulates various processes, including the movement and spread of radioactive plumes through advection and diffusion, radioactive decay, and wet, dry, and fog water deposition on the ground.

To conduct the WRF calculations, we simulated a region of 600 km by 600 km, with MIT at its center (referred to as Domain 1 in Fig. 8). Initial and boundary conditions for the WRF simulations were derived from the Grid Point Value (GPV) data of the Global Spectral Model (GSM), which is provided by the Japan Meteorological Agency (JMA)²⁷ as GRIB data. This dataset is available at a 20 km horizontal grid resolution and is updated every 6 h, and defined at surface pressure as well as at 17 isobaric vertical levels up to 10 hPa.

Furthermore, for enhanced accuracy in a specific area of interest—where we intend to simulate transport using GEARN—we utilized a domain nesting technique. This approach involves running simultaneous simulations in both a finely resolved domain (Domain 2 in Table 3) and a larger surrounding coarser domain (Domain 1). The settings for the WRF and GEARN calculations used in our simulations are summarized in Table 1.

To capture the potential plumes from a release at a location, large training and test plume scenarios were necessary, and generating them through WSPEEDI is computationally expensive (average run time on an Intel Xeon Gold 6240 CPU is about 1.5 h for a single WSPEEDI forecast). Therefore, it was decided to utilize an analytical atmospheric transport and dispersion method, namely the Gaussian plume model, to generate a large ensemble of plumes that represent the different potential plume scenarios.

Gaussian plume model

The Gaussian plume model is based on analytical solutions to quantify the radionuclide concentration for an elevated point source with full ground reflection²⁹:

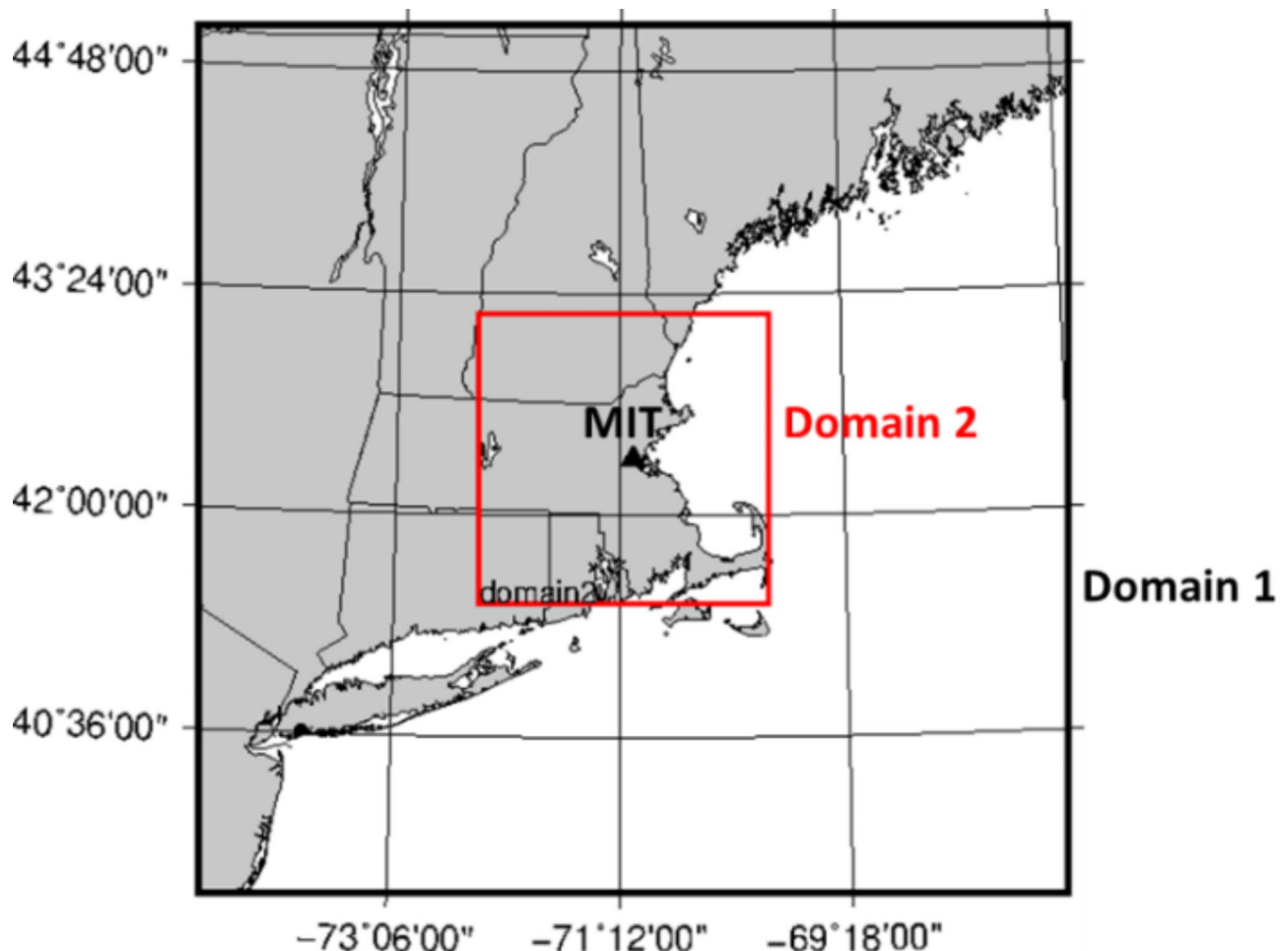


Fig. 8. Computational domain for the WRF and GEARN simulations. Parent domain 1 and high-resolution nested domain 2 are centered at MIT with Latitude and Longitude of 42°21'25" and -71°5'34" respectively.

	Domain 1	Domain 2
Applied WRF calculations	Yes	Yes
Applied GEARN calculations	No	Yes
Simulation Period (until steady state)	8 h	8 h
Horizontal grid number	100 × 100	100 × 100
Horizontal resolution (km)	6 km	2 km
Vertical levels of WRF	35 hybrid levels from surface up to 100 hPa	
Vertical levels of GEARN	29 levels from surface to 10 km	
WRF Integration time step	36 s	12 s
GEARN dispersion time step	–	12 s
Boundary and initial conditions data	GPV (GSM, 6 h frequency, 20 km resolution, at surface pressure and 18 isobaric levels)	
Physics options (microphysics, cumulus, land surface boundary layer, long-wave and short-wave radiation)	Same as in Kadowaki et al. ²⁸	
Source Term	One hour release from the center: 2e16 Bq/hr I131 and 4e14 Cs 137. Release location: Latitude and Longitude of 42°21'25" and –71°5'34"	
Release heights	25 m and 85 m (to simulate releases from both the containment and stack respectively)	

Table 3. Simulation settings of WRF and GEARN.

$$\overline{\chi}(x, y, z) = Q \frac{e^{-\frac{y^2}{2\sigma_y^2}}}{2\pi \overline{u} \sigma_y \sigma_z} \left\{ \exp \left[-\frac{(z-h)^2}{2\sigma_z^2} \right] + \exp \left[-\frac{(z+h)^2}{2\sigma_z^2} \right] \right\} \quad (11)$$

Where $\overline{\chi}$ is the integrated air concentration of radionuclides (Bq.s/m³) and Q the integrated (total) activity of the released radionuclides (Bq). The horizontal and vertical distances, and the height from the plume centerline (m), are noted as y, z, and h. The plume dispersion parameters in the y and z directions (m) are noted as σ_y , σ_z . The average wind speed at the release location is $\overline{u}$ (m/s). For the Gaussian plume model, several parameters were randomly sampled from their probability distributions based on MIT available data. The parameters include wind speed and direction, as well as different atmospheric stability classes and emission height. No terrain aspects were considered in this Gaussian model. In this analytical model, we assumed a source term of 2.0E19 Bq for Iodine-131 and 4.0E17 Bq for the Cesium-137, respectively. The plume dispersion parameters σ_y , σ_z were selected using the Pasquill-Gifford atmospheric stability class of the region. Specifically, we sampled from the class “E” and “D”, which were the two likely atmospheric stability classes at MIT. The emission height was sampled from the two potential release heights, 25 m (containment release) and 85 m (stack release). The wind speed and direction were randomly sampled from the wind rose (Fig. 5), which was generated through two years of WRF simulations for the MIT area by Ayoub and Wainwright¹⁴.

2D static case studies

In the 2D static plume case, the objective was to properly interpolate a single radioactive reference map with GPCAM. This was achieved in a two-step process (Fig. 1):

Step 1.1: Determination of an interpolated map In this case study, an initial set of 5 sensors radially placed around the source term were given GPCAM. In this step, GPCAM uses the sensors it is given (the 5 initial sensors and added ones at previous iterations) to interpolate the reference map. The logarithmic values of the total dose were taken, in order to better grasp the variation in the low dose levels. The dose rate was converted to the logarithmic scale following previous studies^{7,9}. We consider that the log scale is appropriate, since the radiation level distribution tends to have heteroskedasticity such that higher values have higher variability. The log-transformation is necessary for using a stationary Gaussian process model. We acknowledge that taking the log scale could de-emphasize the high-dose region. However, as stated in the introduction, our goal is to capture the heterogeneous distribution of the radiation dose rates and the overall plume shape.

Step 1.2: Placement of a new sensor Once an interpolated map of the radiation is obtained, new bins for additional sensors can be determined. To do so, we selected the appropriate acquisition function from the section “Sensor network optimization: acquisition functions”. The acquisition function then associates a merit score to a set of points to guide the choice of a new bin to add a sensor.

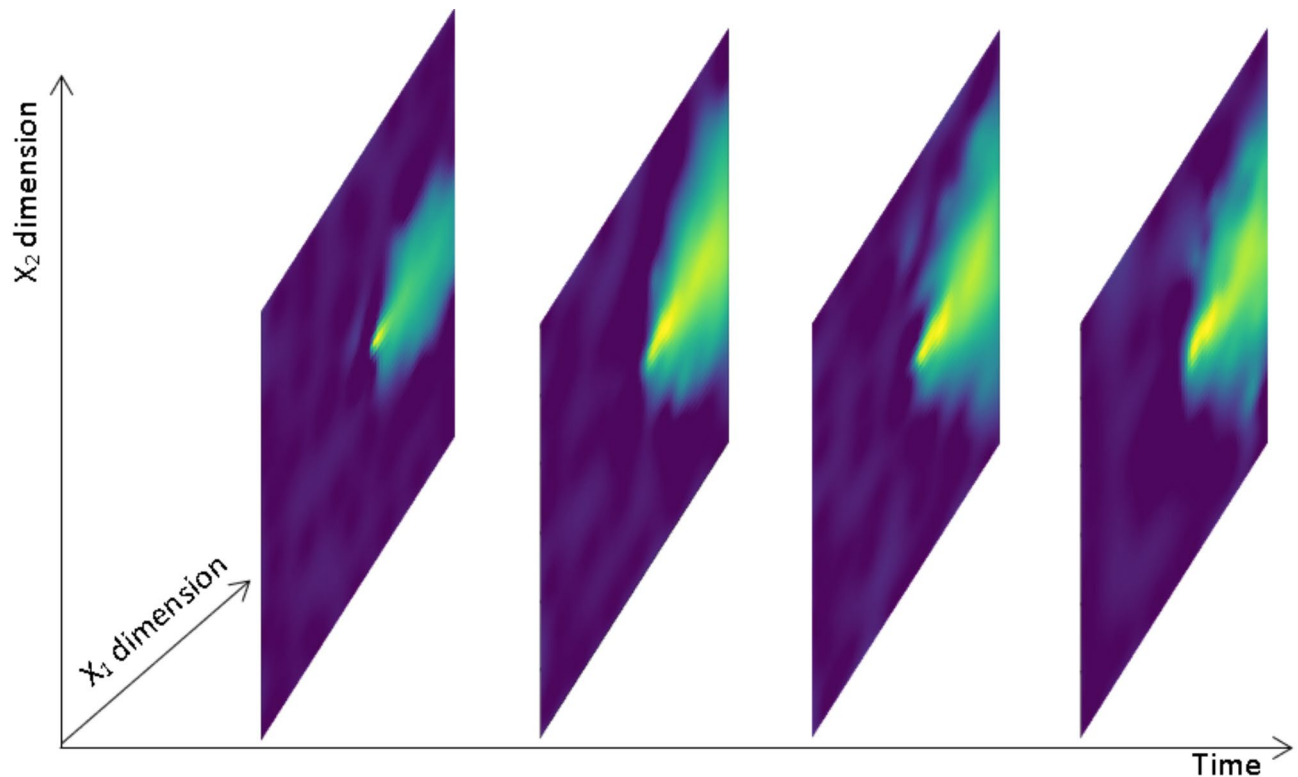


Fig. 9. Example of 3D Plume reference map, with spatial and time dependence (X1, X2 directions).

Steps 1.1 and 1.2 are then repeated until enough sensors are placed. As the number of sensors increases, more data are available to produce an interpolated map (step 1.2). The RMSE score is thus expected to decrease with the number of sensors, because the interpolated map is more representative of the reference map.

Temporal evolution

Assuming an NPP accident happens, one might want to properly capture the expansion of the radioactive plume over time. A natural way to implement the temporal evolution in GPCAM is to add a new dimension to the dataset. The plume is then not only defined by the radiation value as a function of two space coordinates (X1 and X2 dimension shown in Fig. 9), but also as a function of time.

The plume evolution over time was simulated using WSPEEDI, and concatenated to form a 3-dimensional dataset. The maps represent the plume evolution in time with a one-hour interval. Only the first four maps were taken to run the simulation, since the total radiation dose tends to converge and reach a steady state 3 h after a puff release, as pictured in Fig. 9.

Figure 9 shows how to create a 3D dataset by adding a temporal axis to represent the evolution of the radioactivity map.

It is important to note that the measurements are performed at a given moment in time at a specific location. Indeed, the sensors were previously placed according to the 2D space variables when the data was not time dependent. In this case, they are also associated with the time parameter. Therefore, the sensors are not fixed, but are like measurements made at a certain moment and certain location.

Performance evaluation criteria

The Root Mean Square Error (RMSE) and R-Squared (R^2) score were used to evaluate the performance of the Gaussian process, including the French Polyphemus³⁰ air quality modeling system or the Austrian radiation monitoring network ARAD³¹.

The RMSE and R^2 scores were defined as the following:

$$R^2 = \frac{(\overline{x_1 x_2} - \overline{x_1} \overline{x_2})^2}{(\overline{x_1^2} - \overline{x_1}^2)(\overline{x_2^2} - \overline{x_2}^2)} \quad (12)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_{1i} - x_{2i})^2}{N}} \quad (13)$$

where x_1, x_2 are the groundtruth data and predicted data, respectively, and $\overline{x_1}, \overline{x_2}$ are the corresponding mean values.

Ensemble optimization networks

Various network algorithms were tested in order to catch all potential plumes in case of a radioactive release accident. These algorithms (apart from the *homogenous network*) are based on a training set of analytical plumes.

Homogeneous network

This network is the most intuitive network, yet still interesting to consider. In this case, the sensors are spread uniformly across the map with equal distances between the sensors.

Sequential network

In this method, GPCAM simulations are run on each of the training reference maps, which generate one different monitoring network for each map with a set number of sensors each. Then the networks were taken one at a time in a random order. For a given network, the sensors were taken one at a time in the order of GPCAM sequence and added to the final monitoring network (unless there already was a sensor in the final monitoring network within a given radius). The impact of the radius size on the network efficiency is discussed later in the Results section.

The iterative process is stopped when the final monitoring network has reached a given number of sensors, or when all the sensors of the training networks have been iterated over.

Grid majority-voting network

In a similar fashion, training networks are first obtained by running GPCAM on each of the training reference maps. This time, the networks are superposed, and a mesh grid of two X two kilometers is defined on the continuum dense network. For each grid, the number of sensors contained is counted. The grids are then taken by decreasing order based on their number of sensors, and one random sensor from each grid is added to the final monitoring network, until it has reached a given number of sensors, and only if there is no sensor already placed within a given radius.

Adjustable grid majority-voting network, second approach

In a similar way to the 2 k-mgrid network, the independent training networks are superposed and divided in a grid. This time, the size of the grid is defined by the radius value (the minimum allowed distance between two sensors). The grids are then taken by decreasing order based on their number of sensors, and one random sensor from each grid is added to the final monitoring network, until it has reached a given number of sensors.

A combined approach

In this approach, a subset of the total planned sensors is first distributed uniformly (*homogeneous network*), then an *adjustable grid majority-voting network* was used to optimize the rest of the sensors on top of that, until a given number of sensors is reached.

Data availability

The data that support the findings of this study, or the methods to generate the data are included in this article (and its Supplementary Information files).

Received: 9 August 2024; Accepted: 21 March 2025

Published online: 07 April 2025

References

1. Energy Electricity and nuclear power estimates for the period up to 2050. *Int. At. Energy Agency* (2022).
2. Ansolabehere, S. & Konisky, D. M. Public attitudes toward construction of new power plants. *Public. Opin. Q. Vol. 73*(3), 566–577 (2009).
3. Abdulla, A., Azevedo, I. L. & Morgan, M. G. Public perceptions of nuclear power energy risks: bridging the academic and cultural divide. *Energy Policy* **32**(10), 1457–1470 (2019).
4. Ayoub, A., Wainwright, H. M., Sansavini, G., Gauntt, R. & Saito, K. Resilient design in nuclear energy: critical lessons from a cross-disciplinary review of the Fukushima Dai-ichi nuclear accident in *iScience*. **27** (2024).
5. Collins, H. E., Grimes, B. K. & Galpin, F. Planning basis for the development of state and local government radiological emergency response plans in support of light water nuclear power plants. *USNRC-NUREG* **0396**(1978).
6. Hayano, R. S. & Adachi, R. Estimation of the total population moving into and out of the 20 Km evacuation zone during the Fukushima NPP accident as calculated using Auto-GPS mobile phone data. *Proc. Jpn. Acad. Ser. B.* **89**(5), 196–199 (2013).
7. Wainwright, H. M. A multiscale bayesian data integration approach for mapping radionuclide contamination in *Resilience: A New Paradigm of Nuclear Safety* 57–65 (2017).
8. Wainwright, H. M., Seki, A., Mikami, S. & Saito, K. Characterizing regional-scale Temporal evolution of air dose rates after the Fukushima Daiichi nuclear power. *Plant. Accid. J. Environ. Radioact.* **210** (2019).
9. Sun, D. et al. Optimizing long-term monitoring of radiation air-dose rates after the Fukushima Daiichi nuclear power plant. *J. Environ. Radioact.* 220–221 (2020).
10. Noack, M. M. et al. A Kriging-Based approach to autonomous experimentation with applications to X-Ray scattering. *Sci. Rep.* **9** (2019).
11. Chang, E., Beverly, L. & Wainwright, H. Aqueous chemistry and physics constrained autonomous-closed-loop modeling for quantifying adsorption processes as applied to metal-mineral interface geochemistry. *Methods Appl. Auton. Experimentation* 292–303 (Chapman and Hall/CRC, 2024).
12. Chino, M., Yamazawa, H., Nagai, H., Moriuchi, S. & Ishikawa, H. *WSPEEDI (Worldwide Version of SPEEDI): A computer code system for the prediction of radiological impacts on Japanese due to a nuclear accident in foreign countries* 63 (1995).
13. Porcu, E. et al. The matern model: A journey through statistics, numerical analysis and machine learning. <http://arxiv.org/abs/2303.02759> (2023).
14. Ayoub, A., Wainwright, H. M. & Sansavini, G. Machine learning-enabled weather forecasting for real-time radioactive transport and contamination prediction. *Prog. Nucl. Energy* **173** (2024).

15. Ayoub, A., Wainwright, H. M., Wang, L. & Sansavini, G. An enhanced fourier neural operator surrogate for radioactive plume transport forecasting. *Stoch. Env. Res. Risk Assess.* 1–11 (2024).
16. Masson, O. & Beguin-Leprieur, M. Modernisation et redéploiement des moyens de surveillance du compartiment atmosphérique à l'IRSN (2015).
17. Patel, Z. B., Purohit, P., Patel, H. M., Sahni, S. & Batra, N. Accurate and scalable gaussian processes for fine-grained air quality inference. *Proc. AAAI Conf. Artif. Intell.* **36**, 12080–12088 (2022).
18. Araki, S., Iwahashi, K., Shimadera, H., Yamamoto, K. & Kondo, A. Optimization of air monitoring networks using chemical transport model and search algorithm. *Atmos. Environ.* **122**, 22–30 (2015).
19. Hochbaum, D. S. & Fishbain, B. Nuclear threat detection with mobile distributed sensor networks. *Ann. Oper. Res.* **187**, 45–63 (2011).
20. Flanagan, R. R., Brandt, L. J., Osborne, A. G. & Deinert, M. R. Detecting nuclear materials in urban environments using mobile sensor networks. *Sensors* **21**, 2196 (2021).
21. Wainwright, H. M. et al. Model and remote-sensing-guided experimental design and hypothesis generation for monitoring snow-soil-plant interactions. *Front. Water* **5** (2024).
22. Large Eddy Simulations of Turbulent and Buoyant Flows in Urban and Complex Terrain Areas Using the Aeolus Model. <https://www.mdpi.com/2073-4433/12/9/1107>
23. A tutorial on Gaussian process regression. Modelling, exploring, and exploiting functions. *J. Math. Psychol.* **85**, 1–16 (2018).
24. Noack, M. M., Pandolfi, R., Humphrey, I. & Carbone, M. Autonomous data acquisition, uncertainty quantification and HPC optimization. *Zenodo* <https://doi.org/10.5281/zenodo.5975553> (2022).
25. Skamarock, C. et al. A Description of the Advanced Research WRF Version 2. *OpenSky* (2005).
26. Terada, H., Furuno, A. & Chino, M. Improvement of worldwide version of system for prediction of environmental emergency dose information (WSPEEDI), (I): new combination of models, atmospheric dynamic model MM5 and particle random walk model GEARN-new. *J. Nucl. Sci. Technol.* **41**, 632–640 (2004).
27. Amagasa, T., Kitagawa, H. & Komano, T. Constructing a web service system for large-scale meteorological grid data. *Third IEEE International Conference on e-Science and Grid Computing (e-Science 118–124 (2007) (2007).*
28. Kadowaki, M., Nagai, H., Terada, H., Katata, G. & Akari, S. Improvement of atmospheric dispersion simulation using an advanced meteorological data assimilation method to reconstruct the Spatiotemporal distribution of radioactive materials released during the Fukushima Daiichi nuclear power station accident. *Energy Procedia.* **131**, 208–215 (2017).
29. Calabria, J. Introduction to atmospheric dispersion process and models. *7th MEREIA Webinar on Basic Concepts* (2023).
30. Mallet, V. et al. Technical note: the air quality modeling system polyphemus. *Atmos. Chem. Phys.* **7**, 5479–5487 (2007).
31. Olefs, M. et al. The Austrian radiation monitoring network ARAD – best practice and added value. *Atmos. Meas. Techn.* **9**, 1513–1531 (2016).

Author contributions

HW and AA supervised the findings of this work. AA, HW, and AG developed the theory. AA performed the WRF simulations. AG performed all other computations. All authors discussed the results and contributed to the final manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-95571-0>.

Correspondence and requests for materials should be addressed to A.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025, corrected publication 2025

© The Author(s) 2025. This work is published under
<http://creativecommons.org/licenses/by-nc-nd/4.0/>(the “License”).
Notwithstanding the ProQuest Terms and Conditions, you may use this
content in accordance with the terms of the License.