

# DESIGN OF CHEMICAL SENSOR ARRAYS FOR MONITORING DISPOSAL SITES ON THE OCEAN FLOOR

*Aleksandar Jeremić and Arye Nehorai*

EECS Department (M/C 154)  
The University of Illinois at Chicago  
851 S.Morgan Street  
Chicago, Illinois 60607-7053

## ABSTRACT

We develop detection methods for automatic environmental monitoring of disposal sites on the deep ocean floor using chemical sensor arrays. Such sites have been proposed for the relocation of dredge materials from harbors and shipping channels; the monitoring is used to detect possible release of pollutants at the site. We model the underwater transport of the pollutants as a diffusion process, and obtain a measurement model by exploiting the spatial and temporal evolution of the associated concentration distribution. The detection problem is defined by a one side hypothesis test for the case of multiple sources. We derive two detectors, the generalized likelihood ratio (GLR) test and the mean detector, and determine their performance in terms of the probabilities of false alarm and detection. The results are applied to the design of chemical sensor arrays satisfying criteria specified in terms of these probabilities, and to optimally select numbers of sensors and time samples. Numerical examples are used to demonstrate the applicability of our results.

## 1. INTRODUCTION

The continuous buildup of sediment in U.S. ports and harbors has a detrimental impact upon national economic and military security. Dredging is required to maintain channel depths; however, the dredged materials may contain considerable quantities of contaminated sediment and waste materials. It has recently been proposed to dispose of these materials by depositing them in bags on the abyssal ocean seafloor. This proposal is called the deep ocean relocation (DOR) program and it has been pursued by NRL and DARPA over the last 4 years, see [1]. As a result of the presence of contaminants in the dredged material, environmen-

tal monitoring of pollutants near the disposal sites must be performed. In this paper we develop procedures for automatic monitoring of the disposal sites using chemical sensor arrays and signal processing techniques. We derive algorithms for detecting the presence of pollutants outside the sites and design sensor arrays for optimal detection performance. This work extends our previous results in [2].

Although our work is developed in the context of the DOR problem, all the results can easily be extended to different environments, such as open air.

## 2. PROBLEM BACKGROUND

According to the DOR plan [1] the dredged material is enclosed in containers each with a volume of about 800yd<sup>3</sup>. A transporter disposes of 20 such units per trip. The disposal location is a region in which the ocean bottom is topographically flat with a low kinetic energy floor. Due to some randomness that exists in the free fall of the bags, their points of impact with the bottom will approximate a circular Gaussian distribution, see [3]. However, the bags' locations will be measured after impact, and we therefore assume that these locations are known.

A fraction of the bags may leak or rupture upon impact with the bottom. The dispersed material will form an apron deposit which consists of the nearfield, where all the bags are to be deposited, and midfield, which is the larger area that will receive dispersed sediment. Program compliance [1] requires that concentration levels of the contaminants must be insignificant beyond the midfield area. For the remainder of the paper we assume that the sensor array will be located in the midfield area with diameter in the range 5-10km.

There are two basic situations that may create a plume of dissolved materials. The first is bag impact, which can be modeled as impulsive source, and the second is water exchange between the pore waters within

---

This work was supported by the Defense Advanced Research Projects Agency, Program Element 63226E.

the deposit and the overlaying water column, which can be modeled as a continuous source. However, we will not consider continuous release since its effects are negligible compared with that of impulsive release, over the time interval of interest, see [1]. We will consider several phenomena that affect the dispersion of released plumes: molecular diffusion, advective and turbulent flow.

### 3. TRANSPORT MODEL

The underwater transport of substances occurs as a superposition of many scales of motion; these include advective flow, turbulent flow and molecular diffusion. To include the effects of these phenomena we consider the following diffusion equation

$$\frac{\partial c}{\partial t} = \text{div}(\mathcal{K}\nabla \cdot c) - \nabla \cdot c(\mathbf{v}(t)). \quad (1)$$

where  $c$  is the diffusing substance concentration in units of  $\text{kg}/\text{m}^3$ ,  $\mathcal{K}$  is a matrix of diffusivity coefficients in units of  $\text{m}^2/\text{s}$ , and  $\mathbf{v}(t)$  is water velocity. In order to solve (1) we need to reduce problem to a corresponding problem in an isotropic medium, see [3].

#### Impulsive Source

An impulsive source is due to initial stirring up of previously deposited material or to instantaneous release of substances from bags ruptured upon impact. The resulting impact plume can be approximately modeled as an impulsive source at time  $t_0$  of strength  $\mu$  in  $\text{kg}$  units. Then, the solution of (1), detailed in [3], is

$$c(\mathbf{r}, t) = \frac{\mu}{(4\pi\kappa(t-t_0))^{3/2}} \cdot \exp\left\{-\frac{|\mathbf{r} - \mathbf{v}(t-t_0)(t-t_0) - \mathbf{r}_0|^2}{4\kappa(t-t_0)}\right\} \quad (2)$$

where  $\mathbf{r}_0$  is the source location, and  $\kappa$  is the isotropic diffusivity, see [3].

#### 3.1. DOR Scenario

In this part we describe the scenario of the DOR program. We use 20 sources (bags) spatially distributed with 2D Gaussian distribution with  $\sigma_s = 500\text{m}$ . We will consider impact plumes that can be modeled as impulsive sources.

It has been shown [1] that about 5% of the bag' contents material will be subject to advection. Thus, to model the release rate for each bag we use a beta distribution with parameters 1 and 19. Since each bag contains a mixture of different pollutants we assume that the mass of a particular pollutant is 1% of the bag contents and hence that 0.05% is the most probable

value of release rate. The density of released plumes is set to  $\rho = 1100\text{kg}/\text{m}^3$ . Using the above scenario the most probable release of particular pollutant would be approximately 336  $\text{kg}$ .

The advective flow in the DOR scenario is due to a mean flow of magnitude 0.5  $\text{cm}/\text{s}$  and a tidal current which rotates directionally with a period of 12.42 hours and magnitude in the range 2-5  $\text{cm}/\text{s}$ .

### 4. DETECTION AND PARAMETER ESTIMATION

#### 4.1. Measurement Model

To model the measurements, we suppose a spatially distributed array of  $m$  selective chemical sensors located at known positions  $\{\mathbf{r}_i, 1 \leq i \leq m\}$ . Then, the response of each sensor is

$$y(\mathbf{r}_i, t_k) = c(\mathbf{r}_i, t_k) + e(\mathbf{r}_i, t_k), \quad (3)$$

where  $c(\mathbf{r}_i, t)$  denotes the concentration of the substance (pollutant) of interest, and  $e(\mathbf{r}_i, t)$  is the measurement noise. The time samples will be assumed to be taken at uniformly spaced time points  $\{t_k = kT_s, 1 \leq k \leq p\}$ , where  $T_s$  is the sampling interval and  $p$  is the number of time samples.

We lump the measurement model (3) into a matrix form

$$\mathbf{y} = A(\boldsymbol{\theta})\boldsymbol{\mu} + \mathbf{e}, \quad (4)$$

where  $\mathbf{y}$  is an  $(mp)$ -dimensional measurement vector,  $A(\boldsymbol{\theta})$  is an  $mp \times n$  dimensional source-to-sensor transfer matrix,  $\boldsymbol{\theta}$  is a vector of unknown source and medium parameters,  $\boldsymbol{\mu} = [\mu_1, \dots, \mu_n]^T$  is a vector of source intensities,  $n$  is the number of sources, and  $\mathbf{e}$  is a vector of measurement noise.

For the remainder of the paper we will consider the case of known medium characteristics, i.e., the diffusivity matrix  $\mathcal{K}$ , velocity  $\mathbf{v}(t)$ , starting time of diffusion  $t_0$ , and source locations are all known *a priori*. This is a reasonable assumption since these characteristics will be measured according to the DOR plan [1]. Thus, we will omit the explicit dependence of  $A$  on  $\boldsymbol{\theta}$ .

#### 4.2. Source Detection

The detection of pollutant leakage from the disposal site is binary decision:  $H_0$ , only the bias term and noise are present, and  $H_1$ , the source is present as well, i.e. leakage of a pollutant occurs from some bags.

##### 4.2.1. GLR Detector (Known Physical Model)

This detector is based on the assumption that the solution (2) approximates the physical processes reasonably

well and that uncertainties in the model are due mainly to measurement noise.

The GLR test is given by the ratio

$$\text{GLR} = \frac{\sup_{\mu_j \geq 0, \sigma^2 > 0} \{\text{likelihood}(\mathbf{y})\}}{\sup_{\mu_j = 0, \sigma^2 > 0} \{\text{likelihood}(\mathbf{y})\}}, \quad (5)$$

where the numerator (denominator) on the r.h.s. corresponds to the likelihood function under  $H_1$  ( $H_0$ ).

The ML estimates  $\hat{\boldsymbol{\mu}}$  of the source intensities under  $H_1$  are computed as follows [4]. Let

$$\begin{aligned} \tilde{\boldsymbol{\mu}} &= [\tilde{\mu}_1, \dots, \tilde{\mu}_n]^T \\ &= (A^T A)^{-1} A^T \mathbf{y}. \end{aligned} \quad (6)$$

Then,

$$\hat{\boldsymbol{\mu}} = [\hat{\mu}_1, \dots, \hat{\mu}_n]^T, \quad (7)$$

$$\hat{\mu}_j = \begin{cases} 0 & \tilde{\mu}_j \leq 0 \\ \mathbf{g}_j^T A_l^T \mathbf{y} & \tilde{\mu}_j > 0 \end{cases} \quad (8)$$

where  $\mathbf{g}_j$  is the  $j$ th column of matrix  $(A_l A_l^T)^{-1}$ , and  $A_l$  is an  $mp \times l$  dimensional submatrix of matrix  $A$  whose columns correspond to positive estimates,  $\tilde{\mu}_j > 0$  ( $l$  is the number of positive components in  $\tilde{\boldsymbol{\mu}}$ ).

The maximum likelihood estimates of  $\sigma^2$  are

$$\hat{\sigma}^2 = \begin{cases} \frac{1}{mp} (\mathbf{y} - A\hat{\boldsymbol{\mu}})^T (\mathbf{y} - A\hat{\boldsymbol{\mu}}) & \text{under } H_1, \\ \frac{1}{mp} \mathbf{y}^T \mathbf{y} & \text{under } H_0. \end{cases} \quad (9)$$

Inserting the ML estimates  $\hat{\boldsymbol{\mu}}$  and  $\hat{\sigma}^2$  into the likelihood ratio (5) we get

$$\text{GLR} = \left( \frac{\mathbf{y}^T \mathbf{y}}{\mathbf{y}^T \mathbf{y} - \mathbf{y}^T P_{A_l} \mathbf{y}} \right)^{\frac{mp}{2}}, \quad (10)$$

where  $P_{A_l} = A_l (A_l^T A_l)^{-1} A_l^T$  is the projection matrix onto the column space of  $A_l$ .

Applying the monotonic transformation  $\frac{mp}{l} (x^{\frac{2}{mp}} - 1)$ , we redefine the GLR as

$$\text{GLR} = \frac{mp}{l} \frac{\mathbf{y}^T P_{A_l} \mathbf{y}}{\mathbf{y}^T \mathbf{y} - \mathbf{y}^T P_{A_l} \mathbf{y}}. \quad (11)$$

The detection decision is then made by comparing the GLR in (11) with a threshold  $\tau$ : if  $\text{GLR} > \tau$  reject  $H_0$ , otherwise accept  $H_0$ .

The computation of  $\tau$  requires knowledge of the probability distribution of GLR under  $H_0$ . In this case the GLR has a central F distribution with  $l$  and  $mp - l$  degrees of freedom,  $F_{l, mp-l}$ , where  $l$  is the number of positive components of the vector  $\tilde{\boldsymbol{\mu}}$ . This result is generalization of the single source case in [5]. Similarly, under  $H_1$  the GLR has a noncentral F distribution with

noncentrality factor  $\lambda = \boldsymbol{\mu}^T A^T A \boldsymbol{\mu} / \sigma^2$  and the same degrees of freedom.

Thus, the performance measures are

$$P_{\text{fa}} = \sum_{l=1}^n c_l \{1 - \Pr[F_{l, mp-l}(0) \leq \tau]\}, \quad (12)$$

$$P_{\text{d}} = \sum_{l=1}^n c_l \{1 - \Pr[F_{l, mp-l}(\lambda) \leq \tau]\}, \quad (13)$$

where  $c_l$  is probability that  $l$  of the source intensity estimates are positive.

#### 4.2.2. Mean Detector (Unknown Physical Model)

The mean detector makes less assumptions than the GLR about the model, hence it is useful when a reliable model is not available. This detector is computed using the statistic, see [6]

$$T = \frac{\sqrt{N}}{mp} \frac{1}{\sqrt{\mathbf{y}_0^T \mathbf{y}_0}} \sum_{i=1}^{mp} y_i, \quad (14)$$

where  $y_i$  is the  $i$ -th component of the measurement vector  $\mathbf{y}$ , the vector  $\mathbf{y}_0$  is a measurement vector obtained before the detection phase (in the absence of any signal), and  $N$  is the number of time samples used to obtain  $\mathbf{y}_0$ .

Under  $H_0$ ,  $T$  has Student's central  $t$  distribution with  $N$  degrees of freedom, see [7]. Note that as  $N$  increases, the distribution of  $T$  approaches Gaussian. Under  $H_1$ ,  $T$  has a non-central  $t$  distribution with  $N$  degrees of freedom and noncentrality factor  $\lambda$ , see [7].

The performance measures are given by

$$P_{\text{fa}} = 1 - \Pr[t_N(0) \leq \tau], \quad (15)$$

$$P_{\text{d}} = 1 - \Pr[t_N(\lambda) \leq \tau], \quad (16)$$

$$\lambda = \frac{(\frac{1}{mp} \sum_{i=1}^{mp} \xi_i)^2}{\sigma^2}, \quad (17)$$

where  $t_N(\cdot)$  denotes the cumulative  $t$  distribution with  $N$  degrees of freedom and non-centrality factor in parentheses,  $\tau$  is the decision threshold, and  $\boldsymbol{\xi} = A\boldsymbol{\mu}$ .

## 5. SENSOR ARRAY DESIGN

We apply the above results on detection performance to optimally design the array of sensors. The design is with respect to the system parameters defining the array configuration: the array radius  $r$ , number of sensors  $m$ , and number of time samples  $p$ .

We first propose a procedure for selecting the array parameters to achieve a desired detection probability  $P_{\text{d}}^*$  for a fixed  $P_{\text{fa}}$ . This procedure can be summarized

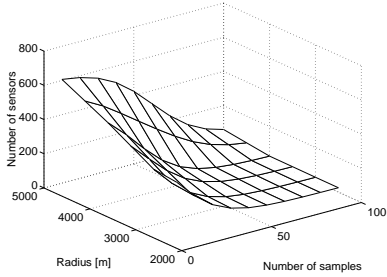


Figure 1: System parameters to achieve  $P_d^* = 95\%$ , GLR detector.

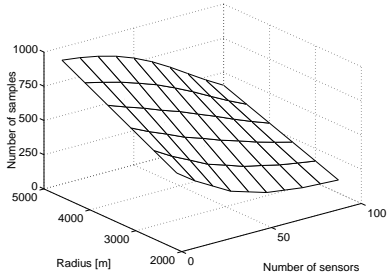


Figure 2: System parameters to achieve  $P_d^* = 95\%$ , mean detector.

as follows: set an acceptably small level of false alarm,  $P_{fa}$  (e.g. 5-10%), compute the corresponding threshold  $\tau$  as a function of  $mp$ , for this  $\tau$ , find combinations of  $m$  and  $p$  that give the desired  $P_d^*$  using (13) and (17), repeat the previous step for different values of array radius.

As a result, we obtain a set of possible choices for the array parameters  $m$ ,  $p$  and  $r$  that guarantees the desired  $P_d^*$ . This set can be presented as a surface in 3D space, see Figures 2 and 3. In both examples we used  $\text{SNR} = 3\text{dB}$ ,  $T_s = 1000\text{s}$ , and  $r = 2500\text{m}$ .

Using the above procedure we design a detector that guarantees the required  $P_{fa}$  and  $P_d^*$ .

An optimum selection of the number of sensors  $m$  and the number of time samples  $p$  can be done by minimizing a cost function  $C$  involving the system parameters, given the  $\text{SNR}$ ,  $P_{fa}$ , and desired  $P_d^*$ . For example,

$$C = C_1 m + C_2 p, \quad (18)$$

where  $C_1$  is the cost per sensor and  $C_2$  is the cost per time sample.

The optimal design is described by the following procedure: fix an array radius  $r$ , use the above algorithm to find the set of candidate choices for  $m$  and  $p$  that give the desired  $P_d^*$ , minimize  $C$  over that set, repeat the previous steps for different values of  $r$ . For any  $r$  this procedure yields the optimal choice of  $m$  and  $p$  and

the corresponding cost. Then, the final decision can be made by selecting the largest radius with acceptable cost.

## 6. SUMMARY

We have proposed detection algorithms for environmental monitoring of disposal sites in the deep ocean using chemical sensor arrays. The algorithms included the GLR and mean detectors. The GLR detector gives a better performance and is applicable when the physical model is reliable, while the mean detector is useful when a precise model is not available. We have analyzed the performance of both detectors using the probability of detection  $P_d$  and false alarm  $P_{fa}$ . We have also proposed algorithms for optimal array design assuming a variety of performance and cost requirements. The design included selection of number of sensors and time samples.

Future research will include improving performance by using flux sensors, dealing with non-selective sensors, investigating effects of more realistic tidal current, pollutant release and measurement models, and including adaptive sampling.

## 7. REFERENCES

- [1] P. J. Valent, and D. K. Young, Abyssal Seafloor Waste Isolation: Environmental Report, NRL/MR/7401-95-7576, Naval Research Laboratory, Stennis Space Center, MS, September 1995.
- [2] A. Nehorai, B. Porat, and E. Paldi, "Detection and Localization of Vapor-Emitting Sources," *IEEE Trans. on Signal Processing*, Vol. 43, No. 1, pp. 243-253, January 1995.
- [3] A. Jeremić, and A. Nehorai, "Design of Chemical Sensor Arrays for Monitoring Disposal Sites on the Ocean Floor," *submitted to IEEE Trans. on Oceanic Engineering*.
- [4] C. L. Lawson, and R. J. Hanson, *Solving Least Squares Problems*, Englewood Cliffs: Prentice Hall, 1974.
- [5] L. Scharf, and B. Friedlander, "Matched Subspace Detectors," *IEEE Trans. on Signal Processing*, Vol. 42, No. 8, pp. 2146-2157, August 1994.
- [6] J. C. Kiefer, *Introduction to Statistical Inference*, New York: Springer-Verlag, 1987.
- [7] C. R. Rao, *Linear Statistical Inference and Its Applications*, 2nd ed. New York: John Wiley & Sons, 1973.