

An approach to leak detection using wireless sensor networks at carbon sequestration sites

James Weimer^a, Bruce H. Krogh^{a,*}, Mitchel J. Small^b, Bruno Sinopoli^a

^a Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

^b Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:

Received 3 September 2010

Received in revised form 23 March 2012

Accepted 3 April 2012

Keywords:

CO₂ sequestration site monitoring

Leak detection

Wireless sensor network

ABSTRACT

This paper concerns the problem of detecting leaks at carbon sequestration sites through the monitoring of CO₂ levels using a wireless sensor network (WSN). By applying a basic linear dynamic model for an advection–diffusion process, a model-based detection strategy called the Iterative Partial Sequential Probability Ratio Test (IPSPRT) can be employed to detect and localize multiple leaks. A 3-D CO₂ transport model is employed to provide a proof of concept simulated evaluation of the IPSPRT against a windowed-average approach in terms of time-to-decision vs. probability of false alarm and probability of a missed alarm. Despite using only a basic linear dynamic model, the IPSPRT is shown to achieve exponentially increasing better time-to-decision than the windowed-average approach as the probability of false alarm and probability of a missed alarm are decreased. A test bed implementation is described and employed to evaluate the performance of the IPSPRT in the presence of common WSN errors.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Currently coal accounts for 25% of the world's energy supply and 40% of carbon emissions, and is likely to be a major source of electricity generation for the foreseeable future ([United States Secretary of Energy Steven Chu, 2009](#)). It has been proposed that CO₂ from coal power plants be sequestered in large underground geological formations for the purposes of enhanced oil recovery in current oil wells, collecting methane from deep un-minable coal beds, and general storage ([Wells et al., 2006](#); [Saripalli et al., 2006](#)). One viable means to ensure sequestration is to monitor the surface CO₂ concentration for changes not explained by the ambient fluctuations caused by seasonal and environmental forces such as respiration and photosynthesis ([Saripalli et al., 2006](#)). Other methods of ensuring CO₂ sequestration include subterranean CO₂ concentration monitoring and monitoring of trace gasses injected during sequestration ([Saripalli et al., 2006](#)). Due to the transport and dispersion of sequestered CO₂ in subsurface formations, monitoring (either on the surface or underground) may need to be performed over vast areas (hundreds of square kilometers) for years where a large number of potential CO₂ leak locations (sources) are possible ([Wells et al., 2006](#); [Saripalli et al., 2006](#)). Due to the inherent cost of installing a wired-network over the CO₂ sequestration site monitoring area ([Saripalli et al., 2006](#)), a wireless sensor

network (WSN) is the most feasible technology for performing large-scale, long-term multiple-leak detection.

Many methods have been proposed for the general problem of detecting potential leaks through noisy observations (see [Willsky, 1976](#); [Kailath et al., 1998](#); [Fox et al., 2007](#); [Weimer et al., 2011](#) and citations therein). These approaches are divided into model-based approaches and sensor-level approaches. Model-based approaches leverage knowledge of underlying dynamics to assist in leak detection, while sensor-level approaches simply monitor for changes in individual sensor measurements to detect leaks. Many researchers have shown that when an accurate model exists, model-based approaches provide better results than sensor-level approaches by leveraging the information gained through knowledge of the dynamics ([Willsky, 1976](#)); however, the primary shortcoming of most model-based approaches is that they do not scale well with the number of potential leaks.

The following section introduces the multiple-leak detection problem for CO₂ sequestration site monitoring and presents a model-based test capable of addressing the inherent scalability issues. [Section 3](#) formulates a CO₂ transport model that can be employed as part of the model-based test for identifying leaks. [Section 4](#) describes a 3-D CO₂ atmospheric transport model, developed to provide a proof of concept evaluation of the model-based test against a windowed-average test using simulation in [Section 5](#). [Section 6](#) presents a WSN test bed implementation and robustness analysis of the leak detector's performance with respect to common WSN errors. The concluding section summarizes the contributions of this paper.

* Corresponding author. Tel.: +1 412 243 6578; fax: +1 412 243 3890.

E-mail addresses: weimerj@kth.se (J. Weimer), krogh@ece.cmu.edu (B.H. Krogh).

2. Problem formulation

At CO₂ sequestration sites, ensuring the CO₂ remains sequestered requires monitoring the CO₂ concentrations (or other trace gases injected during sequestration Saripalli et al., 2006) over large areas. When the monitoring area is divided into several potential leak locations, the leak detection problem requires not only identifying whether CO₂ is leaking, but also identifying where the leak occurs. A test for performing leak detection makes a decision whether there is a leak at each location while bounding the probability of a *false alarm*, the probability of a *missed alarm*, and the probability of a *missed leak*. In our formulation, a false alarm occurs when no leaks exist at any locations, but the test declares there is at least one leak. Conversely, a missed alarm occurs when a leak exists and the test decides there are no leaks. A missed leak occurs when the test correctly decides that some leaks exist, but does not identify all the locations where leaks exist.

One common sensor-level approach to leak detection is a *windowed-average test*. The windowed-average test assumes no knowledge of the underlying dynamics. Each sensor independently measures and averages its respective CO₂ concentration measurements over a window of the most recent measurements, where the window size defines the time-to-decision. A leak is declared to exist if for any sensor, the windowed-average concentration exceeds an a priori specified threshold; otherwise, it is decided that no leaks exists. While the windowed-average test is easy to implement, identifying the window size and test threshold to achieve a desired level of performance in terms of the probabilities of false alarm, missed alarm, and missed leak may not be possible since the threshold varies significantly with the sensor location, climate, season, and changing environmental dynamics (such as wind speed, atmospheric stability).

To improve detection performance, model-based approaches leverage knowledge of the environmental dynamics. Our model-based test, the Iterative Partial Sequential Probability Ration Test (IPSPRT), uses a basic linear dynamic model relating the effects of potential leaks on the CO₂ sensor measurements. We note that the purpose of the IPSPRT model is not to precisely model the dynamic effects of CO₂ atmospheric fate and transport and CO₂ leak rates, but rather to capture the general trends in a completely tractable framework. Keeping this in mind, we model the CO₂ concentration dynamics using a state-space model of the form

$$x_{k+1} = A_k x_k + B_k (\Gamma_k z_k + e_k) + w_k \quad (1)$$

$$q_k = C x_k + v_k,$$

where each element of the vector x_k represents the CO₂ concentration in *parts-per-million* (PPM) at a specific location and time, the elements of q_k are the measured CO₂ concentration values at time k in PPM, and each element of z_k and e_k represents the expected CO₂ flux rates at a specific location in PPM-meters-per-second caused by leaks and environmental processes¹ (such as respiration and photosynthesis), respectively.²

Γ_k is a diagonal binary matrix representing which of the potential leak locations are actually leaking. A unit entry in Γ_k indicates that there is a leak at the corresponding leak location, while a zero entry denotes that no leak is present. A_k and B_k specify a lumped parameter model of the environmental dynamics, and C is a binary matrix that identifies the sensor locations, that is, which element of x_k are being measured. w_k and v_k are zero-mean uncorrelated

¹ For the purposes of this work, the expected environmental CO₂ generation, e_k , is assumed to be known. For detailed studies on the modeling background CO₂ generation, we direct the interested reader to (Yang et al., 2011, 2011) and the citations within.

² Model (1) is derived in the following section.

Gaussian signals that account for the environmental modeling uncertainty (including background environmental CO₂ generation uncertainty) and sensor measurement noise, respectively.

The form of the IPSPRT used in this paper assumes all CO₂ generation rates are constant, that is,

$$z_{k+1} = z_k + h_k, \quad (2)$$

where h_k is a Gaussian zero-mean signal that accounts for the CO₂ generation rate uncertainty.³ Assuming a lossy WSN is used to gather the measured CO₂ concentrations for centralized processing, the received sensor observations are modeled according to

$$y_k = \Lambda_k q_k, \quad (3)$$

where Λ_k is a binary selection matrix that captures the effects of packet loss inherent in WSNs, that is, Λ_k selects which elements of q_k comprise the elements of y_k . By combining the models in (1)–(3), the resulting model relating the potential CO₂ leak locations to the collected CO₂ concentration measurements is

$$\begin{bmatrix} x_{k+1} \\ z_{k+1} \end{bmatrix} = \begin{bmatrix} A_k & B_k \Gamma_k \\ 0 & I \end{bmatrix} \begin{bmatrix} x_k \\ z_k \end{bmatrix} + \begin{bmatrix} B_k & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} e_k \\ w_k \\ h_k \end{bmatrix} \quad (4)$$

$$y_k = [\Lambda_k C \quad 0] \begin{bmatrix} x_k \\ z_k \end{bmatrix} + \Lambda_k v_k.$$

We again remark for clarity that the model in (4) is not intended to precisely simulate the environmental fate and transport of CO₂, but rather to just capture the general effects. It will be shown in Section 4 that the IPSPRT performs well even when the dynamics for atmospheric advection-diffusion and CO₂ are simulated using a more complex model (which will be described in Section 4).

Using the model in (4), the large-scale multiple-leak detection problem results in a multiple hypothesis testing problem between all possible combinations of potential leaks. Even under the simplifying assumptions that the potential leaks are persistent (leak continuously once becoming active), if there are J leak locations and leaks can become active at K time steps, the number of potential combinations of leaks, $S(J, K)$, is

$$S(J, K) = (K + 1)^J. \quad (5)$$

Even for a small problem containing 30 different potential leak locations ($J=30$) and a single time when leaks can become active ($K=1$), the number of different potential combinations of leaks is $2^{30} \simeq 10^9$. As the number of leaks and times at which leaks can become active increase, testing between all possible combinations of leaks becomes infeasible.

To avoid this exponential computational growth, the IPSPRT iteratively solves a sequential multiple hypothesis testing problem for a partial subset of the possible leak combinations by sequentially performing *aggregate source detection* followed by *active source identification*. The remainder of this section discusses the IPSPRT in detail, following the flow chart provided in Fig. 1.

As illustrated by block A in Fig. 1, the IPSPRT is initialized to assume that leaks can only become active at time zero, as denoted by the hypothesis time, $K_H \leq k$, when the current time $k=0$. Additionally, the IPSPRT assumes that none of the J potential leaks were previously active, $\Gamma_{-1}=0$. After initialization, the CO₂ measurements are collected using a WSN and the environmental model is formulated (block B in Fig. 1). Using the model and collected CO₂

³ More sophisticated models for the CO₂ generation rates can be used in place of the constant-generation model in (2), but it will be shown in Section 5 that the IPSPRT, assuming constant generation, performs well even for non-static CO₂ generation rates.

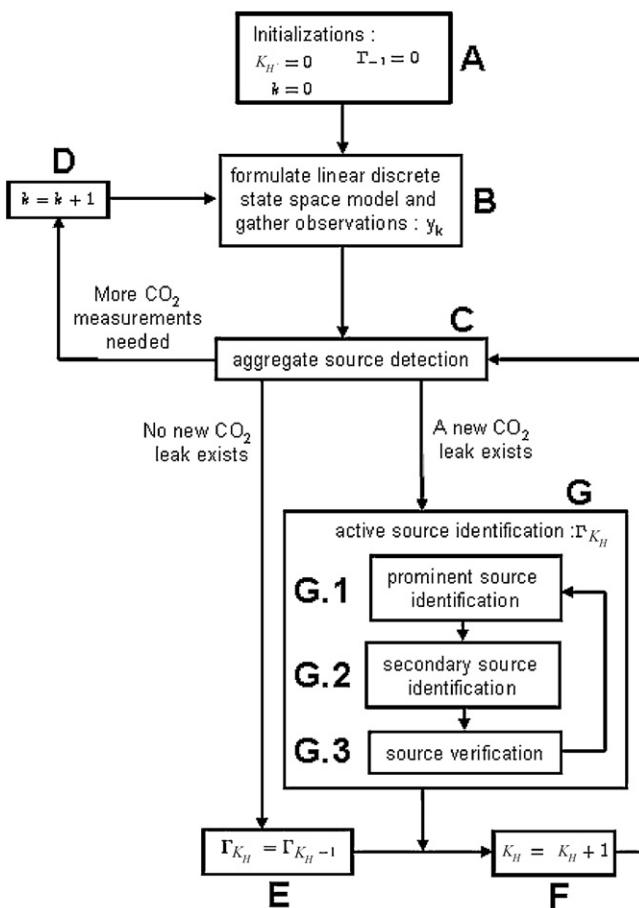


Fig. 1. IPSPRT flowchart.

measurements, aggregate source detection (block C) attempts to decide whether *no leaks* or *some leaks* became active at the hypothesis time, K_H . This decision is made such that both the probability of false alarm and probability of missed alarm are bounded using the Sequential Probability Ratio Test (SPRT) developed by Wald (1947). The test for aggregate source detection can make one of three decisions: no new leaks began at time K_H , some leak began at time K_H , or more CO₂ measurements are needed to make a decision.

In Fig. 1, if the test for aggregate source detection (block C) decides that more measurements are needed, then at the next time step ($k = k + 1$ as denoted by block D), additional CO₂ measurements are gathered and the test continues testing for leaks that began at the hypothesis time K_H . If the test decides no leaks began at the current hypothesis time, K_H , then it is assumed that no additional leaks became active at time K_H (i.e. $\Gamma_{K_H} = \Gamma_{K_H-1}$ in block E), and the hypothesis time is incremented, $K_H = K_H + 1$ (block F). The test for aggregate source detection is then performed using the incremented value of K_H . If the test detects a leak then it is assumed that some leak became active at time K_H and the IPSPRT proceeds to active source identification, represented by block G.

Active source identification consists of three parts: *prominent source identification*, *secondary source identification*, and *source verification* (as denoted by blocks G.1, G.2, G.3, respectively in Fig. 1). Prominent source identification identifies the most likely leak locations through an iterative likelihood maximization. After identifying the most likely leak locations, secondary source identification (under the assumption that the most likely leak locations actually contain leaks) identifies the leak locations which must also be checked to meet the performance requirement on the probability of a missed leak. We call these leak locations the *nuisance*

locations. Both the likely leak locations and the nuisance locations are checked as part of source verification. If after performing source verification the probability of a missed leak still meets the performance criteria, then active source identification terminates by logging the active sources in Γ_{K_H} , the hypothesis time is incremented ($K_H = K_H + 1$), and aggregate source detection is performed; otherwise, active source identification is performed again until the probability of a missed leak meets the performance criteria after performing source verification, that is, personnel are sent to the field and more accurate local monitoring is performed to determine if there are leaks at the identified locations.

3. Advection–diffusion model

Although the IPSPRT is well suited to handle the computational issues imposed by performing large-scale leak detection, it requires a linear dynamic model describing, in a general sense, the leaks on the measured CO₂ concentrations in the form of (4). This requires a model describing the transport of CO₂ from potential leaks to the sensors through the atmosphere. In general, gas transport through the atmosphere is a complex process with the most accurate United States Environmental Protection Agency (EPA) approved models requiring physical terrain data, detailed meteorological data for wind characterization and atmospheric stability, and source emissions data for the chemical(s) of interest (in the case of ground-level ambient CO₂, the principal sources occur naturally, including respiration from surface vegetation and soil microbes) (Federal Register). While these models are considered the most accurate (Federal Register), identifying the necessary parameters requires significant empirical studies and remains an open research issue (Sharan et al., 1996; Sharan and Yadav, 1998; Hernandez et al., 1991; Seinfeld and Pandis, 1998; Rao, 1983; Turner, 1994). Due to the complexity of the EPA approved models, a simplified linear dynamic model, suitable for the IPSPRT, is developed for use in a first demonstration.

Virtually all models for atmospheric fate and transport originate from a first-principles model of an advection–diffusion process (Seinfeld and Pandis, 1998),

$$\frac{\delta c(p, t)}{\delta t} + \phi(p, t) \frac{\partial c(p, t)}{\partial p} = \alpha(p, t) \frac{\partial^2 c(p, t)}{\partial p^2}, \quad (6)$$

where $c(p, t)$ denotes the concentration of CO₂ in *part per million* (PPM) as a function of space and time, $p = [x, y, z]$ is the 3-D location vector, t is time, $\phi(p, t) = [\phi_x(p, t), \phi_y(p, t), \phi_z(p, t)]^T$ and $\alpha(p, t) = [\alpha_x(p, t), \alpha_y(p, t), \alpha_z(p, t)]^T$ are the advection and dispersion coefficients, respectively, in units of m/s and m²/s. The surface boundary condition is

$$\left(\phi_z(p, t)c(p, t) - \alpha_z(p, t) \frac{\delta c(p, t)}{\delta z} \right) |_{p=(x,y,0)} = \lambda(x, y, 0, t) + \epsilon(x, y, 0, t), \quad (7)$$

where $\lambda(x, y, z, t)$ and $\epsilon(x, y, z, t)$ represent the CO₂ surface leak rate and the surface CO₂ absorption/expulsion rate, respectively, each known to be non-zero only at the surface ($z = 0$) and having units of PPM-meters-per-second.

To generate a model of the form in (4), we spatially discretize the partial differential equation in (6) using an Euler's approximation⁴ as in Hernandez et al. (1991) with $x - y - z$ granulations of Δ_x , Δ_y ,

⁴ A TVD-based discretization could also be incorporated in place of an Euler's approximation (Tannehill et al., 1997).

and Δ_z , respectively, and write the continuous-time above-surface ($z \geq 0$) advection–diffusion model as

$$\begin{aligned} \frac{\delta c(x, y, z, t)}{\delta t} = & \alpha_x(t) \left(\frac{c(x + \Delta_x, y, z, t) - 2c(x, y, z, t) + c(x - \Delta_x, y, z, t)}{\Delta_x^2} \right) \\ & + \alpha_y(t) \left(\frac{c(x, y + \Delta_y, z, t) - 2c(x, y, z, t) + c(x, y - \Delta_y, z, t)}{\Delta_y^2} \right) \\ & + \alpha_z(t) \left(\frac{c(x, y, z + \Delta_z, t) - 2c(x, y, z, t) + c(x, y, z - \Delta_z, t)}{\Delta_z^2} \right) \quad (8) \\ & - \phi_x(t) \left(\frac{c_x - c(x, y, z, t)}{\Delta_x} \right) - \phi_y(t) \left(\frac{c_y - c(x, y, z, t)}{\Delta_y} \right) \\ & - \phi_z(t) \left(\frac{c_z - c(x, y, z, t)}{\Delta_z} \right) + \frac{\lambda(x, y, z, t)}{\Delta_z} + \frac{\epsilon(x, y, z, t)}{\Delta_z} \end{aligned}$$

where

$$\begin{aligned} c_x &= \begin{cases} c(x + \Delta_x, y, 0, t) & \text{if } \phi_x(t) > 0 \\ c(x - \Delta_x, y, 0, t) & \text{if } \phi_x(t) < 0 \end{cases} \\ c_y &= \begin{cases} c(x, y + \Delta_y, 0, t) & \text{if } \phi_y(t) > 0 \\ c(x, y - \Delta_y, 0, t) & \text{if } \phi_y(t) < 0 \end{cases} \\ c_z &= \begin{cases} c(x, y, z + \Delta_z, t) & \text{if } \phi_z(t) > 0 \\ c(x, y, z - \Delta_z, t) & \text{if } \phi_z(t) < 0 \end{cases} \end{aligned}$$

The advection parameters in (8) denote the wind speed and direction; in the physical world, determining this parameter typically requires approximation since the wind is continuously changing. The eddy diffusion parameters are characterized by several parameters, including the crosswind intensity, vertical height, atmospheric stability, and wind speed, all of which vary (Rao, 1983; Seinfeld and Pandis, 1998; Sharan and Yadav, 1998). To generate the IPSPRT model, we assume that the horizontal advection dominates vertical advection, the horizontal wind vectors are spatially uniform, and the eddy diffusion parameters are

$$\begin{aligned} \alpha_x(t) &= \frac{10}{|\phi_y(t)| + 1} \\ \alpha_y(t) &= \frac{10}{|\phi_x(t)| + 1}, \end{aligned} \quad (9)$$

which represents the trend that eddy diffusion parameter values decrease as the wind speed increases in the orthogonal direction. Additionally, since the WSN is only deployed on the surface, there will be no vertical concentration monitoring assumed by the IPSPRT. As a heuristic, we account for the unobservable vertical effects of advection and diffusion by assuming that the CO₂ surface flux rates are two-thirds their expected values.⁵

Applying these approximations, the spatially discretized 3-D advection–diffusion model described by (8) can be written in a surface plane ($z=0$) 2-D advection–diffusion model, with suitable characteristics for the IPSPRT, as

$$\begin{aligned} \frac{\delta c(x, y, 0, t)}{\delta t} = & \alpha_x(t) \left(\frac{c(x + \Delta, y, 0, t) - 2c(x, y, 0, t) + c(x - \Delta, y, 0, t)}{\Delta^2} \right) \\ & + \alpha_y(t) \left(\frac{c(x, y + \Delta, 0, t) - 2c(x, y, 0, t) + c(x, y - \Delta, 0, t)}{\Delta^2} \right) \quad (10) \\ & - \phi_x(t) \left(\frac{c_x - c(x, y, 0, t)}{\Delta} \right) - \phi_y(t) \left(\frac{c_y - c(x, y, 0, t)}{\Delta} \right) \\ & + \frac{2}{3\Delta_z} \lambda(x, y, 0, t) + \frac{2}{3\Delta_z} \epsilon(x, y, 0, t) \end{aligned}$$

⁵ This modeling heuristic captures the general trend that as CO₂ is released a portion does not remain in the observable surface plane.

By applying the same process as Hernandez et al. (1991) to (10), the 2-D advection–diffusion process can be written in a continuous-time state-space model as

$$\frac{\delta x(t)}{\delta t} = A(t)x(t) + B(t)(u(t) + e(t)), \quad (11)$$

where $x(t) \in R^N$ is the row-by-column concatenation of the planar monitoring area, $A(t) \in R^{N \times N}$ is the lumped parameter model governing the time evolution of $x(t)$, the vector of environmental background generation/absorption rates is denoted by $e(t)$, and $u(t) \in R^N$ is the vector of leak rates at time t .⁶ In (11), each element of $x(t)$ represents the concentration over the corresponding discretized area at time t . The continuous state-space model in (11) is discretized according to the sensor sampling period, resulting in a discrete-time state space model for the advection diffusion process

$$x_{k+1} = A_k x_k + B_k(u_k + e_k) + w_k. \quad (12)$$

Similar to (11), each element of x_k in (12) represents the concentration over the corresponding discretized area at time k . The lumped-parameter model (12) is used in (4) to describe the transport of CO₂ from leaks to sensor locations, where $u_k = \Lambda_k z_k$.

The model developed in this section describing the transport of CO₂ is designed to capture the general trends occurring in atmospheric fate and transport. The model in (12) is not intended to accurately simulate the CO₂ concentrations, but it will be shown in Section 5 that when this basic model is employed by the IPSPRT, accurate detection and identification is achieved. Before evaluating the IPSPRT, the following section describes how the eddy diffusion parameters are determined and describes the environmental simulator employed in the evaluation.

4. Environmental simulation

In this study, we are concerned with short-range transport (a few kilometers) under non-steady-state conditions, and wish to also account for the effects of the change in the ambient CO₂ concentration. To address these concerns, this section is divided into three subsections describing the short-range transport parameters, the ambient CO₂ modeling, and the 3-D environmental simulator, respectively.

4.1. Short-range transport parameters

Both the 3-D transport model in (8) and the IPSPRT model in (12) require known advection and eddy-diffusion parameters. To identify the advection and eddy-diffusion parameters in (6) requires the following user inputs: mixing height, monitoring area size, time of day, season, cloud cover, terrain (either grass or forest), and the wind vector at 10 m above the surface. Using the user-defined inputs, the Pasquill stability class can be approximated using the approximations in Table 1 (Seinfeld and Pandis, 1998), where the Pasquill stability class provides a qualitative measure of how turbulent the atmosphere is vertically. In Table 1, class F denotes the most stable atmosphere, class A is the most unstable atmosphere, and class D is a neutral atmosphere. Based on the Pasquill stability class, the Monin–Obukov length, L , is determined using the straight-line approximation (Seinfeld and Pandis, 1998) as

$$L = \frac{1}{a + b \log(z_0)}, \quad (13)$$

⁶ The element-wise comparison between the lumped-parameter model, $A(t)$, in (11) and the spatially discretized partial differential equation in (10) is provided in Appendix C of Weimer (2010).

Table 1
Estimation of Pasquill stability classes.

Wind speed at 10 m (m/s)	Solar radiation			Night cloud cover	
	Strong	Moderate	Slight	≥50%	≤38%
<2	A	A–B	B		Very stable
2–3	A–B	B	C	E	F
3–5	A–B	B	C	D	E
5–6	B	B–C	C	D	E
>6	C	D	D	D	D

Table 2
Coefficients for straight line approximation to Monin–Obukov distance in (13).

Pasquill stability class	$a(\text{m}^{-1})$	$b(\text{m}^{-2})$
A	−0.096	0.029
B	−0.037	0.029
C	−0.002	0.018
D	0	0
E	0.004	−0.018
F	0.035	−0.036

where z_0 is the surface roughness for a given terrain (Seinfeld and Pandis, 1998) and a and b are constants defined for different Pasquill stability classes in Table 2 (Seinfeld and Pandis, 1998).

To identify the advection and eddy diffusion parameters requires the wind speed at different heights and the atmospheric friction velocity. The vertical wind speed is generally approximated using the wind-speed power-law formulation (Heinsohn and Kabel, 1999)

$$U(z) = U_{10} \left(\frac{z}{10} \right)^p \quad (14)$$

where U_{10} is the wind speed at 10 m above the surface, and p is a unitless value based on the Pasquill stability class according to Table 3 (Heinsohn and Kabel, 1999). The friction velocity, u_* , can be approximated using the Pasquill stability class according to Seinfeld and Pandis (1998),

$$\frac{u_*}{\kappa U_{10}} = \begin{cases} \left[\ln \left(\frac{10}{z_0} \right) + \frac{4.7}{L} (10 - z_0) \right]^{-1} & \text{if: stable (E,F)} \\ \left[\ln \left(\frac{10}{z_0} \right) \right]^{-1} & \text{if: neutral (D)} \\ \left[\ln \left(\frac{10}{z_0} \right) + \ln \left[\frac{(\eta_0^2 + 1)(\eta_0 + 1)^2}{(\eta_r^2 + 1)(\eta_r + 1)^2} \right] + 2[\tan^{-1} \eta_r - \tan^{-1} \eta_0] \right]^{-1} & \text{if: unstable (A, B, C)} \end{cases} \quad (15)$$

where κ is Karman's constant and

$$\eta_0 = \left[1 - 15 \frac{z_0}{L} \right]^{1/4} \quad \text{and} \quad \eta_r = \left[1 - \frac{150}{L} \right]^{1/4}.$$

Using the friction velocity, Pasquill stability class, and wind speed profile, the vertical eddy diffusion parameter in the surface layer is Seinfeld and Pandis (1998)

$$\alpha_z(p, t) = \frac{\kappa u_* z}{\theta(z/L)} \quad (16)$$

Table 3
Wind speed power-law exponent, p , in (14).

Pasquill stability class	A	B	C	D	E	F
p	0.07	0.07	0.10	0.15	0.35	0.55

Table 4
 CO_2 generation rates, ϵ , vs. time of day ($\text{mol m}^{-2} \text{s}^{-1}$ of CO_2).

Terrain	Photosynthesis	Respiration
Forest	$40 \pm 2 \times 10^{-9}$	$-40 \pm 2 \times 10^{-9}$
Grassland	$30 \pm 2 \times 10^{-9}$	$-30 \pm 2 \times 10^{-9}$

where

$$\theta \left(\frac{z}{L} \right) = \begin{cases} 1 + 4.7 \frac{z}{L} & \text{if: stable (E,F)} \\ 1 & \text{if: neutral (D)} \\ \left[1 - 15 \frac{z}{L} \right]^{-1/2} & \text{if: unstable (A, B, C)} \end{cases}.$$

From the vertical eddy diffusion parameter, the horizontal eddy diffusion parameters can be determined according to Fatehifar et al. (2006). The CO_2 transport model described above is based on multiple common approximations. Since this CO_2 transport model is intended for only short range transport, the Coriolis effect and upper level transport (above the surface layer) are neglected.

4.2. Ambient CO_2 modeling

The ambient or background CO_2 concentration is known to vary daily based on photosynthesis and respiration rates of the surrounding vegetation. Although the rates of respiration and photosynthesis are generally not equal (with photosynthesis often consuming more CO_2 than respiration expels), vegetation respiration and photosynthesis are known to exchange approximately

$$\begin{aligned} &\text{if: stable (E,F)} \\ &\text{if: neutral (D)} \\ &\text{if: unstable (A, B, C)} \end{aligned} \quad (15)$$

the same amount of carbon annually (Smith, 1981). Currently, the atmosphere contains (on average) 390 PPM CO_2 , and is known to fluctuate approximately 15 PPM annually (depending on the surrounding environment) (Larcher, 1995).

Modeling of the ambient CO_2 dynamics is a topic of continued research (e.g. Yang et al., 2011, 2011 and citations within). In this study, to evaluate the ISPRT, we assume the expected ambient CO_2 generation/absorption rate of the surface is known (i.e. the expected value of $\epsilon(x, y, 0, t)$ is known in (8)). As a proof of concept evaluation of the IPSRT, which is not intended

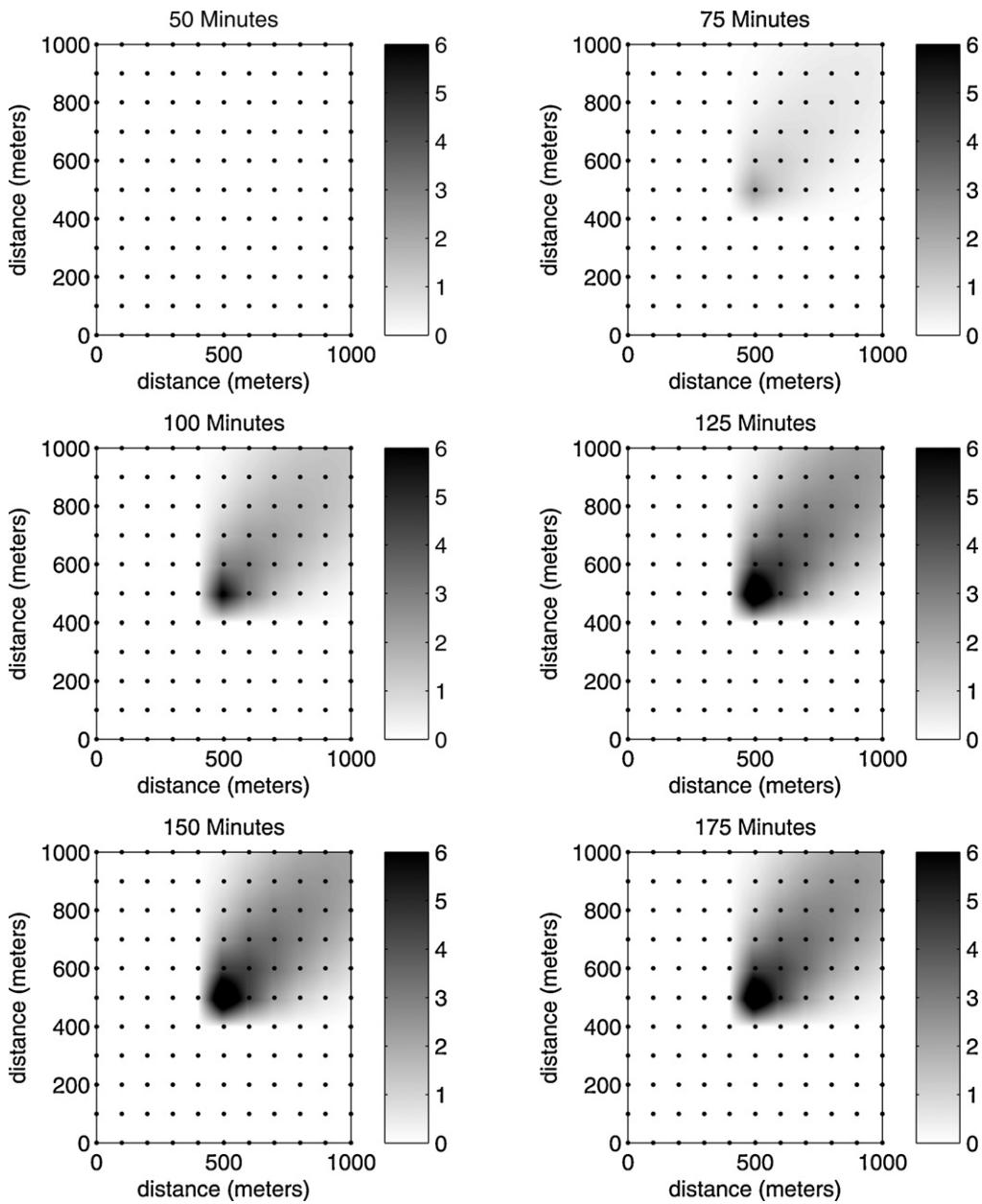


Fig. 2. normalized CO₂ concentrations in PPM for a ramping source at 50, 75, 100, 125, 150, and 175 min.

to satisfy the EPA testing requirements, we develop a simple model for the ambient CO₂ concentration. To model the fluctuation in ambient CO₂, the respiration and photosynthesis rates given in Table 4 are used. The rates in Table 4 are chosen such that the expected fluctuation in annual CO₂ concentration is 15 PPM based on the general assumption that plants expel CO₂ at night (through respiration) and consume CO₂ during the day (through photosynthesis). These rates, $\epsilon(x, y, 0, t)$, along with the leak rates, $\lambda(x, y, 0, t)$, define the boundary condition in (6).

While the model for ambient CO₂ concentrations employed in this study is simplified and does not necessarily capture the full effect, improved models can be incorporated as part of the IPSPRT since the natural CO₂ generation and absorption present in the environment are merely inputs to the IPSPRT. We remark that for field-testing purposes, improved models of ambient CO₂ concentration levels would be needed.

4.3. Environmental simulator

To provide a proof of concept evaluation of the IPSPRT, a basic 3-D environmental simulator is employed based on the 3-D advection-diffusion model in (8) and the parameter values provided earlier in this section. To generate spatial CO₂ concentrations at specific locations and times, a spatial Euler's

Table 5
CO₂ transport model parameter values.

Parameter	Value
Mixing height	1000 m
Monitoring area	1000 m × 1000 m
Time of day	Midnight
Season	Summer
Cloud cover	Clear
Wind vector	[330] m/s
Terrain	Grass

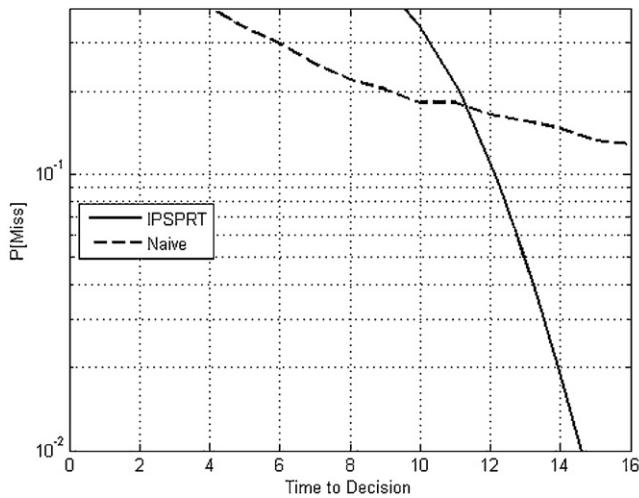


Fig. 3. Probability of miss vs. time-to-decision for probability of false alarm = 0.10.

discretization (Melman, 1997) is used in conjunction with a Taylor series approximation for temporal discretization to generate the 3-D CO₂ concentration profile over time. For evaluation purposes, the parameter values in Table 5 are assumed for the CO₂ transport model described above assuming $\Delta_x = \Delta_y = 10$ m and $\Delta_z = 1$ m. Fig. 2 presents the simulated CO₂ concentrations normalized to the expected ambient CO₂ concentration at times 50, 75, 100, 125, 150, and 175 min, assuming surface leaks become active at 60 min and their rates increase linearly until reaching a saturation rate of 100 PPM-meters-per-second at time 120 min. The subplots in Fig. 2 illustrate that before the leak begins, the only difference between the expected ambient CO₂ concentration and actual CO₂ concentration is due to the random variations in the photosynthesis and respiration rates. The subplots in Fig. 2 illustrate the effects of a leak on the spatial distribution of CO₂ concentrations over time. Using the CO₂ transport model developed in this subsection, the following section evaluates the performance of the IPSPRT.

5. Performance evaluation

For comparison, we evaluate the performance of the IPSPRT and the windowed-average test described in Section 2 in terms of time-to-decision, probability of false alarm, and probability of miss. The CO₂ concentration data obtained at each sensor is generated using the CO₂ transport model described in the previous section assuming the parameter values in Table 5, with a horizontal discretization of 100 m and a vertical discretization of 20 m. One thousand (1000) Monte Carlo runs of the CO₂ transport process are performed assuming a single leak exists which starts at a rate of zero PPM per second and saturates 1 h later with a random rate drawn from a normal distribution with mean 100 PPM per second and standard deviation of 50 PPM per second, where any leaks resulting in a negative leak rate are discarded. The sensor noise is assumed to be zero-mean Gaussian with a standard deviation of 1 PPM.

Since an a priori threshold is unknown for the windowed-average test, to compare its performance, 8000 different test thresholds ranging from -4 to 4 and 35 window sizes (time-to-decisions) ranging from 1 to 35 time steps are evaluated. A false alarm occurs when, for a specific threshold and time-to-decision, a leak is incorrectly detected. Similarly, a missed alarm occurs when a leak goes undetected. For the IPSPRT, the probability of false alarm, α , and miss, β , are design parameters and the average time-to-decision is calculated for all combinations of $\alpha, \beta \in \{0.01,$

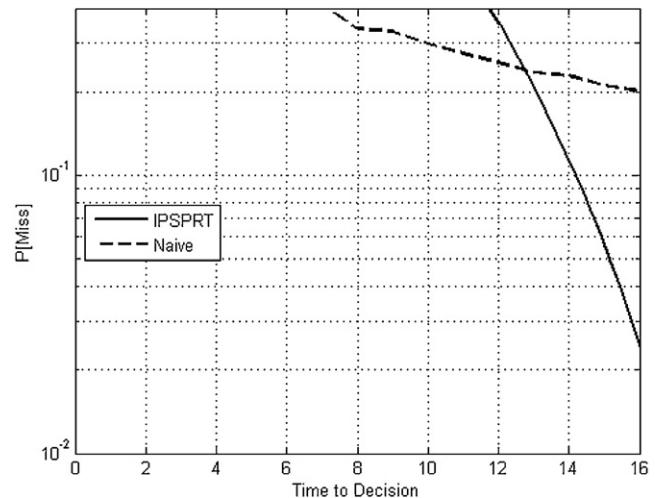


Fig. 4. Probability of miss vs. time-to-decision for probability of false alarm = 0.01.

0.02, 0.05, 0.10, 0.20, 0.50} (using the same Monte Carlo runs as the windowed-average test).

Fig. 3 shows the expected probability of miss vs. the time-to-decision for each test when the maximum probability of false alarm is assumed to be 0.10 ($\alpha = 0.10$). In Fig. 3 the dotted line and solid line represent the windowed-average test and the IPSPRT, respectively. One test is preferred over the other test if for a given probability of miss, it results in the smallest time-to-decision. Under this criterion, Fig. 3 suggests that if a probability of miss above 0.108 is acceptable, then the windowed-average test is the better option because it requires fewer time steps to make a decision (as denoted by the dotted line being below both the dashed and solid line). If the desired probability of miss is less than 0.108, then the IPSPRT should be employed. The reason the IPSPRT does not always perform better than the windowed-average test is due to the conservative decision thresholds obtained through Wald's approximation associated with performing the sequential probability ratio test (SPRT) (Wald, 1947). The results in Fig. 3 illustrate that the complexity of the IPSPRT test pays dividends as the desired probability of miss decreases.

Fig. 4 provides the results for when the maximum probability of false alarm is decreased to 0.01. As in Fig. 3, the dotted line and solid line in Fig. 3 represent the windowed-average test and the IPSPRT,

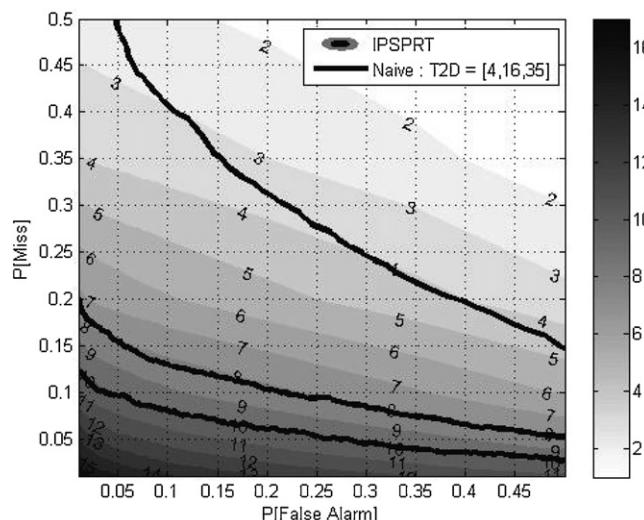


Fig. 5. Receiver operator characteristic vs. time-to-decision.

respectively. Similar to the results in Fig. 3, the results in Fig. 4 indicate the windowed-average test is preferred when the acceptable probability of miss is above 0.115, otherwise the IPSPRT is preferred. Comparing the results in Fig. 4 to the results in Fig. 3, we find that when the maximum probability of false alarm is decreased, the range of desired probability of miss where the IPSPRT is preferred increases.

The results in both Figs. 3 and 4 suggest that as the acceptable probability of miss decreases, the number of time steps required by the windowed-average test grows exponentially when compared to the IPSPRT. To illustrate this point, Fig. 5 shows a graph of the probability of false alarm vs. the probability of miss for each test, known as the receiver-operator characteristic (ROC). In Fig. 5, the underlying shaded contour plot represents integer values of the average time-to-decision for the IPSPRT, ranging from 2 to 17 for probability of false alarm and probability of miss ranging from 0.01 to 0.50. The color bar on the right and the integer values on the plot indicate the IPSPRT time-to-decision. The solid lines in Fig. 5 represent the windowed-average test, where the lines from highest to lowest denote the time-to-decisions of 4, 16, and 35, respectively.

To relate the results in Fig. 5 to the results in Fig. 3, the middle solid line (corresponding to a time-to-decision of 16 for the windowed-average test) represents all combinations of probability of false alarm and probability of miss where a time-to-decision of 16 can be achieved for the windowed-average test. When the probability of false alarm equals 0.10, the minimum probability of miss that can be achieved is 0.15 for the windowed-average test, assuming a window size of 16. In Fig. 3, these values correspond to the probability of miss attainable by each strategy when the time-to-decision is 16. A similar relation can be made between the results in Figs. 5 and 4 when the probability of false alarm equals 0.01 and the time-to-decision equals 16.

In Fig. 5, the results for a time-to-decision of 35 are plotted to illustrate the marginal savings in time-to-decision for the IPSPRT when compared to the windowed-average test. To achieve a probability of false alarm of 0.05 while also achieving a probability of miss of 0.10, the windowed-average test requires a window size (time-to-decision) of 35 while the IPSPRT averages a time-to-decision of 10. As expected, for both tests, as the probability of false alarm and probability of miss decrease, the time-to-decision increases. As would be expected, more observations are needed to make a more accurate decision. To summarize, despite using only a 2-D model for the advection–diffusion process in the IPSPRT, the performance in terms of time-to-decision performs better than a windowed-average approach. This suggests that the IPSPRT is robust with respect to moderate modeling errors.

In this study, only a single sensor density was assumed. Determining the sensor density that yields a desired detection level depends on the desired leak rate, environmental dynamics, background CO₂ stability, and desired granularity of leak detection. While this work does not address the issues of sensor placement, previous work on sensor placement (Weimer et al., 2008; Srivastava et al., 2009; Fox et al., 2007) could be used to estimate the most cost-effective network that ensures a high probability of detection.

In this evaluation, we observe that the IPSPRT provides better performance than a simple sensor-level approach, despite the use of a 2-D model and an incorrect assumption that the sources are constant. Moreover, the IPSPRT does not require an a priori specification of the decision threshold (which may not be possible), unlike simple sensor-level approaches. For the CO₂ concentration monitoring problem, the IPSPRT is shown to be robust to modeling errors resulting from incorrect leak-rate dynamics in the design model, where incorporation of such dynamics would only serve to improve the performance more. A primary shortcoming of the IPSPRT is that it requires significantly more computational resources than

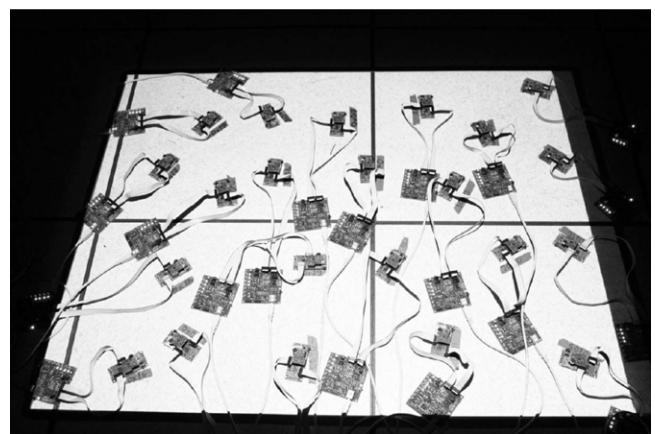


Fig. 6. 22 wireless sensor test bed.

sensor-level approaches. This computational requirement forces the IPSPRT calculations to be performed on a central machine. Additionally, the IPSPRT is prone to systematic errors occurring from preemptive decisions occurring before sufficient data has been received. This result is indicated in Fig. 5 by the fact that when the average time-to-decision is set low, the resulting test incurs significant probability of miss, although the corresponding performance of the naive approach (windowed-average) is shown to suffer much more. These systematic errors are a direct result of the modeling errors inherent in both the IPSPRT and naive approaches. However, it is shown in Fig. 5 that as the expected time to decision increases, the detector performance improves significantly.

As a final note, we reiterate that the IPSPRT evaluation provided in this section is intended to be a proof of concept. To meet the strict EPA guidelines would require (at the very least) an in-depth simulated evaluation using a state-of-the-art environmental simulation platform such as AERMOD (Cimorelli et al., 2005).

6. Wireless sensor network implementation

This section describes a test bed implementation for evaluating the robustness of the IPSPRT in the presence of common errors associated with using a WSN to collect data. The test bed, shown in Fig. 6, consists of 22 Firefly sensor nodes (Mangharam et al., 2007). Each firefly node in Fig. 6 runs the Nano-RK operating system (Eswaran et al., 2005), contains a light intensity sensor, and is connected to a unique programming board that supplies power to the firefly nodes and allows for quick reprogramming of the entire network and background monitoring through a wired network.

Fig. 7 shows a flowchart describing the functionality of the test bed. In addition to the 22 firefly nodes and programming boards, the complete test bed incorporates two computers (named *Ramathorn* and *Coolstore* in Fig. 7), and a light projector. Coolstore is a Linux machine that operates as a network manager, and Ramathorn is a Windows XP machine that executes the IPSPRT routines and simulates the environment, which is projected onto the 22 firefly nodes as different light intensities ranging from 0 to 255. The light sensor on each firefly node outputs a light intensity value ranging from 0 to 1023 corresponding to *bright* and *dark*, respectively. A second-order least squares approximation is used to convert the firefly light intensity value to the environment variable used for detection and localization, calibrated using 15 different light intensities.

At each sampling instant, Ramathorn first updates the environmental data and projects the corresponding light data onto the sensor network. The network manager gathers the light sensor observations using the SAMPL data acquisition protocol Rowe

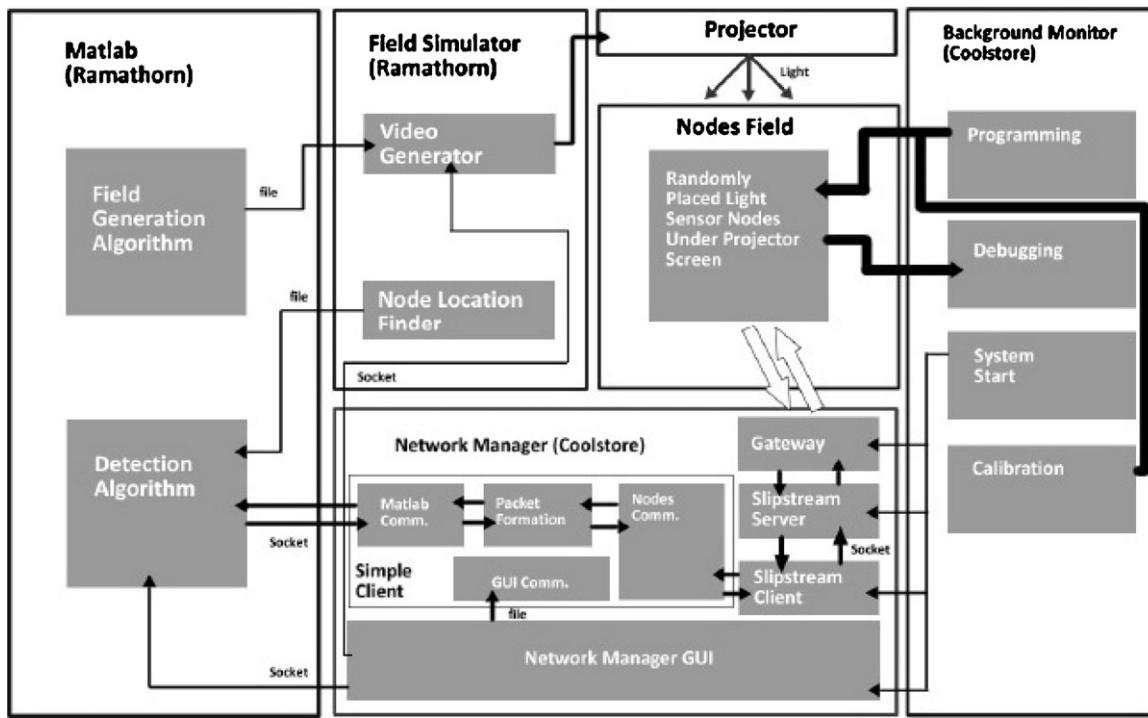


Fig. 7. Test bed architecture.

et al. (2008). After data acquisition is complete, the network manager relays the collected light sensor observations to Ramathorn. Ramathorn then executes IPSPRT routine. When the IPSPRT routine completes for the current time step, the time step is incremented and the procedure repeats.

In the previous section, the IPSPRT is shown to perform much better than the windowed-average test for large-scale CO₂ sequestration site monitoring as the desired probably of false alarm and miss decrease, despite the fact that the model used by the IPSPRT is a crude 2D approximation of the 3D CO₂ transport process. This result suggests that the IPSPRT is robust with respect to moderate modeling errors. When a WSN is used to gather observations, additional robustness issues arise. In this subsection, we further investigate the robustness of the IPSPRT in the presence of common error/failure scenarios arising when data is gathered using a WSN, namely, packet loss and sensor failures. In the following

evaluation, the environmental CO₂ concentrations are generated using the 2D model assumed by the IPSPRT. The 2D model is used in place of the 3D model as an attempt to investigate the effect of potential WSN errors independent of any modeling errors.

To evaluate the robustness of the IPSPRT with respect to WSN errors/failures, we assume a sensor network of 22 sensors and 49 potential leaks, distributed as in Fig. 8, where a square denotes a sensor location and a dot represents a potential leak location. One thousand (1000) simulations were performed, each lasting for 1200 time steps with a randomly located single leak becoming active at time step 600. We assume the wind is always blowing in the direction (1,1) and tested six different wind intensities ranging from 0 km/h to 8.5 km/h.

To evaluate sensor failures, we consider two scenarios. The first sensor failure considered, and henceforth referred to as *sensor failure 1*, assumes that the interior sensors denoted by the filled boxes in Fig. 9 drop out of the network permanently. Similarly, the second sensor death considered, referred to as *sensor failure 2*, assumes that the exterior sensors denoted by the filled boxes in Fig. 10 drop off the network.

To evaluate the robustness of the IPSPRT in the presence of the sensor failures, we consider four different active leak scenarios

1. *Synchronous distributed leaks*: Two leaks located far apart from one another that become active at the same time.
2. *Synchronous clustered leaks*: Two leaks located near one another that become active at the same time.
3. *Asynchronous distributed leaks*: Two leaks located far apart from one another that become active at different times.
4. *Asynchronous clustered leaks*: Two leaks located near one another that become active at different times.

In the following, we assume clustered leaks are located at positions (0.15 km, 0.25 km) and (0.2 km, 0.2 km), while distributed leaks are located at positions (0.15 km, 0.25 km) and (0.30 km, 0.10 km). Each active leak scenario is evaluated in low wind (2.8 km/h) and in high wind (8.5 km/h). The test scenarios assume

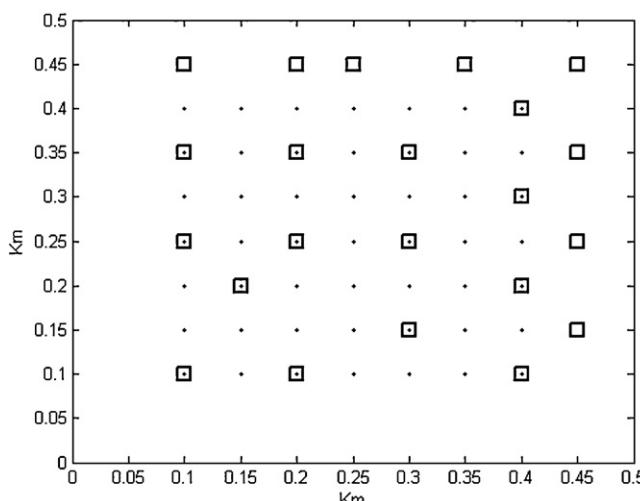


Fig. 8. Sensor and potential leak locations.

Table 6

Average time-to-decision (in sampling periods) for deciding no leaks exist.

	Wind speed	Packet loss	Sensor failure 1	Sensor failure 2
Synchronous distributed leaks	Low	6.79	6.89	7.77
	High	7.71	7.72	11.68
Synchronous clustered leaks	Low	6.81	6.90	7.79
	High	7.72	7.74	11.72
Asynchronous distributed leaks	Low	6.82	6.92	7.80
	High	7.72	7.75	11.74
Asynchronous clustered leaks	Low	6.81	6.87	7.79
	High	7.74	7.80	11.78

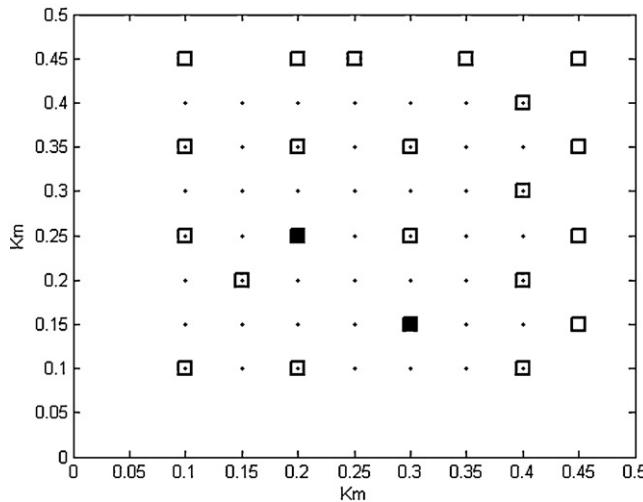


Fig. 9. Sensor failure 1 (interior sensor failure).

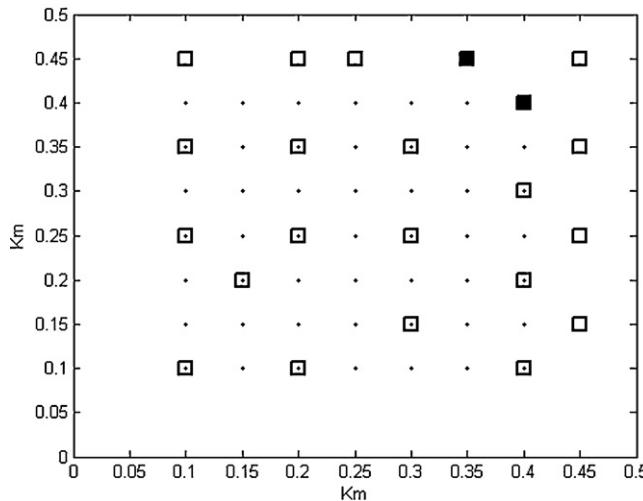


Fig. 10. Sensor failure 2 (exterior sensor failure).

a spatial discretization of 50 m and the temporal sampling rate of 10 min. All leak strengths (regardless of position) are scalar constants and are taken from the Gaussian distribution with an expected value of 200 parts-per-million (PPM) per second with a standard deviation of 100 PPM per second. The results are classified by the IPSPRT's decision to accept or reject the null hypothesis (no leaks exist). It is important to consider the IPSPRT's decision when evaluating performance because the time-to-decision can differ significantly between deciding no leaks exist and some leaks exist.

Table 6 illustrates the average time-to-decision in sampling steps when the IPSPRT decides that no leaks exist. In Table 6 the results indicate that the expected time-to-decision when an exterior sensor fails is about 50% greater than if an interior sensor fails or packet loss occurs. There is no significant difference in the average time-to-decision for each error between the different leak scenarios when deciding no leaks are active.

Table 7 illustrates the average time-to-decision in sampling steps for deciding some leak exists. The results show that an interior sensor failure has the largest effect on the average time-to-decision. Observing from Fig. 9 and recalling that clustered leaks are located at positions (0.15 km, 0.25 km) and (0.2 km, 0.2 km), while distributed leaks are located at positions (0.15 km, 0.25 km) and (0.30 km, 0.10 km), we note that the interior nodes that failed were also the closest downwind nodes from the active leaks. Thus, in high-wind situations, where the concentration of CO₂ is small except for near the leak, having nodes fail in the downwind proximity can cause a significant increase in the time-to-decision.

The results in Table 7 also illustrate that in low-wind situations, leaks are detected quickly, and in high-wind situations leaks are detected slowly. It is intuitive that when sensors observe a significant increase in the CO₂ concentration level, a detection occurs quickly. As the CO₂ concentration decreases (due to either a decrease in the leak magnitude or an increase in the wind speed), the time-to-decision increases. Based on the results for deciding no leaks exists in Table 6 and deciding some leak exists in Table 7, one could institute a preemptive detection scheme based on the duration of the monitoring period. As the time-to-decision increases, it becomes more likely (based on the empirical data) that leaks exist, but can't be accurately detected yet using the IPSPRT.

We observe in Table 7 that when two leaks become active at the same time and are near one another, the time-to-decision is smaller

Table 7

Average time-to-decision (in sampling periods) for deciding a leak exists.

	Wind speed	Packet loss (packet loss)	Sensor failure 1	Sensor failure 2
Synchronous distributed leaks	Low	1.92	3.76	2.61
	High	24.82	31.05	25.60
Synchronous clustered leaks	Low	1.61	2.78	2.01
	High	12.61	20.10	12.79
Asynchronous distributed leaks	Low	2.01	3.91	2.83
	High	24.79	31.11	25.52
Asynchronous clustered leaks	Low	2.02	4.01	2.96
	High	26.61	34.58	26.91

than in any other corresponding leak scenario. This is due to the overall increase in the CO₂ concentration at each of the down-wind sensor nodes. When multiple leaks are active in the same proximity, the results are similar to a single leak with a larger leak rate. These multiple proximate active leak scenarios improve the time to detection. When comparing the time-to-decision for deciding no leaks exist (**Table 6**) vs. some leak occurs (**Table 7**), we observe that an exterior node sensor failure significantly increases the time-to-decision for deciding no leaks exist, while a failure of down-wind nodes close to the leak has a similar effect when claiming a leak exists. These observations lead us to believe that sensor failures (nodes dropping out of the network) are of a key concern when a WSN is used to perform leak detection.

7. Discussion and future work

In this paper, the problem of large-scale multiple-leak detection using a WSN at CO₂ sequestration sites is addressed. Using a simplified 2-D model for a 3-D atmospheric CO₂ transport process, the IPSPRT can be applied for the purposes of identifying leaks. It is shown that even though the 2-D model is known to be only a basic representation of the atmospheric CO₂ transport, the improvement in performance from implementing such a strategy increases as compared to a windowed-average approach as the desired accuracy of detection increases. A test bed implementation is presented and employed to evaluate the robustness of the IPSPRT in the presence of common WSN errors and failures.

Future work on this problem includes improved model parameter identification and ambient CO₂ concentration characterization. The more accurate the ambient CO₂ concentration characterization, the better the detector's performance will be. In this light, further investigation into the effects of model parameter errors and WSN errors on detection performance is also warranted. Additionally, studies into the distribution of leak rates is warranted to determine whether the leaks are Gaussian, bi-Gaussian, etc. Lastly, this study illustrates that the IPSPRT shows promise for CO₂ sequestration site surface monitoring. A next step would be to further evaluate the IPSPRT using state-of-the-art atmospheric simulation environments and to commence preliminary field testing.

Acknowledgement

This work was supported in part by the National Energy Technology Laboratory under RDS contract DE-AC26-04NT41817.

References

- Cimorelli, A.J., Perry, S.G., Venkatram, A., Weil, J.C., Paine, R.J., Wilson, R.B., Lee, R.F., Peters, W.D., Brode, R.W., 2005. Aermod: A dispersion model for industrial source applications. part i: general model formulation and boundary layer characterization. *Journal of Applied Meteorology* 44 (5), 682–693.
- Eswaran, A., Rowe, A., Rajkumar, R., 2005. Nano-RK: an energy-aware resource-centric rtos for sensor networks. In: Real-Time Systems Symposium, 2005. RTSS 2005. 26th IEEE International, pp. 10–265. doi:10.1109/RTSS.2005.30. URL <http://dx.doi.org/10.1109/RTSS.2005.30>.
- Fatehifar, E., Elkamel, A., Taheri, M., 2006. A MATLAB-Based modeling and simulation program for dispersion of multi-pollutants from an industrial stack for the educational use in a course on air pollution control. *Computer Applications in Engineering Education* 14 (4), 300–312.
- Revision to the guideline on air quality models: adoption of a preferred general purpose (flat and complex terrain) dispersion model and other revisions, final rule. *Federal Register* 40.
- Fox, E.B., Fisher, J.W., Willsky, A.S., 2007. Detection and localization of material releases with sparse sensor configurations. *IEEE Transactions on Signal Processing*, 55(May (5)), 1886–1898. doi:10.1109/TSP.2006.890914 [see also *IEEE Transactions on Acoustics, Speech, and Signal Processing*].
- Heinsohn, R., Kabel, R., 1999. Sources and Control of Air Pollution. Prentice Hall, New Jersey.
- Hernandez, E., Martan, F., Valero, F., 1991. State-space modeling for atmospheric pollution. *Journal of Applied Meteorology* 30 (6), 793–811.
- Kailath, T., Fellow, L., Poor, H.V., 1998. Detection of stochastic processes. *IEEE Transactions on Information Theory* 44, 2230–2259.
- Larcher, W., 1995. Physiological Plant Ecology. Springer, New York.
- Mangham, R., Rowe, A., Rajkumar, R., 2007. Firefly: a cross-layer platform for real-time embedded wireless networks. *Real-Time Systems*, 37(3), 183–231. doi:10.1007/s11241-007-9028-z. URL <http://dx.doi.org/10.1007/s11241-007-9028-z>.
- Melman, A., 1997. Geometry and convergence of Euler's and Halley's methods. In: SIAM Review, vol. 39. Society for Industrial and Applied Mathematics, pp. 728–735.
- Rao, K.S., 1983. E. R. L. (U.S.), Plume Concentration Algorithms with Deposition, Sedimentation, and Chemical Transformation [Microform]. U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, Washington, DC.
- Rowe, A., Lakshmanan, K., Rajkumar, R., 2008. Sample: a simple aggregation and message passing layer. In: Real-Time Systems Symposium.
- Saripalli, P., Amonette, J., Rutz, F., Gupta, N., 2006. Design of sensor networks for long term monitoring of geological sequestration. *Energy Conversion and Management* 47 (13–14), 1968–1974.
- Seinfeld, J.H., Pandis, S.N., 1998. Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. John Wiley and Sons Inc, New York.
- Sharan, M., Yadav, A.K., 1998. Simulation of diffusion experiments under light wind, stable conditions by a variable k-theory model. *Atmospheric Environment* 32 (20), 3481–3492.
- Sharan, M., Yadav, A.K., Singh, M.P., Agarwal, P., Nigam, S., 1996. A mathematical model for the dispersion of air pollutants in low wind conditions. *Atmospheric Environment* 30 (8), 1209–1220.
- Smith, W.H., 1981. Air Pollution and Forests: Interaction between Air Contaminants and Forest Ecosystems. Springer-Verlag, New York.
- Srivastava, V., Plarre, K., Bullo, F., 2009. Randomized sensor selection in sequential hypothesis testing. Available from: <http://arxiv.org/abs/0909.1801>.
- Tannehill, J., Anderson, D., Pletcher, R., 1997. Computational Fluid Mechanics and Heat Transfer, Series in Computational and Physical Processes in Mechanics and Thermal Sciences. Taylor & Francis.
- Turner, D., 1994. Workbook of Atmospheric Dispersion Estimates: An Introduction to Dispersion Modeling, 2nd edition. CRC Press.
- United States Secretary of Energy Steven Chu, 2009. Letter on carbon capture and sequestration.
- Wald, A., 1947. Sequential Analysis. John Wiley and Sons Inc., New York.
- Weimer, J., Sinopoli, B., Krogh, B., 2008. A relaxation approach to dynamic sensor selection in large-scale wireless networks. In: Distributed Computing Systems Workshops, 2008. 28th International Conference on ICDCS '08, pp. 501–506. doi:10.1109/ICDCS.Workshops.2008.82.
- Weimer, J., Sinopoli, B., Krogh, B., 2011. Large-scale source localization with a wireless sensor network application. In: 18th International Federation on Automatic Control (IFAC) World Congress.
- Weimer, J.E., 2010. Large-scale multiple-source detection using wireless sensor networks. Ph.D. Thesis.
- Wells, A., Hammack, R., Veloski, G., Diehl, J., Strazisar, B., 2006. Monitoring, mitigation, and verification at sequestration sites: seque technologies and the challenge for geophysical detection. Tech. Rep., NETL, Pittsburgh, PA.
- Willsky, A., 1976. A survey of design methods for failure detection in dynamic systems. *Automatica* 12, 601–611.
- Yang, Y.-M., Small, M., Junker, B., Bromhal, G., Strazisar, B., Wells, A., 2011. Bayesian hierarchical models for soil CO₂ flux and leak detection at geologic sequestration sites. *Environmental Earth Sciences* 64, 787–798.
- Yang, Y.-M., Small, M.J., Ogretim, E.O., Gray, D.D., Bromhal, G.S., Strazisar, B.R., Wells, A.W., 2011. Probabilistic design of a near-surface CO₂ leak detection system. *Environmental Science and Technology* 45 (15), 6380–6387.