# Pollution Source Identification

Arthur Wiedmer

CE 291F

December 17, 2007



## Outline

## Outline







## Why is this problem interesting ?

- 300000 to 400000 contaminated sites

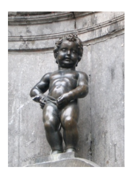

Figure: Water pollution

## Why is this problem interesting ?

- 300000 to 400000 contaminated sites
- Expected clean-up cost : 1 trillion $ (4000$ per US citizen)

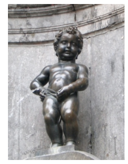

Figure: Water pollution





## Why is this problem interesting ?

- 300000 to 400000 contaminated sites
- Expected clean-up cost : 1 trillion $ (4000$ per US citizen)
- Chemical fingerprinting and records are seldom sufficient to determine the source.

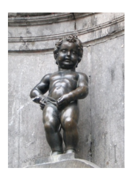

Figure: Water pollution

## Forward problem

Governing equation :

$$\begin{cases} \dfrac{\partial C}{\partial t} = \nabla \left[ D \cdot \nabla C \right] - \nabla \cdot \left[ \boldsymbol{v} \cdot C \right] - \lambda C \\ C(x_1, t) = C_{in}(t) \quad = 0 \leq t \leq T_{final} \\ C(x, 0) = C_0(x) \text{ initial concentration profile} \\ C(x_{sensor}, t) = f(t) \text{ measurements} \end{cases}$$

Some analytical (closed-form) solutions in simple cases. Solved numerically most of the time.





## Inverse problem

- The problem of finding the source and possibly the time history of the pollution is a "time inversion" problem.
- Due to the irreversibility of diffusion it is fundamentally ill-posed. This means that we do not know in general :
  1. if a solution exist
  2. if it is unique
  3. how much it is sensitive to the observed data.

$\Longrightarrow$Since the problem is severly underdetermined we have to make assumptions on the unknown function.

## Outline







## Several methods are used

- Use of closed-form solutions
- Brute-force forward solving and optimization
- Probability and backward solving
- Statistical Estimation.

## A groundwater model

Figure: Henri Darcy

$$\frac{\partial C}{\partial t} = D\mathbf{\Delta} C + q$$

We can consider the equation far from the sources, and a decay rate for the contaminant :

$$\frac{\partial C}{\partial t} = D\mathbf{\Delta} C - \lambda C$$





## A groundwater model

We can add carefully chosen boundary conditions to the model :

$$\lim_{||(\boldsymbol{x},t)||\to\infty} C(\boldsymbol{x},t) = C_0$$

$$\int_{\mathbb{R}^n} [C(\boldsymbol{x},t) - C_0]\, d\boldsymbol{x} = M_0 e^{(t-t_0)}$$

$$\lim_{t\to t_0} [C(\boldsymbol{x},t) - C_0] = 0$$

The solution to the direct problem is then :

$$C(\boldsymbol{x},t) = \frac{M_0 e^{(t-t_0)}}{\sqrt{[2\pi(t-t_0)(2D)]^n}} exp\left(\frac{-||\boldsymbol{x}-\boldsymbol{x_0}||^2}{4D(t-t_0)}\right) + C_0$$

## Identifying the source

We can take the log :

$$-\frac{||\mathbf{x}-\mathbf{x_0}||^2}{4D(t-t_0)} - \lambda(t - t0) - \frac{n}{2} ln(t - t_0) + ln(M_0) - \frac{n}{2} ln(4\pi D)$$

We also introduce additional parameters to simplify the notation : $\kappa = (2D)^{-1}$ and $\alpha = ln(M_0) - \frac{n}{2} ln(4\pi s)$ and consider the accordingly modified measurements $\ln(C(\mathbf{x_l}, t_k) - C_0) = \nu_{k,l}$. We thus have :

$$(k, l)\ \nu_{k,l} = -\frac{||\mathbf{x_l}-\mathbf{x_0}||^2}{(t_k - t_0)} \frac{\kappa}{2} - \lambda(t_k - t0) - \frac{n}{2} ln(t_k - t_0) + \alpha$$





**Theorem**

*For a space of any dimension $\mathbb{R}^n$, four measurements of $C(\mathbf{x_l}, t_k)$ at two different times $t_1 > t_0$ and $t_2 > t_0$ common at each to distinct locations $\mathbf{x_1}$ and $\mathbf{x_2}$ suffice to express $t_0$ in closed form :*

$$t_0 = \frac{(\nu_{k+1,l+1} - \nu_{k+1,l})t_{k+1} - (\nu_{k,l+1} - \nu_{k,l})t_k}{(\nu_{k+1,l+1} - \nu_{k+1,l}) - (\nu_{k,l+1} - \nu_{k,l})}$$

*Two additional measurements at the same locations at a third time lead to closed-form solution for the decay rate $\lambda$ and for auxiliary parameter $\alpha$ (They are solutions of a system function of the $t_k$ with a non-zero determinant)*

With the condition that sensor locations $(x_l)_l$ are in general position

$$\Leftrightarrow \forall m \in \{2, \ldots, n+1\} \text{ no subset of } m \text{ points is entirely included in an affine (hyper)plane of dimension } m-2.$$

**Theorem**

*For $\mathbb{R}^n$, $n+2$ measurements of $C(\mathbf{x_l}, t_k)$ at the same time $t_k > t_0$ at each $n+2$ points $\mathbf{x_l} \neq \mathbf{x_0}$ in general position suffice to express the location of the source $\mathbf{x_0}$ in closed form as well as parameter $M_0$ and $D$.*





By using the previous equations and by substituting the values of $t_0, \alpha, \lambda$ in $(k, l)$ we get

$$\kappa||\mathbf{x_l} - \mathbf{x_0}||^2 = -2(t_k - t_0)\left(\nu_{k,l} + \lambda(t_k - t0) + \frac{n}{2}ln(t_k - t_0) - \alpha\right)$$

By forming the ratio of the following form $(h \neq l)$, we obtain $n+1$ independent equations :

$$\frac{||\mathbf{x_h} - \mathbf{x_0}||^2}{||\mathbf{x_l} - \mathbf{x_0}||^2} = r^2_{k,h,l}$$

Those are the equations of the median plane if $r^2_{k,h,l} = 1$ or of a sphere which center is the centroid of $(\mathbf{x_h}, 1)$ and $(\mathbf{x_l}, -r^2_{k,h,l})$. $\mathbf{x_0}$ is the intersection of all these.

## Probabilistic model

We define the probability density function $F_q$ as $\frac{C(x,t)}{M_0}$ For a divergence-free flow and non-reactive transport :

$$\frac{\partial F_q}{\partial t} = -q \cdot \nabla F_q + \nabla [D \cdot \nabla F_q] \quad (1)$$

$$F_q(x_0, 0) = \delta(x - x_0) \quad (2)$$

The backward model :

$$\frac{\partial F_{-q}}{\partial t} = q \cdot \nabla F_{-q} + \nabla [D \cdot \nabla F_{-q}] \quad (3)$$

$$F_{-q}(x_1, 0) = \delta(x - x_1) \quad (4)$$

The boundary conditions are

$$\nabla F_{\pm q} = 0 \text{ on } \partial\Omega$$





## Concept

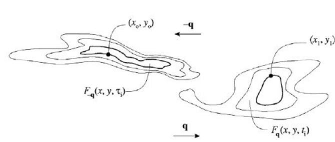


Figure: Backward model concept

$$F_q(x_1, t_i) = F_{-q}(x_0, \tau_i)$$

where $\tau_i = t_{detection} - t_i$

$$C_0(x_1, \tau_i) = \frac{1}{M_0} \cdot C(x_1, t_i)^2 dx_1$$

## Solving

We can run the backward flow model and watch for the isocontour obtained by equation :

$$C_0(x_1, \tau_i) = \frac{1}{M_0} \cdot C(x_1, t_i)^2 dx_1$$

- Running this from a single point does not yield to a unique solution. But it can be done from several points, since the equations are true $\forall x_1$.
- Since the equation is linear we can use as initial condition for the backward model the whole concentration profile at $t_detection$. We can then backtrack it until it shrinks discovering the source.





## Outline

## What to do next ?

For the project :

- Numerical example : Have the simulation working

After the project : Maybe investigating on other methods, there is still a lot to do.





Questions ?