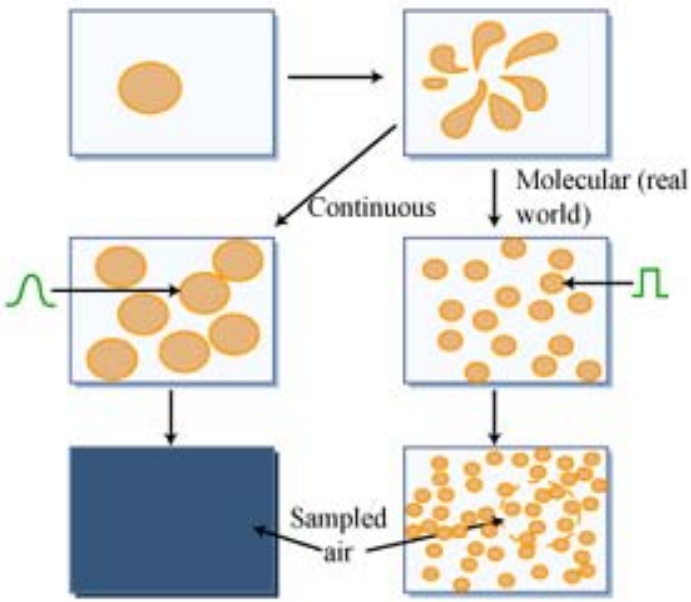
Estimating Surface/Internal Sources and Sinks



Cartoon illustrating the fate of an injection of a pure exotic chemical into a real (molecular) fluid and idealized (continuous) fluid. Graphs of the mole fraction variation across the miniscule fluid elements in the two cases are shown in the two middle panels.

Figure by MIT OCW.

Notation	
Bold Symbols	**Vector or Matrix**
$\mathbf{T}$	Transpose
$\mathbf{I}$	Identity Matrix
$\mathbf{P}(\ )$	Probability of ( )
$\mathbf{E}(\ )$	Expectation value of ( )
$\mathbf{Y}^o$	Measurement Vector
$\boldsymbol{\tau}$	Error in Y^o
$\mathbf{Y}$	Estimate of Y^o
$\mathbf{R}$	Error Covariance matrix of Y^o
$\mathbf{x}$	Estimate of State Vector
$\boldsymbol{\upsilon}$	Error in x
$\mathbf{P}$	Error Covariance matrix of x
$\mathbf{x}^t$	True State Vector
$\mathbf{H}$	Model Partial Derivative (Observation) Matrix
$\mathbf{H}^t$	True Value of H
$\tilde{\mathbf{H}}$	Error in H
$\mathbf{K}$	Kalman Gain Matrix
$\mathbf{M}$	State Extrapolation (Transition) Matrix
$\boldsymbol{\eta}$	System Random Forcing Vector
$\mathbf{Q}$	System Random Forcing Covariance Matrix
$(\ )^a$	Estimate After Use of Measurement
$(\ )^f$	Estimate Before Use of Measurement

Measurement Equation

In Lagrangian framework:

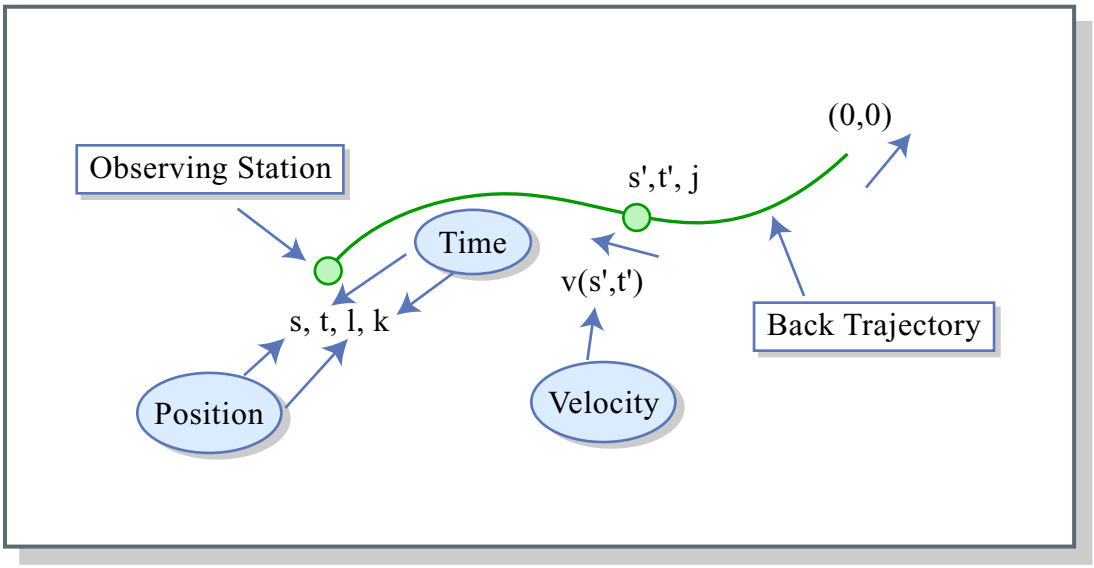


Figure by MIT OCW.

Change in mole fraction $y(s,t)$ from its initial condition $y^*(0,0)$ is given by

$$y(s,t) = y^*(s,t) - y^*(0,0)$$

$$= \int_0^t x\,dt'$$

$$= \int_0^s \frac{x}{v}\,ds' \qquad \left(v = \frac{ds'}{dt'}\right)$$

where x is the net chemical production (ignoring molecular diffusion and assuming perfect definition of back trajectory)

Real measurements of y have errors ε

$$y^0(s,t) = \int_0^s \frac{x^t}{v}\,ds' + \varepsilon(s,t)$$

Real measure-ment	True value if perfect model	Measure-ment error

In discrete form:

$$y_i^0 = \sum_j h_{ij} x_j^t + \varepsilon_i$$

In vector-matrix form for multiple observing sites i

$y^0 = Hx^t + \varepsilon$ measurement equation

where $H = \left[h_{ij}\right]$ is the partial derivative (or observation) matrix, and

$$h_{ij} = \frac{\partial y_i^0}{\partial x_j^t} = \frac{\partial y_i}{\partial x_j} \Bigg\}$$ compute in model

where (y_i, x_j) denotes an estimate of (y_i^0, x_j^t).

Eulerian models: Here x refers to grid points in the Eulerian model rather than to points along a back trajectory. To utilize the above equations we must define:

$$x = x(\text{grid point}) - x(\text{ref}, \text{grid point})$$

$$y = y(\text{grid point}) - y(\text{ref}, \text{grid point})$$

where y(ref) is the mole fraction computed in a reference run using best available estimates $x(\text{ref})$ of the state vector prior to estimating x^t.

The measurement equation expresses an apparent linear relation between the observation vector y^0 and the unknowns contained in the state vector x^t. It is an expression of the forward problem. The theory of the Linear Kalman Filter allows us to perform the inverse problem in which we optimally estimate x^t given a model and a time series of observations y^0 at one or more sites.

Optimal estimation – the "cost" function (J)

Seek the estimate x of x^t that minimizes J (i.e. $\frac{\partial J}{\partial x} = 0$)

(Notation: $p(\ ) =$ probability distribution function of $(\ )$)

(a) No knowledge of $p(x^t)$, $p(y^0)$

J = sum of squares of differences between predictions $(y = Hx)$ and observations (y^0)

$= (y^0 - Hx)^T (y^0 - Hx)$ (called "least squares" minimization)

(b) Know $p(y^0)$ and hence R = expectation$[\varepsilon\varepsilon^T]$

Then seek estimate x of x^t that maximizes the conditional probability $p(y^0 | x^t)$:

$J = (y^0 - Hx)^T R^{-1} (y^0 - Hx)$ (R^{-1} weights each squared difference by inverse of variance of each relevant observation)

(called "weighted least squares" minimization or "maximum likelihood")

(c) Know $p(y^0)$ and $p(x^t)$

Then seek x that maximizes $p(x^t | y^0)$ which is defined from Bayes Theorem:

$$p(x^t \mid y^o) = \frac{p(y^o \mid x^t)p(x^t)}{p(y^o)}$$

(called the "minimum variance Bayes estimate" – see below for the definition of J in this case for the Kalman filter)

KALMAN FILTER

Suppose we have an optimal estimate of the state vector available prior to consideration of the kth measurement y_k^o in a data series and we wish to obtain a new optimal estimate x_k^a and its error ν_k^a using this measurement. We can thus propose in general

$$x_k^a = K_k' x_k^f + K_k y_k^o \tag{13}$$

$$\nu_k^a = K_k' \nu_k^f + K_k \varepsilon_k \tag{14}$$

and seek to specify the matrices K_k and K_k'. Using the measurement equation (5)

$$y^o = Hx^t + \varepsilon$$

to define $y_k^f = H_k x_k^f$, the definition $x_k = x_k^t + \nu_k$, the demand that the random measurement errors have zero mean $(E[\varepsilon_k] = 0)$, and finally, the demand that estimations are unbiased $(E[\nu_k^a] = 0)$, we can show that $K_k' = I - K_k H_k$. Hence the new estimate is

$$x_k^a = x_k^f + K_k[y_k^o - H_k x_k^f] \tag{15}$$

with an error

$$\nu_k^a = (I - K_k H_k)\nu_k^f + K_k \varepsilon_k \tag{16}$$

and estimation error covariance matrix

$$P_k^a = E[\nu_k^a (\nu_k^a)^T] \tag{17}$$

Substituting (16) into (17), using the definition of the measurement error covariance matrix $R_k = E[\varepsilon_k \varepsilon_k^T]$, and demanding that measurement errors and state errors are uncorrelated (so that $E[\nu_k^f \varepsilon_k^T] = E[\varepsilon_k (\nu_k^f)^T] = 0$ we obtain

$$P_k^a = (I - K_k H_k) P_k^f (I - K_k H_k)^T + K_k R_k K_k^T \tag{18}$$

We now use the criterion of optimality to determine K_k. Since we will assume we know $p(y^o)$ and $p(x^t)$, we will choose a value for K_k which minimizes the cost function J (equation 12) for the minimum variance Bayes estimate. Specifically

$$\begin{aligned} J_k &= E[(\nu_k^a)^T \nu_k^a] \\ &= \text{trace}[P_k^a] \end{aligned} \tag{19}$$

Evaluating $\partial J_k / \partial K_k = 0$ and solving for the so-called "Kalman Gain" matrix K_k we have

$$K_k = P_k^f H_k^T [H_k P_k^f H_k^T + R_k]^{-1} \tag{20}$$

Substituting (20) into (18) then yields

$$P_k^a = [I - K_k H_k] P_k^f \tag{21}$$

Finally, using the state space equation (7)

$$x(t) = M(t, t_o) x(t_o) + \eta(t, t_o)$$

we then obtain the estimates of x_k^f needed in (15) and P_k^f needed in (21)

$$x_k^f = M_{k-1} x_{k-1}^a \tag{22}$$

$$P_k^f = M_{k-1} P_{k-1}^a M_{k-1}^T + Q_{k-1} \tag{23}$$

where $Q_{k-1} = E[\eta_{k-1} \eta_{k-1}^T]$, and x_{k-1}^a and P_{k-1}^a are the optimal outputs from the previous iteration of the filter. From our earlier discussion (Section 3), Q could represent random forcing in the system model due to transport model errors.

To use the filter we must provide initial (a priori) estimates for x and P. Then from any prior output estimates (x_{k-1}^a, P_{k-1}^a), we use measurement k information (y_k^o, R_k) and model information (H_k, Q_k) together with equations (22), (23), (20), (15), and (21) to provide outputs x_k^a and P_k^a for inputs to the next step. The filter equations are summarized in Table 1.

Some intuitive concepts regarding the DKF are useful in understanding its operation. First, from equation (20), the gain matrix $K_k \to H_k^{-1}$ (its "maximum" value) as the measurement error covariance (noise) matrix $R_k \to 0$, and $K_k \to P_k^f H_k^T R_k^{-1}$ (its "minimum" value) as $R_k \to \infty$. Since the update in the state vector $x_k^a - x_k^f$ varies linearly with K_k, it is clear that measurements noisy enough so that R_k much exceeds $H_k P_k^f H_k^T$, will contribute much less to improvement of the state vector estimation.

In this respect we can usefully consider $H_k P_k^f H_k^T$ as the error covariance matrix for the measurement estimates y_k. This emphasizes the importance of the weighting of the data inherent in R_k and the distortions created if erroneous R_k are used. Note that R_k can include model error, mismatch error, and instrumental error as noted earlier.

Second, using (21), and recognizing that the maximum value of $K_k H_k = I$, we see $P_k^a \le P_k^f$ with equality occurring for infinitely noisy measurements. Hence, the error covariance matrix P_k (whose diagonal elements are the variances of the state vector element estimates) decreases by amounts sensitively dependent on the measurement errors.

Third, we note from (23), that random forcings η in the system (state-space) model [equation (7)], which are represented here by Q, will increase the extrapolated error covariance matrix P_k^f by amounts depending on the relative values of Q_{k-1} and the extrapolation matrix $M_{k-1}P_{k-1}^a M_{k-1}^T$ in the absence of system (state-space) model noise. The inclusion of Q lessens the influence (or memory) of previous iterations in the filter operation. In the extreme, sufficiently large values of Q will prevent the capability of even non-noisy measurements to decrease P_k and hence increase the confidence in the state vector estimate. In other words excellent (non-noisy) measurements are of little use if the system (state-space) model is very noisy (e.g., through random variations η introduced by random transport errors).

Table 1: Kalman Filter Equations*

Definition	**Equation**
Measurement equation (model)	$y_k^o = H_k x_k^t + \varepsilon_k; \quad y_k = H_k x_k^f$
System (state) equation (model)	$x_k = M_{k-1} x_{k-1} + \eta_{k-1}$
State update	$x_k^a - x_k^f = K_k (y_k^o - y_k)$
Error Update	$P_k^a = (1 - K_k H_k) P_k^f$
Kalman gain update	$K_k = P_k^f H_k^T (H_k P_k^f H_k^T + R_k)^{-1}$
State time extrapolation	$x_k^f = M_{k-1} x_{k-1}^a$
Error time extrapolation	$P_k^f = M_{k-1} P_{k-1}^a M_{k-1}^T + Q_{k-1}$
System random forcing covariance	$Q_k = E(\eta_k \eta_k^T)$
Measurement error covariance	$R_k = E(\varepsilon_k \varepsilon_k^T)$
Estimation error covariance	$P_k = E(\nu_k \nu_k^T)$
Input measurement matrix	$= H_k = \partial y_k / \partial x_k$
Input system random forcing covariance	$= Q_k$
Input state extrapolation	$= M_k$
Input measurement	y_k^o
Input measurement error covariance	$= R_k$
Filter iteration	$--- \rightarrow (k-1)^f, \rightarrow$ estimate $\rightarrow (k-1)^a, \rightarrow$ extrapolate $\rightarrow (k)^f, \rightarrow ---$

*A superscript a or superscript f denotes respectively the value before (f) or after (a) an update of an estimate using measurements, and k denotes the measurement number. In general, errors are assumed random with zero mean and measurement and estimation errors are uncorrelated.