SAND2007-7702C

A Comparison of Markov Models and the Ensemble Kalman Filter for Estimation of Sorption Rates

AGU Fall Meeting,

San Francisco, CA

December 11, 2007

Eric Vugrin, Sean McKenna, and Kay Vugrin
Sandia National Laboratories

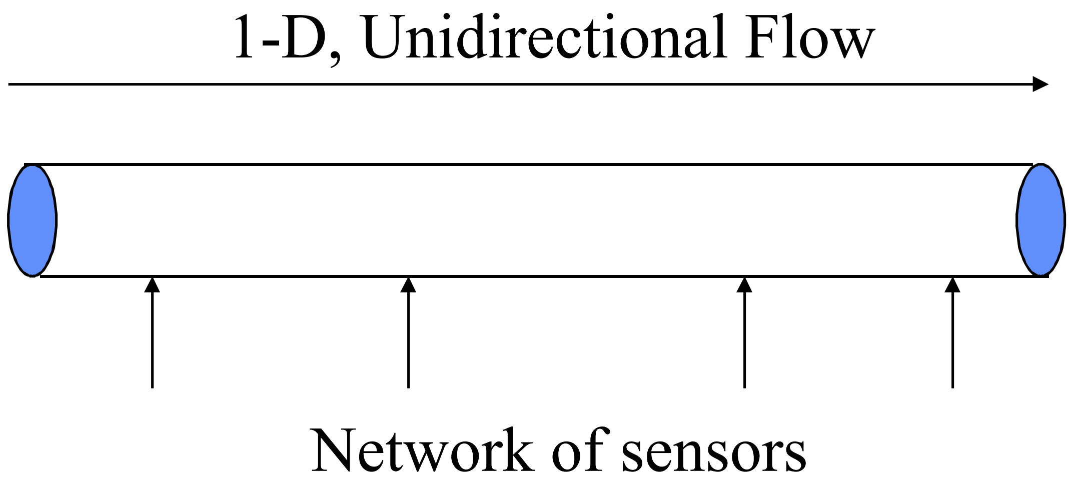
Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.

Outline

- Set Up
- Particle Tracking Model
- Prediction of de/sorption rates
- Application of the EnKF
- Results

Sandia National Laboratories

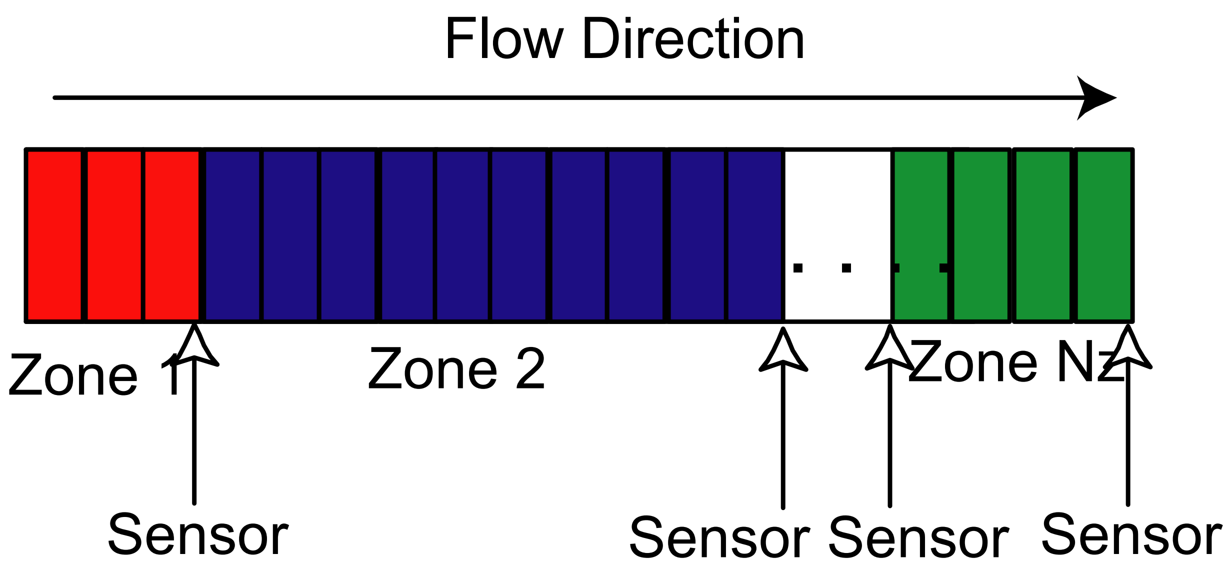
Physical Setup



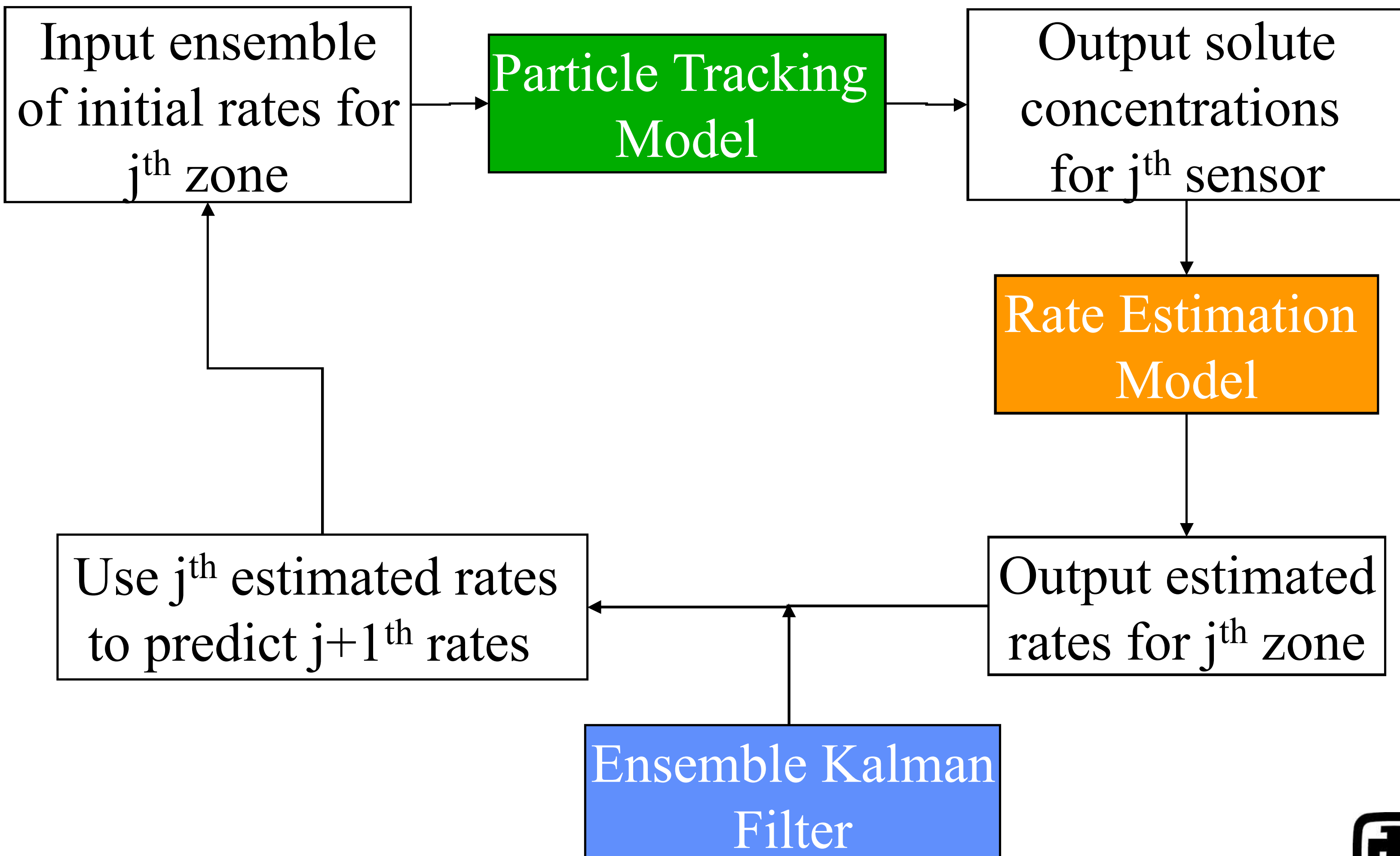
- Setup
  - Streamline with spatially varying de/sorption rates
  - Network of sensors to measure concentration
- Goal: using measured concentration data from sensors for an injected tracer, identify the de/sorption rates in real time

Modeling Set Up

- The streamline is divided into N_z zones
- Each zone is divided into subintervals of uniform length $\Delta x = v \Delta t$
- There is a sensor at the end of each zone
- De/sorption rates are constant in each zone and with respect to time
- At t=0, all particles are injected into the inlet of Zone 1



Rate Estimation Process



Particle Tracking Model: Probabilities

- Mishra et al. developed a Markov model that used constant de/sorption rates to define probabilities that particles transition between sorbed and aqueous phases
- Some key probabilities:
  - Particle in aqueous phase stays in aqueous phase for 1 time step:

$$\hat{p} = 1 - k_f \Delta t$$

  - Sorbed particle transfers to aqueous phase for 1 time step:

$$\tilde{p} = k_r \Delta t$$

- These probabilities form foundation of the particle tracking model

Particle Tracking Model: De/sorption Process

- Phase (aqueous or sorbed) determined at the beginning of each time step
- Particle in solution travels a distance of $\Delta x = v \, \Delta t$ in one time step
- Assume a well-mixed solution so that dispersivity is not an issue
- Implementation in code is a Monte Carlo simulation:
  - A “large” number of particles are tracked
  - Randomly sampled numbers are compared with probabilities to determine phase (and location) of each particle

Rate Estimation Model

- Normalized concentrations are "inverted" to estimate rates
- Example: Assume Zone 1 has 3 subintervals
  - Probability that a particle reaches 1st sensor in 3 time steps is

$$\left[\hat{p}(1)\right]^3 = \left[1 - k_f(1)\Delta t\right]^3$$

  - Particle must always remain in aqueous phase
  - Therefore,

$$C_1(3\Delta t) \approx \left[\hat{p}(1)\right]^3 = \left[1 - k_f(1)\Delta t\right]^3$$

$$\Rightarrow k_f(1) \approx \Delta t^{-1}\left(1 - \sqrt[3]{C_1(4\Delta t)}\right)$$

Sandia National Laboratories

Ensemble Kalman Filter (EnKF) Application

- Kalman Filter (KF) is predictor-corrector technique
  - "State" is predicted by linear model, and KF uses observation data and system statistics to "correct" state estimation
- EnKF is popular technique for non-linear applications
- For our application,
  - State = rates + concentrations
  - Predictor= Particle tracking model + rate prediction model
  - Observations = concentrations calculated with particle tracking code and true rates

Sandia National Laboratories

Markov Model+EnKF

- Create ensemble of rates by perturbing "true rates"
- For each ensemble member
  1. Run particle tracking model based upon previously defined probabilities to get concentrations at j^{th} sensor
  2. Use rate prediction model to estimate j^{th} rates
  3. Kalman Filter Update
  4. Re-run particle tracking code with new estimates of j^{th} rates to get concentrations at $(j+1)^{th}$ zone
  5. Repeat steps 2-4 until all rates have been estimated

Final rate estimates for each zone are calculated by averaging rates over all ensemble members

Test Case: True Rates

Zone	Boundary (m)	True k_f (1/day)	True k_r (1/day)	True k_d
1	0.0-10	0.005	0.005	1
2	10-15	0.005	0.000005	10^3
3	15-30	0.001	0.009	0.11
4	30-45	0.009	0.001	9.0
5	45-50	0.001	0.0001	10

Sandia National Laboratories

Creating the Ensemble

- Ensemble rate = true rate x (1+p) where p is randomly sample from U[min,max]
- Two ensembles created
  - U[min,max] = U[-0.4,0.4]
    - Average perturbation is zero
  - U[min,max] = U[0,0.8]
    - Average perturbation is not 0 so there is a “bias”

Results

- Without bias,
  - With large enough ensembles, both approaches estimated sorption rates fairly well (< 5% error)
  - Desorption rates were not as accurately estimated as well since they generally had less impact on concentrations and they rely on accurate estimation of sorption rates
  - LSEs were comparable for both methods
  - Addition of EnKF does not result in better rate estimates that justify additional computational cost

Sandia National Laboratories

Results

- With bias,
  - Markov estimates of sorption rates resulted in REs that converged to mean perturbation
  - Markov estimates of desorption rates resulted in much scatter
  - Addition of EnKF resulted in rejection of “bias” and more accurate estimates of forward rates
  - Addition of EnKF decreased LSE by two orders of magnitude

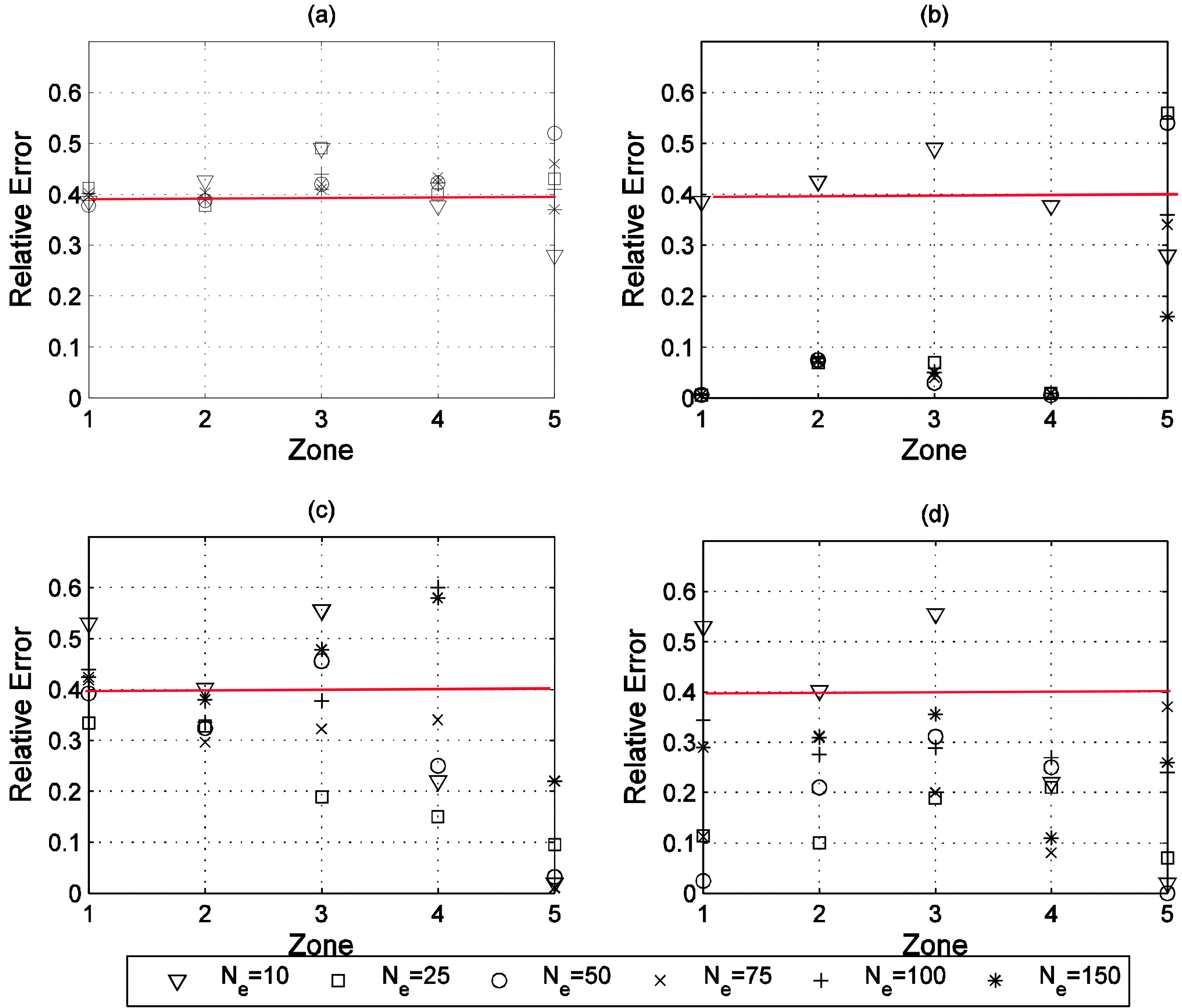
Sandia National Laboratories

Results with Bias

W/O EnKF | W/EnKF

Sorption

De-sorption



Sandia National Laboratories

Results: LSE's with Bias

Ensemble Size	Markov	Markov+EKF
10	5.12e-2	2.96e-4
25	5.20e-2	2.06e-4
50	4.79e-2	1.57e-4
75	5.11e-2	2.05e-4
100	4.97e-2	2.71e-4
150	5.07e-2	2.39e-4

Sandia National Laboratories

Questions?