Adjoint approach to parameter identification with application to the Richards Equation

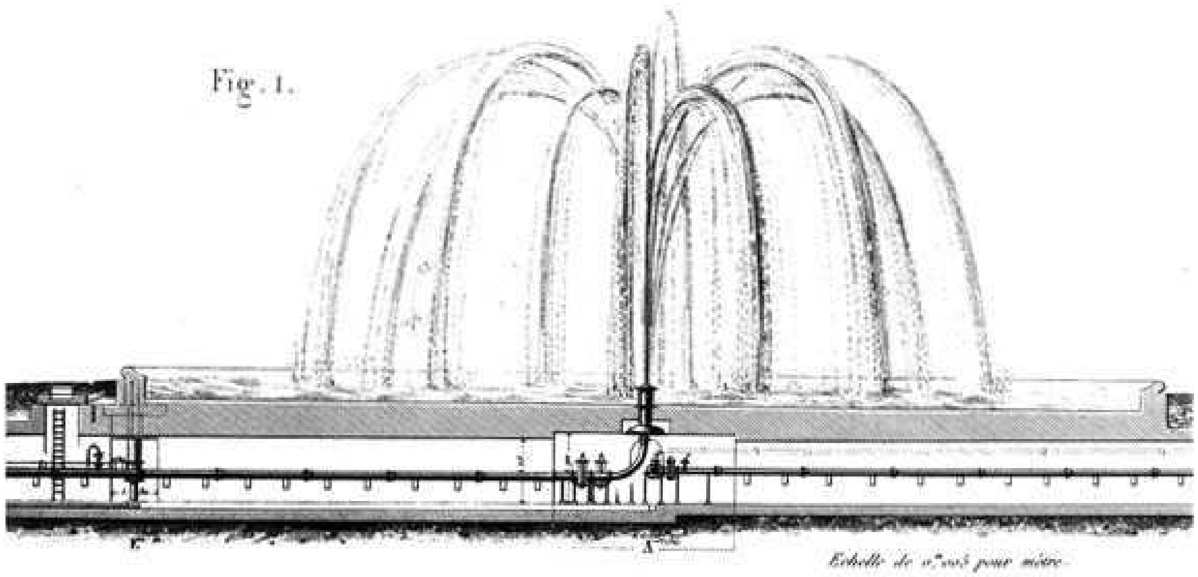
Roger Thelwell

June 2, 2004

Outline

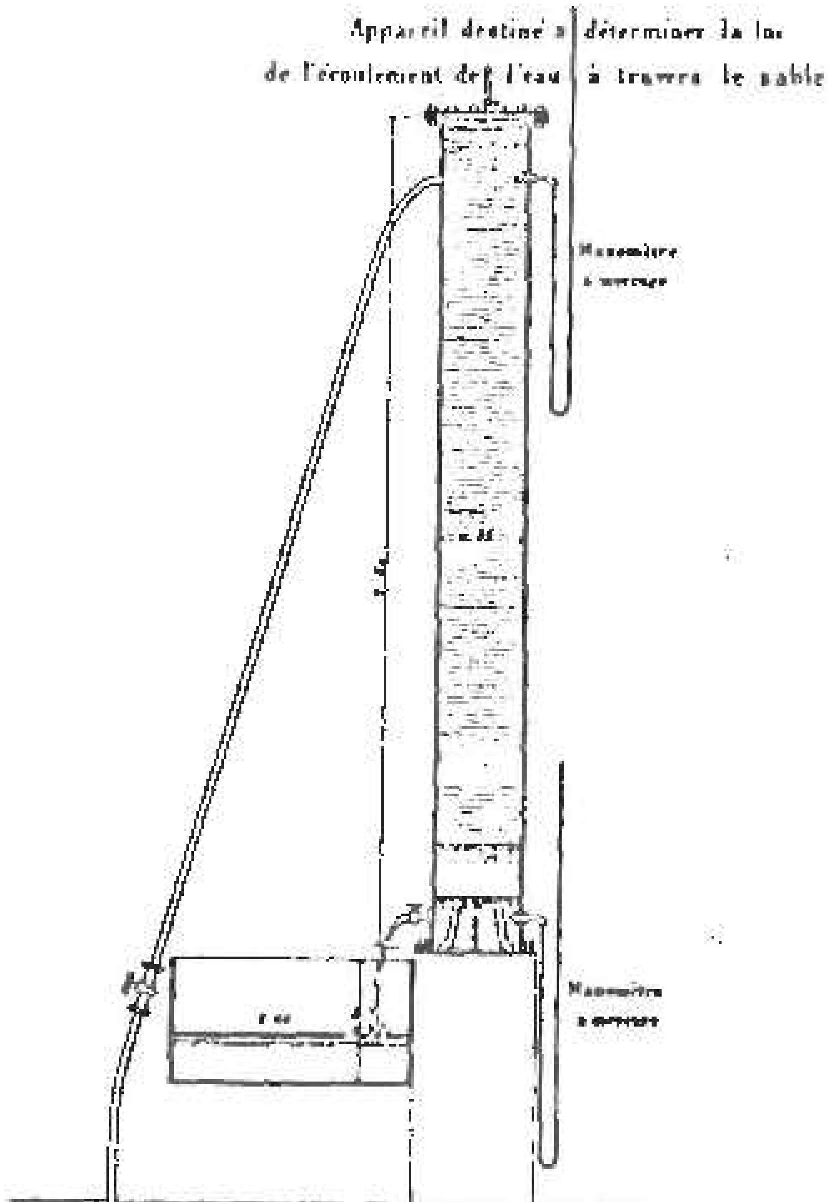
- Introduction
- Two parameter Theory
- Two parameter Numerics
- Conclusions

Introduction



The story begins in 1835. M. Henri Darcy was commissioned to enlarge and modernize the town water works of Dijon, France. He utilized large sand filters. In an effort to determine the correct capacity of these filters, he designed an experiment to provide the required information. In doing so, he gave birth to ground-water hydrology.

Figure 1: Darcy's Experiment



Darcy's superb experimental technique and design allowed him to realize his famous Law:

$$q = -\kappa \frac{(h_1 - h_0)}{z_1 - z_0},$$

the mathematical statement that the flux in a soil column is proportional to the difference in head pressures at two points in the column divided by vertical difference between this points.

A conservation of mass statement for water content $\Theta(h)$ is given by

$$\partial_t \Theta + \partial_x q = 0.$$

But Darcy's law describes q, and so we have ...

Richards Equation - here describing a vertical flow situation.

$$\partial_t(\Theta(h)) - \partial_z(\kappa(\partial_z h - 1)) = 0.$$

Assuming the soil to be homogeneous, unsaturated, and neglecting hysteresis, we write

$$\begin{aligned} \partial_t \Theta(h) - \partial_z (K(h)(\partial_z h - 1)) &= 0, \\ C(h)\partial_t h - \partial_z (K(h)(\partial_z h - 1)) &= 0 \end{aligned}$$

where $C(h) = d\Theta/dh$ represents soil capacity and $K(h)$ hydraulic conductivity. This is called the 1d capacity conductivity form of the Richards equation.

The Richards equation is perhaps the most widely applied model in porous media flow - in areas including ground-water flow and contaminant transport.

In general, a complete model requires

- state inputs - initial and/or boundary data
- structure inputs - coefficients, source terms related to physical properties of the system.

We call these ingredients.

In this talk, we focus on the determination of the unknown ingredients C and K of the Richards Equation. We do so using a novel adjoint equation approach, utilizing over-determined boundary measurements in the recovery.

Integral expressions are presented which relate changes in the unknown coefficients to corresponding changes in the measured output.

Other methods have been developed to reconstruct parameter information from over-determined boundary methods. The most well known is the method of Output Least Squares (OLS), while another is the Equation Error Method (EEM). However, general information about an input to output mapping is not readily available from these methods.

In contrast, the adjoint approach offers a nearly transparent view of the recovery process.

What does a typical solution to the Richards Equation look like?

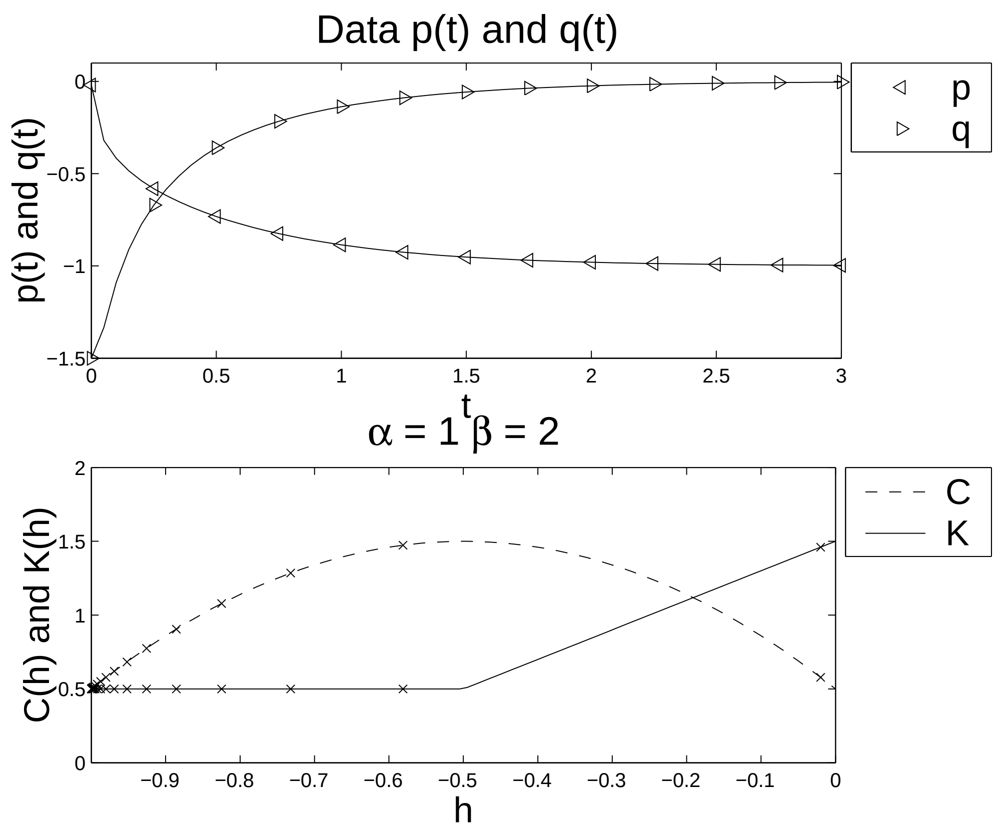
We solve

$$\begin{aligned} C(h)\partial_t h(z,t) - \partial_z(K(h)(\partial_z h(z,t) - 1)) &= 0, \\ h(z,0) &= 0, \\ \partial_z h(0,t) - 1 = 0 \quad h(L,t) &= g(t) \end{aligned}$$

Let's visualize one ...

We would like to use boundary measurements in order to gain information about the coefficients. These measurement look like

Figure 2: Data and coefficients with $\alpha = 1$ and $\beta = 2$



We use an approach based on adjoint versions of the direct problems to derive equations explicitly relating changes in inputs (coefficients) to changes in outputs (measured data).

We will often be considering terms that take the form

$$a(h(z,t))\partial_z h(z,t)$$

If we let

$$A(h(z,t)) = \int_0^{h(z,t)} a(s)$$

then

$$a(h)\partial_z h = \partial_z(A(h(z,t)) \text{ and } a(h)\partial_t h = \partial_t A(h(z,t))$$

We begin the discussion by introducing a mapping Φ. This is the solution map for the PDE. Under certain restrictions on the coefficients, we are able to write

$$\begin{aligned} \Phi : W(J) &\longrightarrow L^2[0,T] \\ \Phi[C,K] &= h(z,t) \end{aligned}$$

with $W(J)$ representing the class of admissible coefficients.

We define the the projection map Γ, which takes h and generates:

- $p = h(0,t)$, a state quantity
- $q = K(h(L,t))(\partial_z h(L,t) - 1)$, the flux

Now we consider the pair of IVBPs

$$\begin{aligned} \partial_t a_i(h_i) = \partial_z(K_i(h_i)(\partial_z h_i - 1)) \\ h_i(z, 0) = 0 \\ \partial_z h_i(0, t) - 1 = 0 \quad h_i(1, t) = 0 \end{aligned}$$

for $i = 1$ and 2.

We have two pairs of coefficients, (C_1, K_1) and (C_2, K_2). One pair we consider to represent the physical system, the other an approximation to the physical system.

We also can generate two pairs of observable quantities (p_1, q_1) and (p_2, q_2)

Now subtract, and multiply the result by a smooth function ϕ and integrate (almost all) terms by parts. In doing so, we notice that by choosing ϕ to satisfy the associated adjoint problem

$$
\begin{aligned}
C_1^* \partial_t \phi(z,t) + K_1^* \partial_{zz} \phi(z,t) + K_1^{'*}(z,t) \partial_z \phi(z,t) &= 0 \\
h_i(z,0) &= 0 \\
K_1^*(0,t) \partial_z \phi(0,t) = P^*(t), \quad \phi(L,t) &= Q^*(t)
\end{aligned}
$$

We control P^* and Q^*, and denote the solution via the map Φ^*, where $\Phi^*[P^*, Q^*] = \phi$

To ease notation, let

$$\begin{aligned}
\Delta K(h) &= K_1(h) - K_2(h), \\
\Delta C(h) &= C_1(h) - C_2(h), \\
\Delta q(t) &= q_1(t) - q_2(t) \quad \text{and} \\
\Delta p(t) &= p_1(t) - p_2(t).
\end{aligned}$$

And a miracle occurs ...

The problem reduces to the integral identity - relating changes in input to changes in output.

$$
\begin{aligned}
\int_0^T \int_0^L & \quad \Delta C(h_2)\partial_t h_2 \phi(z,t) dz\ dt \\
+ \int_0^T \int_0^L & \quad \Delta K(h_2)(\partial_z h_2 - 1)\partial_z \phi(z,t) dz\ dt \\
= & \quad \int_0^T \Delta q Q^*(t)\ dt + \int_0^T \Delta p P^*(t) dt,
\end{aligned}
$$

This the the general integral identity.

If $\phi = \Phi^*[P^*, 0]$ then the general identity becomes

$$\int_0^T P^*(t)[p(t) - p_2(t)]\, dt = \int_0^T \int_0^L \Delta K(h_2))\partial_z \phi(\partial_z h_2 - 1)\ dx\ dt + \int_0^T \int_0^L \Delta C(h_2)\partial_t h_2 \phi\ dt\ dx,$$

which we refer to as the p-integral identity, and

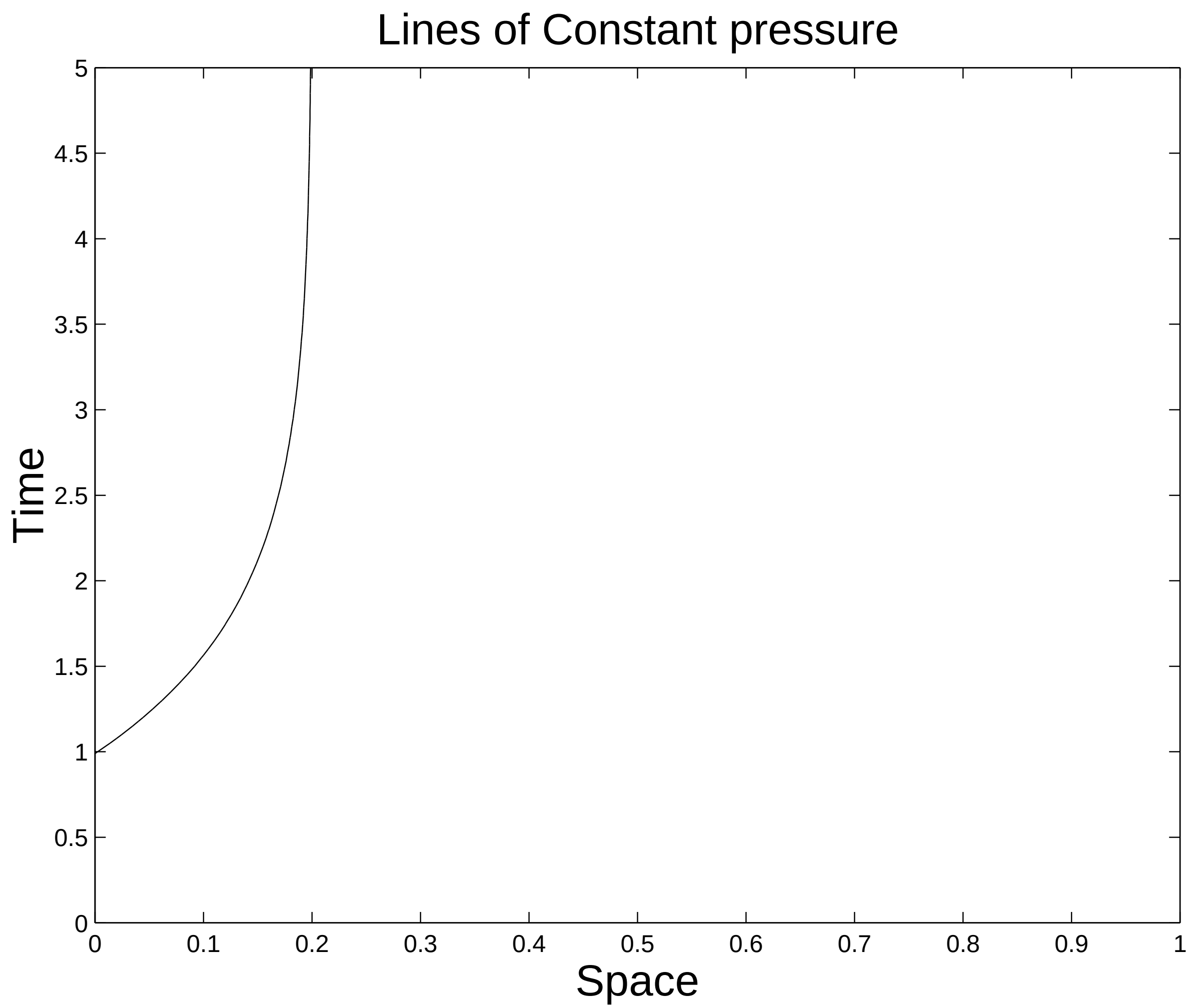
If $\psi = \Phi^*[0, Q^*]$, the identity becomes

$$\int_0^T Q^*(t)[q(t) - q_2(t)]\, dt = \int_0^T \int_0^L \Delta K(h_2))\partial_z \psi(\partial_z h_2 - 1)\ dx\ dt + \int_0^T \int_0^L \Delta C(h_2)\partial_t h_2 \psi\ dt\ dx$$

which we call the q-integral identity.

Since the pressure is monotone in time, we consider an approximation of the coefficients in a piecewise linear basis $\{\Lambda(h)\}$, which naturally follows from a time discretization.

Here we have forced a nodal break at $h = -0.8$ or, alternatively, at $t = 1$.



We now have a linear system, relating changes in the coefficients (ΔC and ΔK) to changes in the boundary observations (Δp and Δq). The explicit relation is:

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \Delta K \\ \Delta C \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (1)$$

Where

$$
\begin{aligned}
M_{11} &= \int_0^T \int_0^L \Lambda(\partial_z h_2 - 1)\partial_z \phi \; dt \; dx \\
M_{21} &= \int_0^T \int_0^L \Lambda(\partial_z h_2 - 1)\partial_z \psi \; dt \; dx \\
M_{12} &= \int_0^T \int_0^L \Lambda(\partial_t h_2)\phi \; dt \; dx \\
M_{22} &= \int_0^T \int_0^L \Lambda(\partial_t h_2)\psi \; dt \; dx \\
b_1 &= \int_0^T \Delta p P^*(t) dt \\
b_2 &= \int_0^T \Delta q Q^*(t) dt
\end{aligned}
$$

We now have a foundation for a numerical scheme!

- choose nodes for coefficient basis - either time, state or some combination.
- Build approximation to coefficients C and K.
- Solve linear system to compute updates to C and K , and apply them.

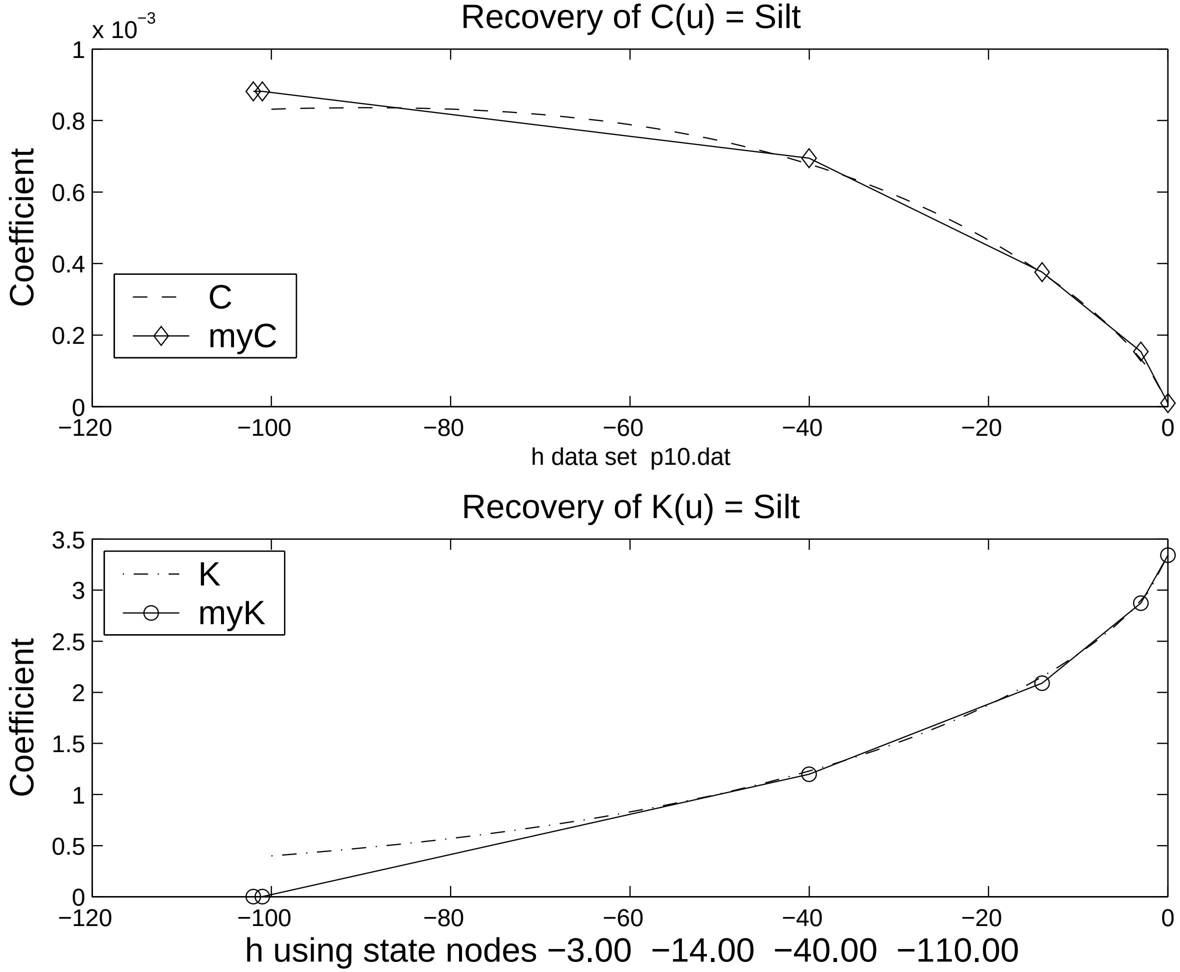
The work is in the third step.

The Bad News

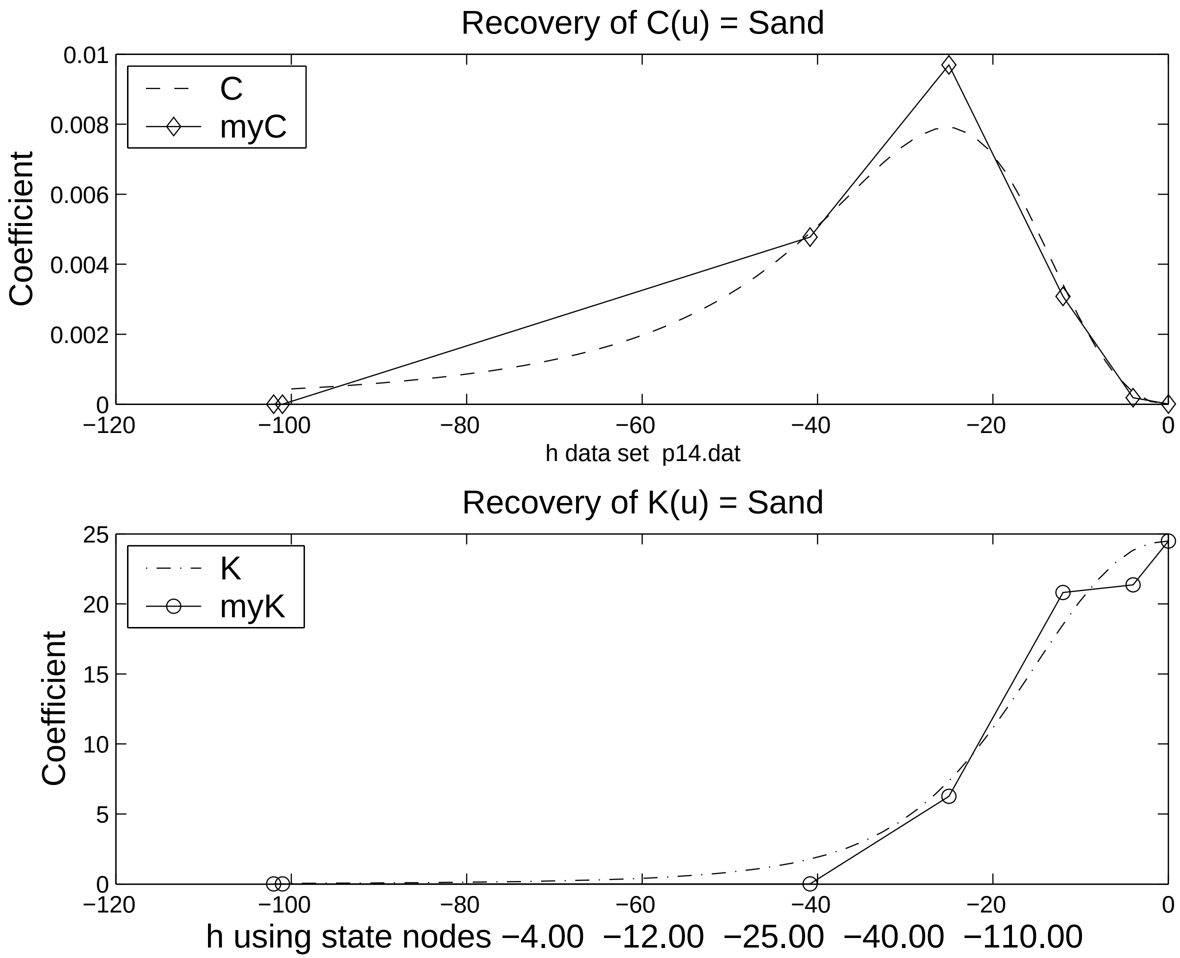
1. Dual solution requires knowledge of true coefficient. We don't know this.
2. We only approximate the integrals numerically.

But now, the Good News ...

The algorithm seems to work!



Here we successfully recovery parameters associated with silt



Here we recovered parameters associated with Sand

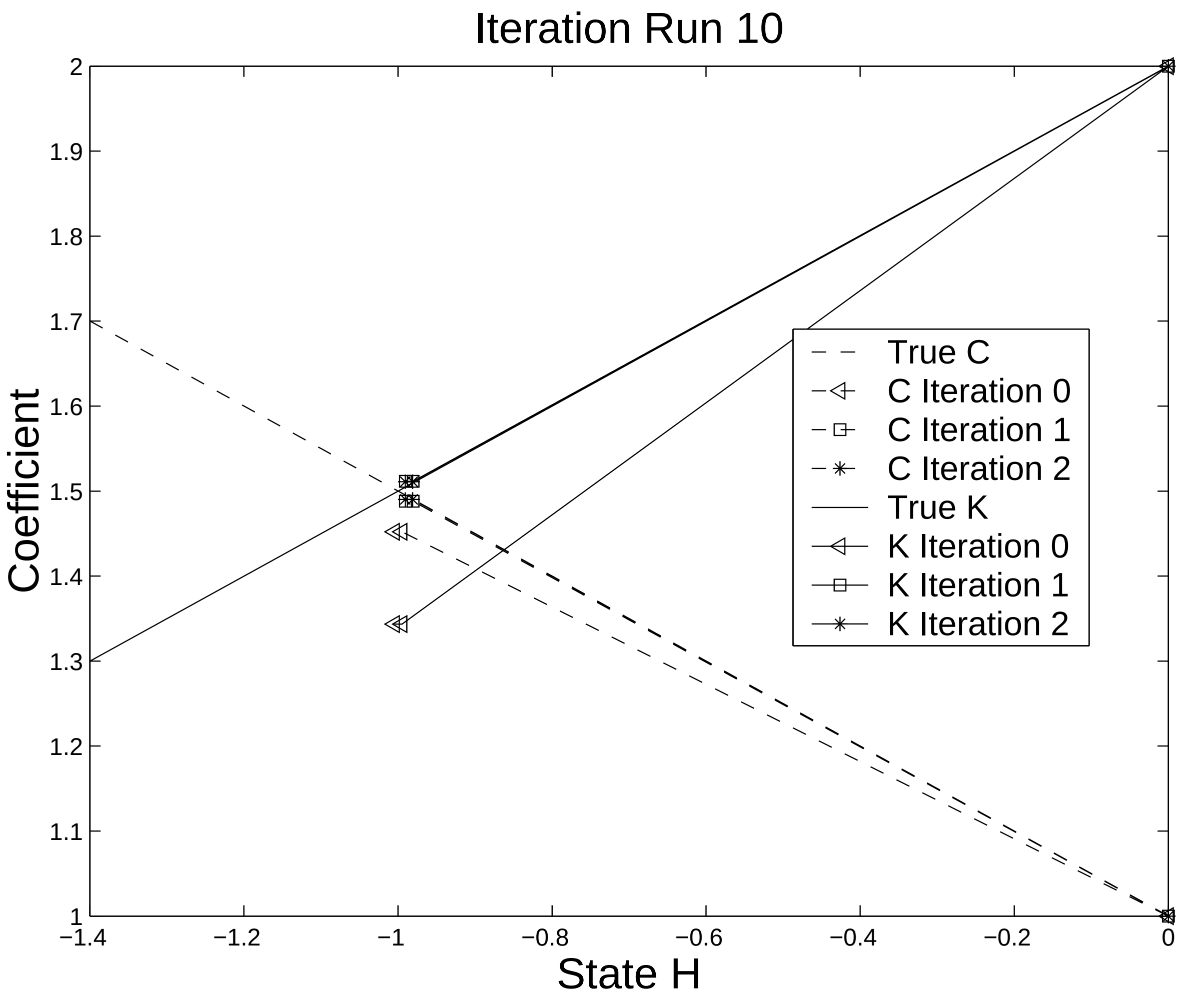
A few details

- All PDEs were discretized in space using Finite Difference.
- All PDEs were numerically integrated by a Matlab ODE suite solver.
- C^* and K^* were approximated by constants.
- Trapezoidal integration was used.

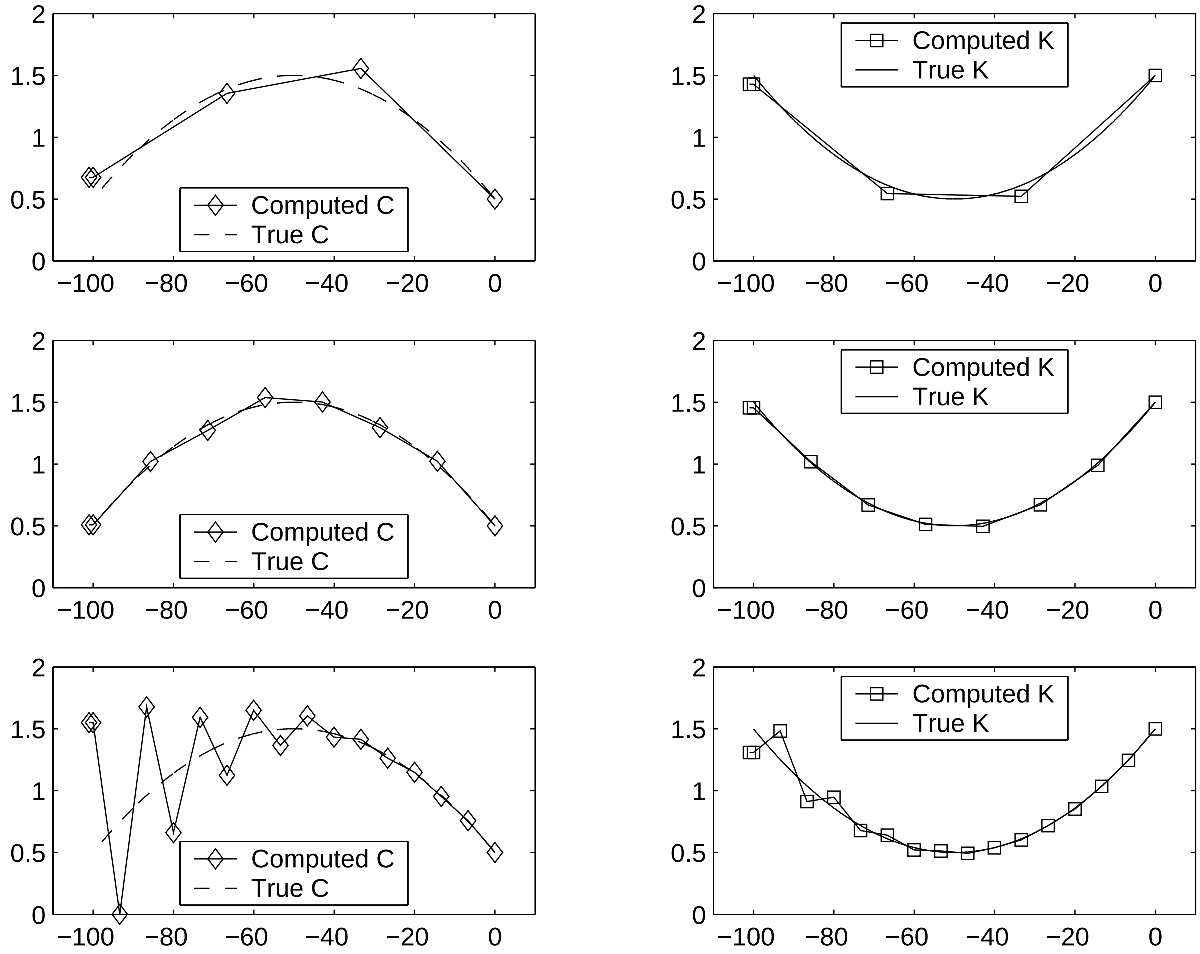
We looked at several features of the recovery

- Iteration
- Dimension of Nodal Basis
- Boundary Control
- Scaling of Inversion
- Dual Data
- Scaling of C to K
- Noise

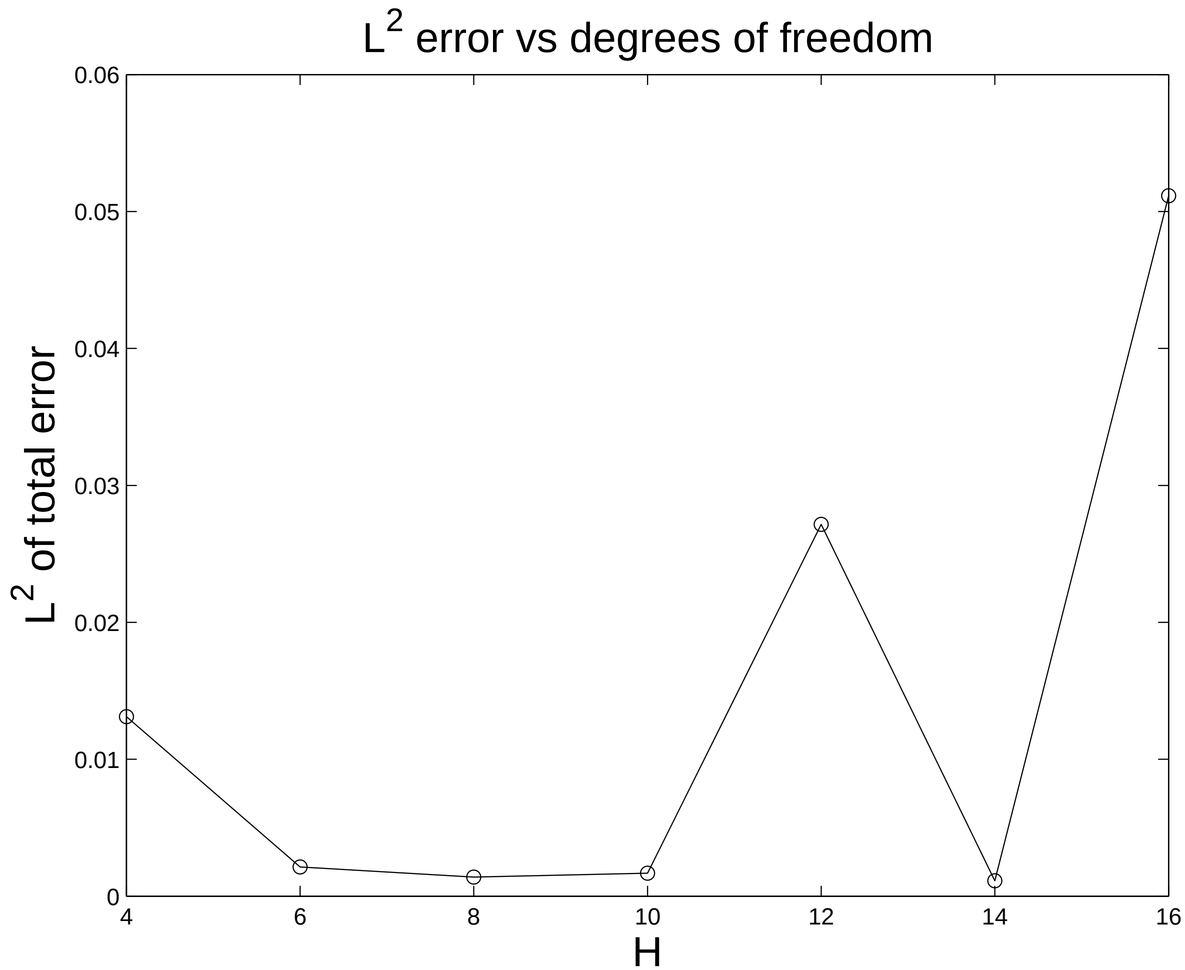
Here we iterated - by repeatedly computing C and K over a single time strip.



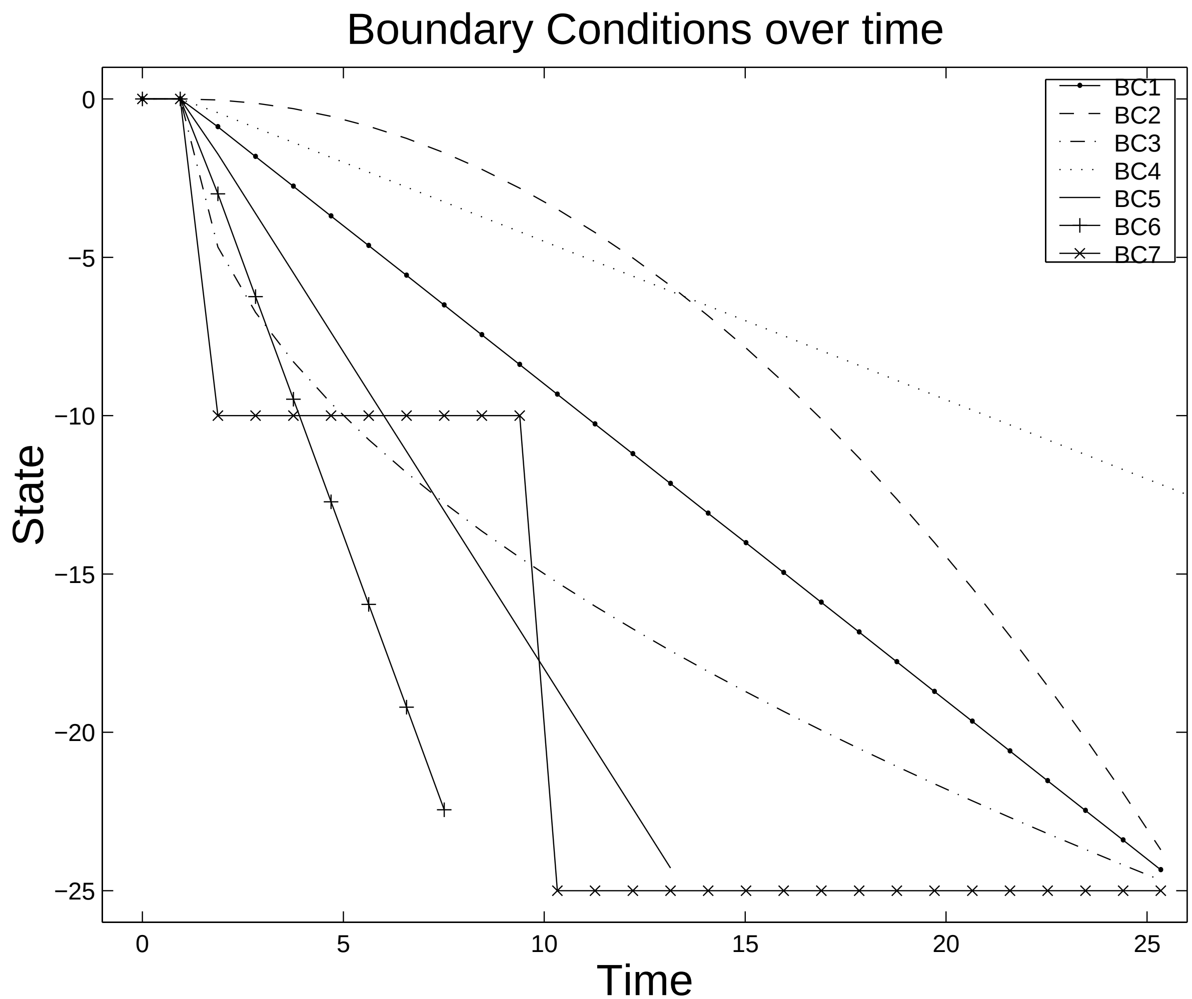
We considered the dimension the nodal basis.



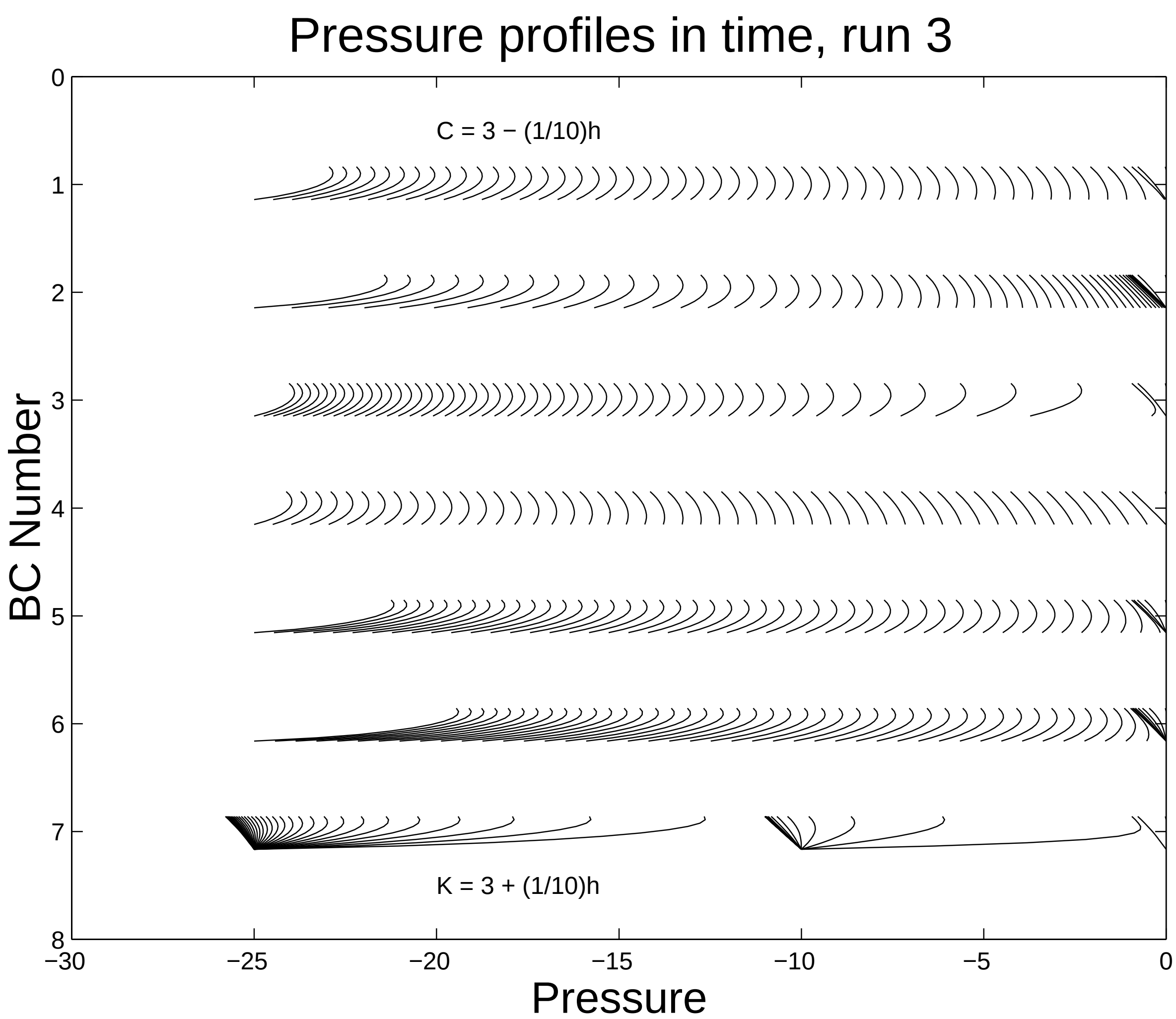
A summary of the dimension error experiments.



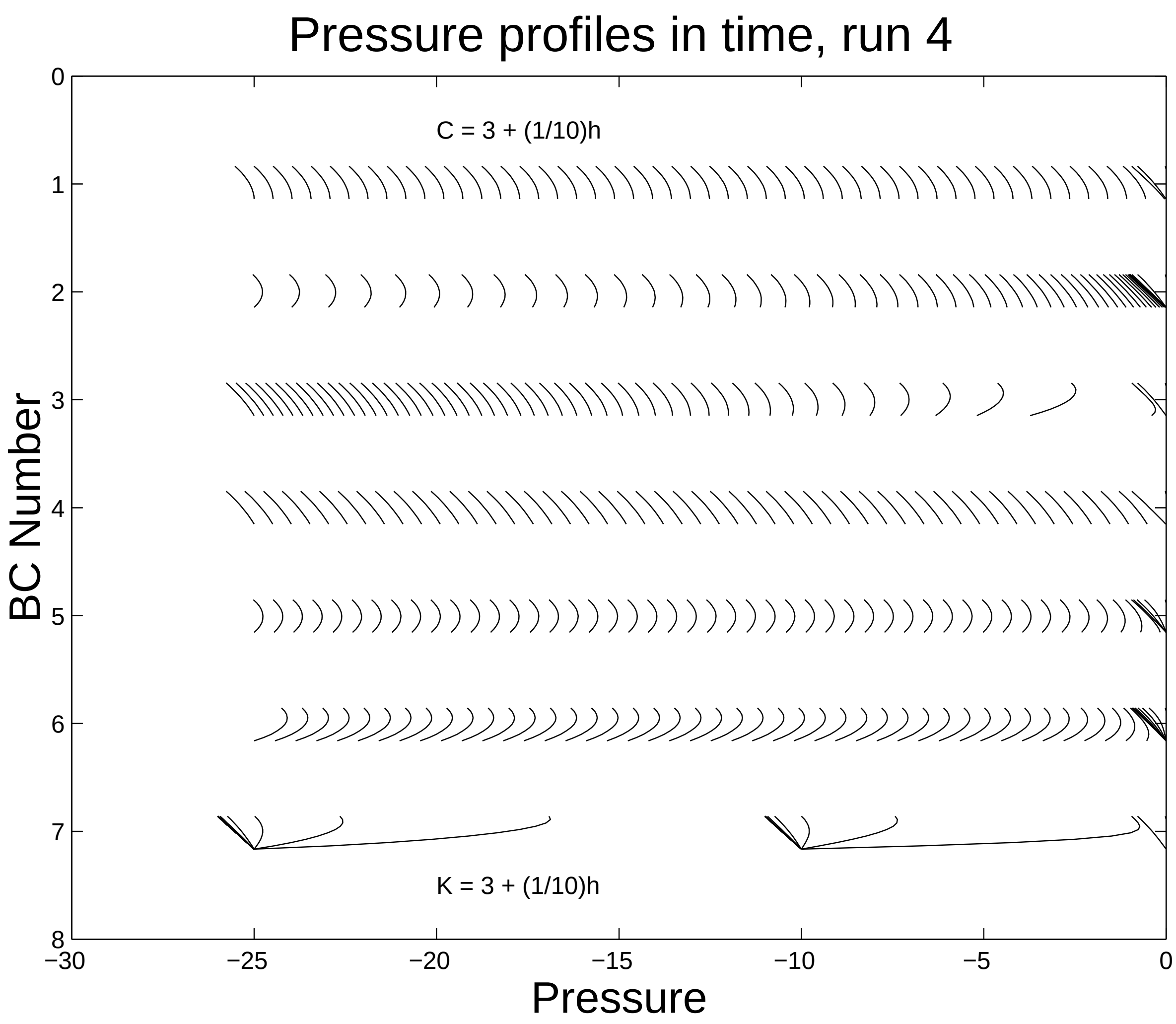
We also considered the influence of several boundary functions,



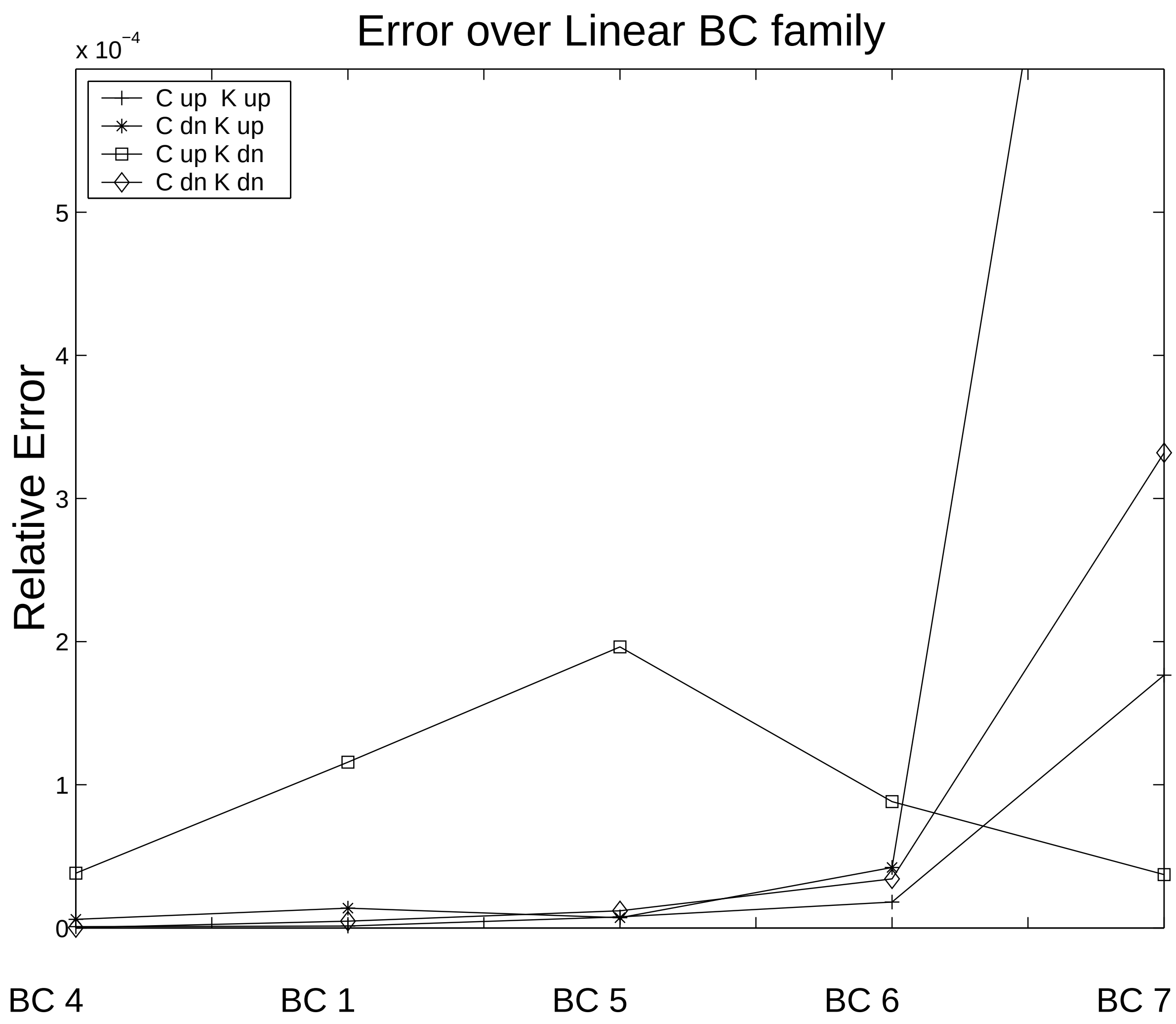
which generated the pressure profiles



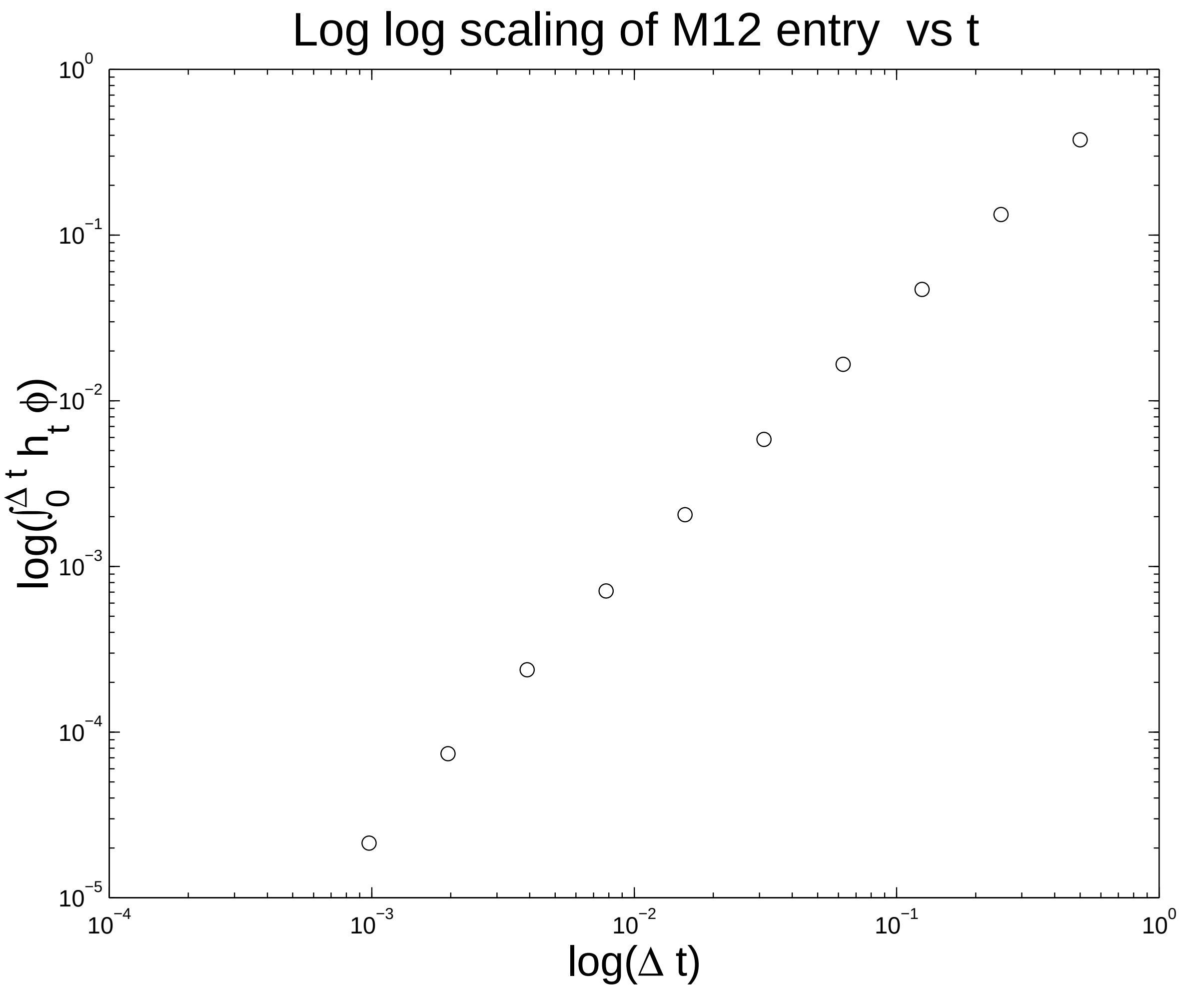
And with another set of coefficients



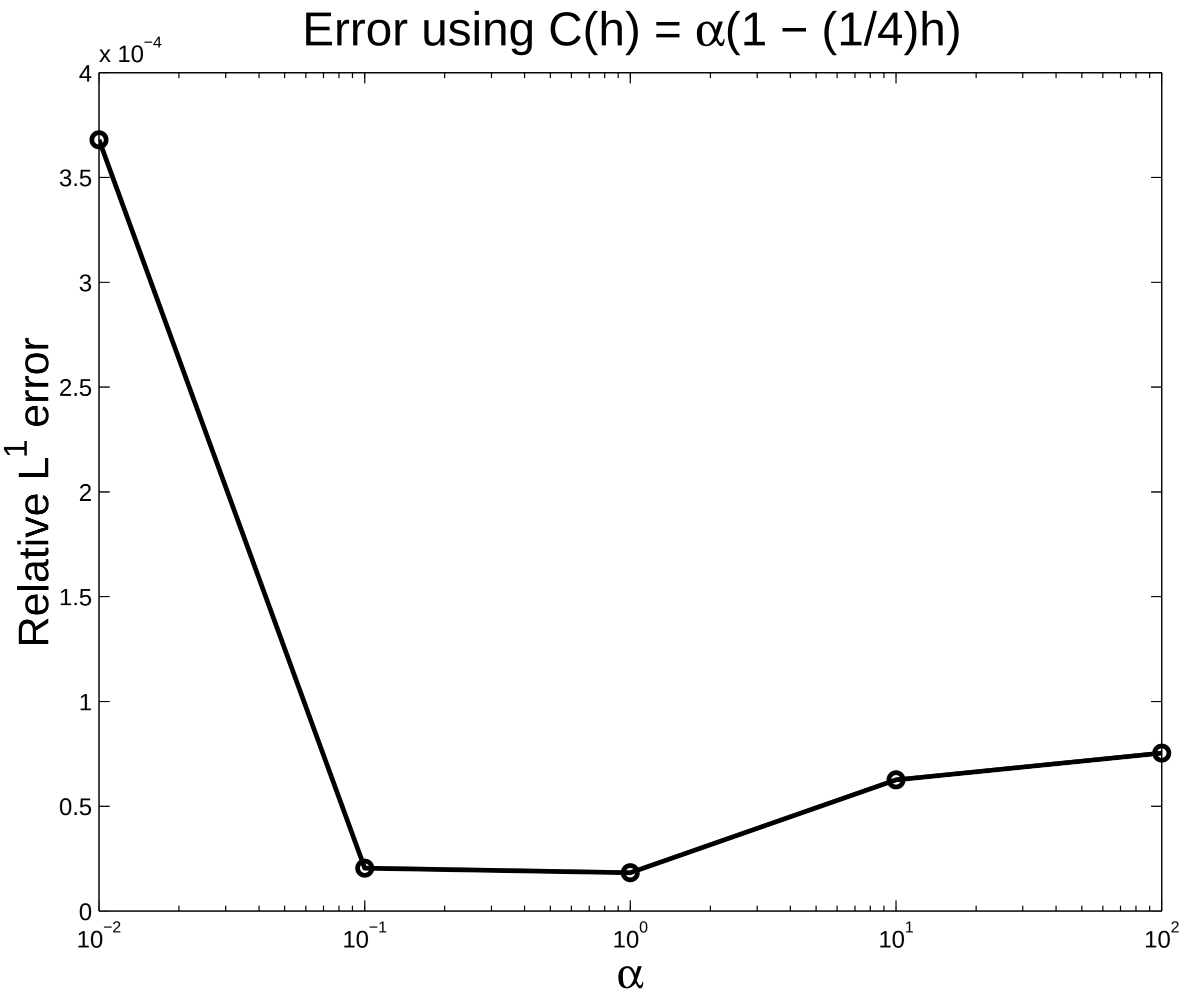
What happened to the recovery process?



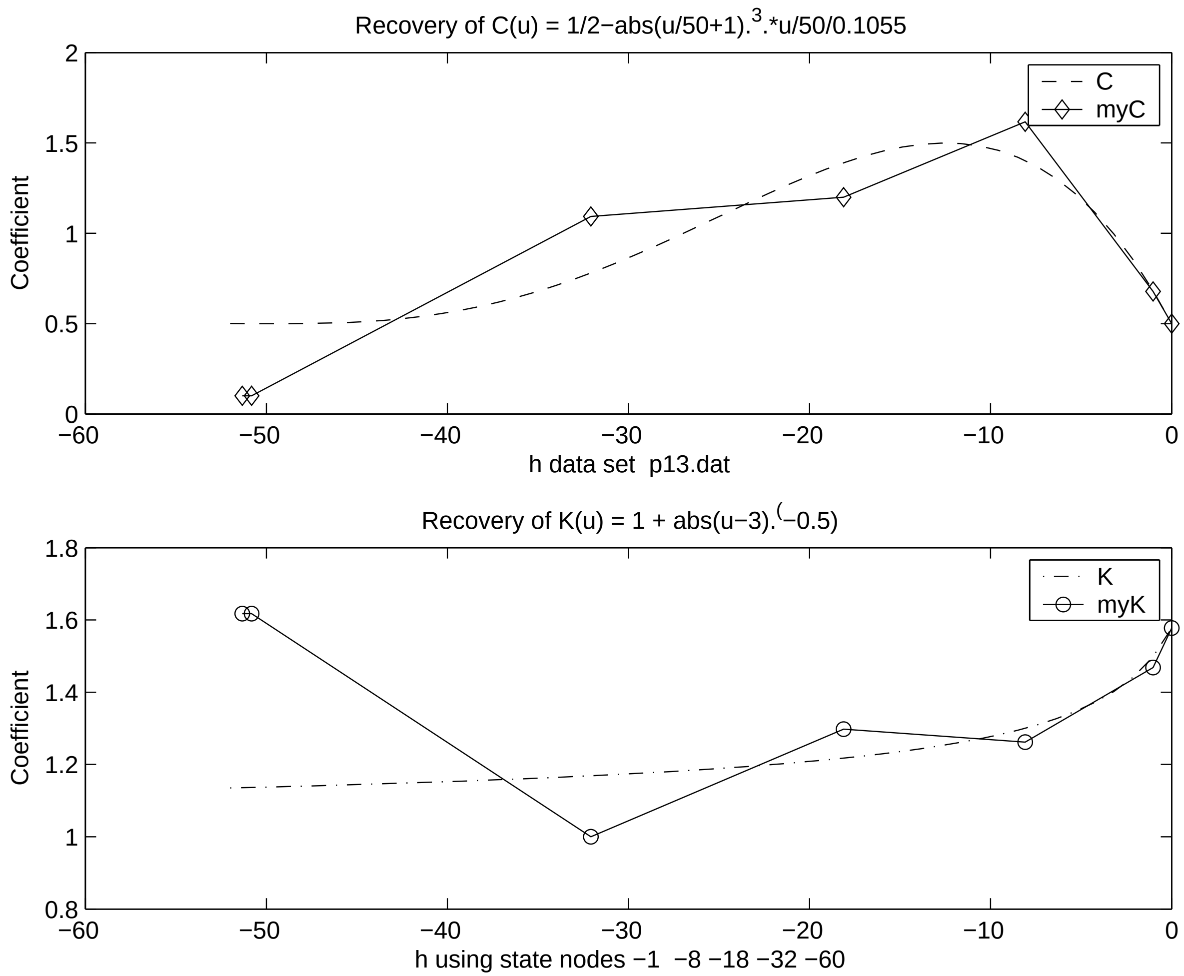
Scaling of the entries of the matrix M



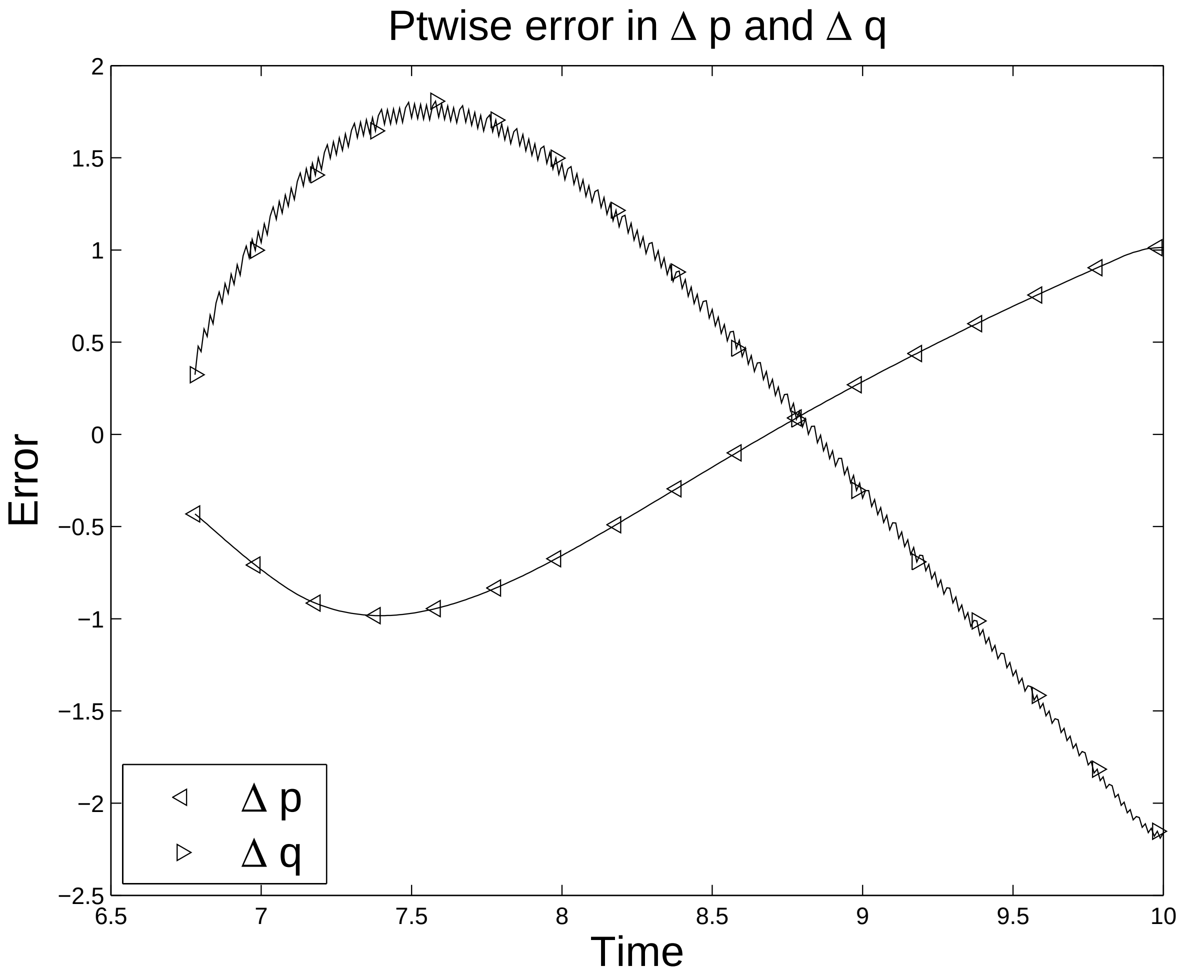
Scaling of C to K



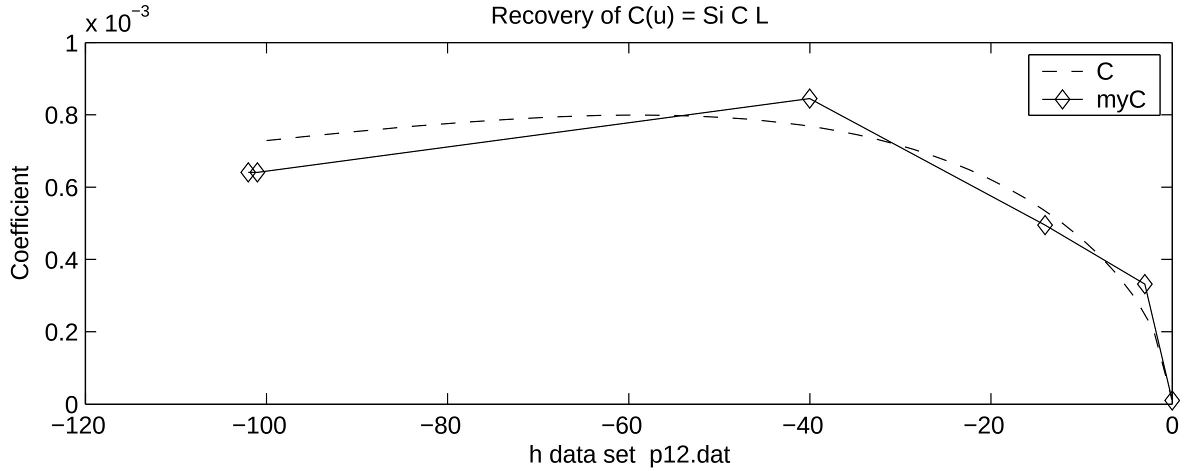
An (unphysical) example


Recovery of C(u) = 1/2−abs(u/50+1).^3.*u/50/0.1055
Coefficient
h data set p13.dat
C
myC
2
1.5
1
0.5
0
−60
−50
−40
−30
−20
−10
0
Recovery of K(u) = 1 + abs(u−3).^(−0.5)
Coefficient
h using state nodes −1 −8 −18 −32 −60
K
myK
1.8
1.6
1.4
1.2
1
0.8

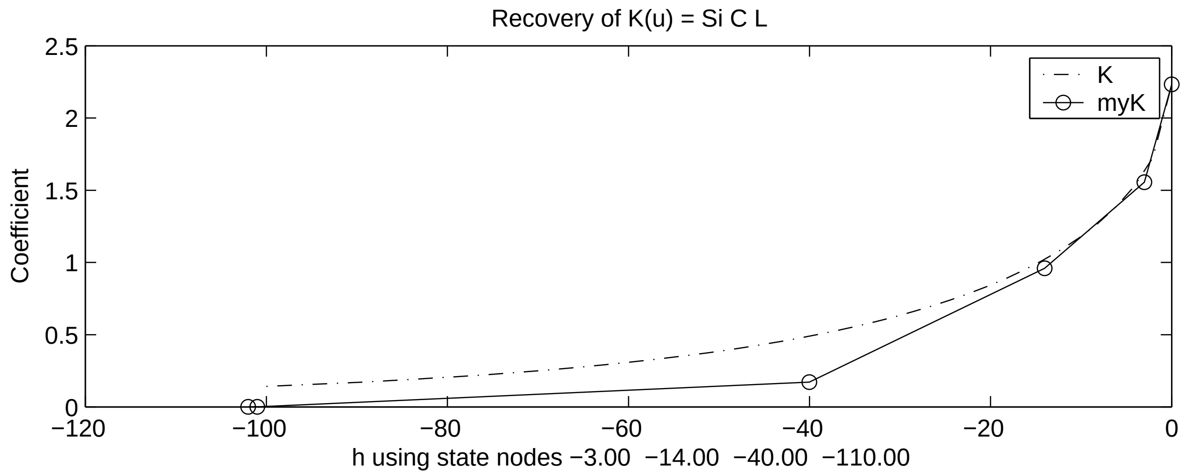
Recovered Data observations



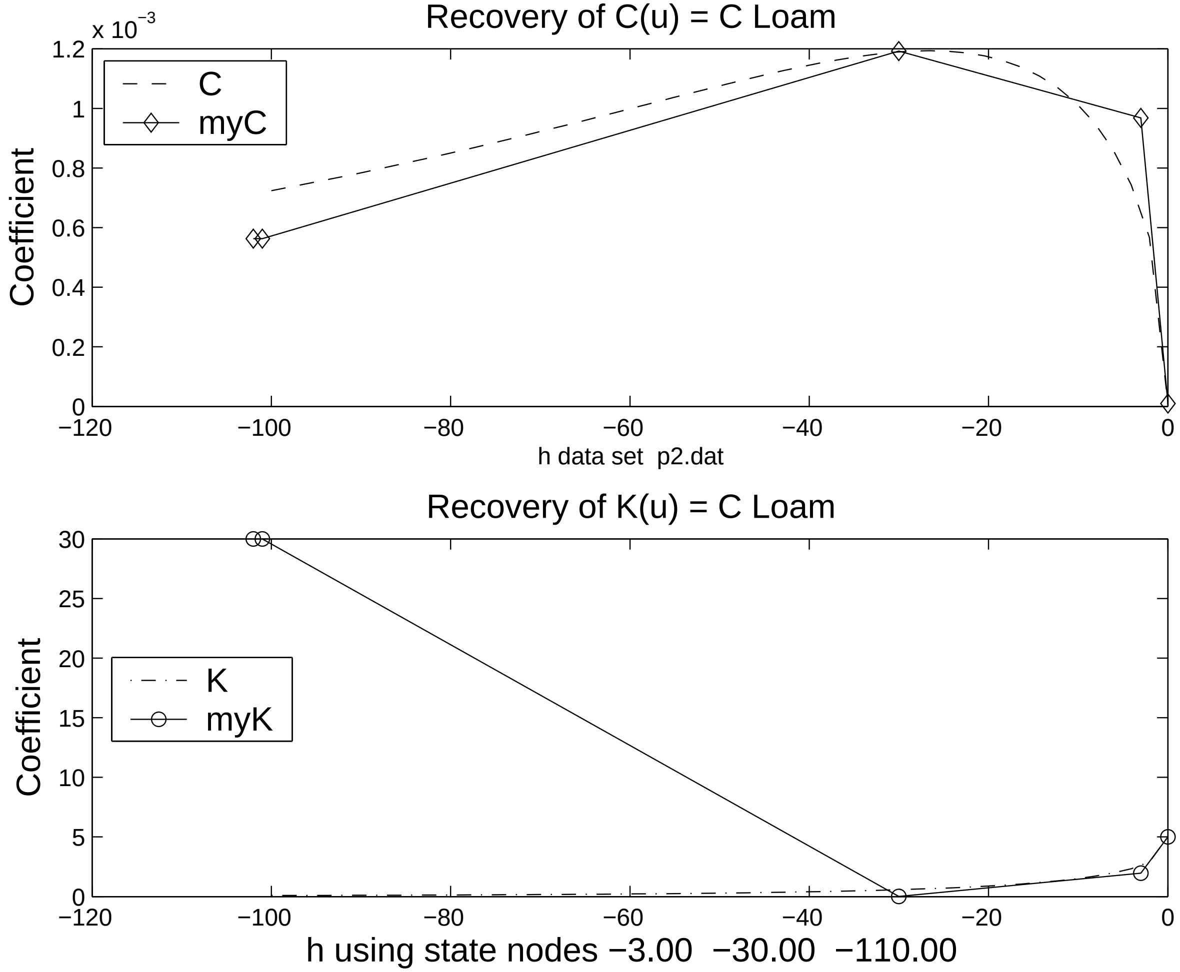
Now for some van Genuchten functions ...



Recovery of C(u) = Si C L
x 10^−3
Coefficient
C
myC
h data set p12.dat

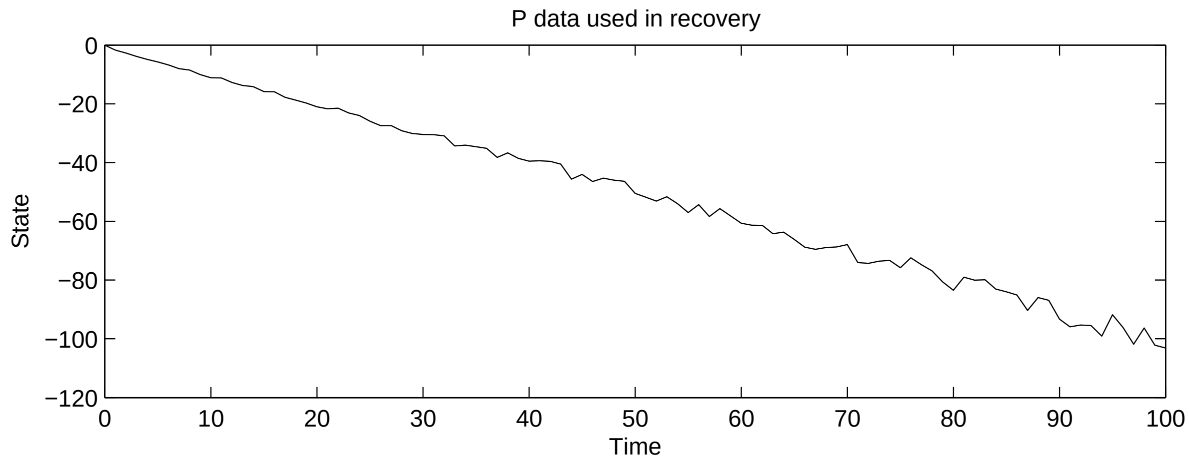


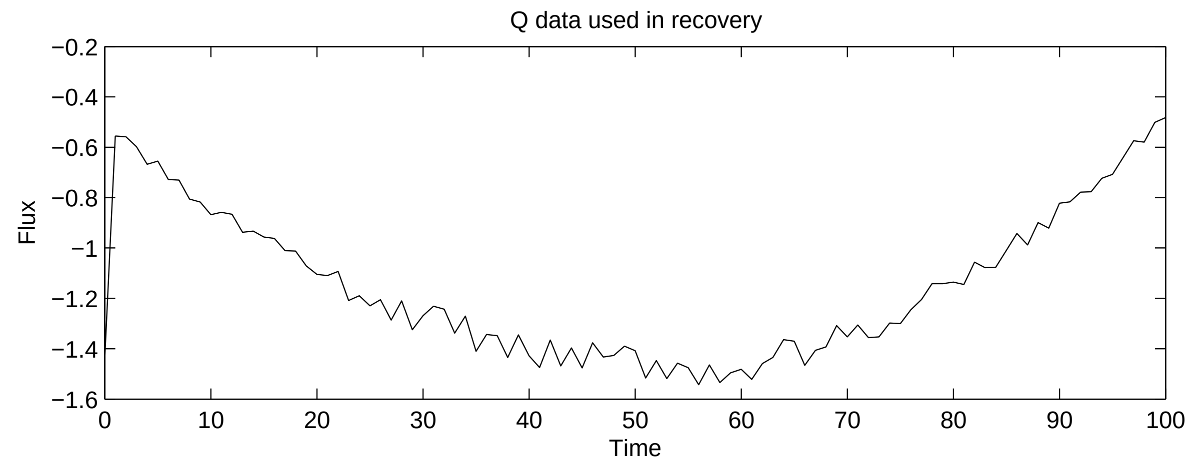
Recovery of K(u) = Si C L
Coefficient
K
myK
h using state nodes −3.00 −14.00 −40.00 −110.00


Recovery of C(u) = C Loam
x 10^-3
1.2
1
0.8
0.6
0.4
0.2
0
Coefficient
C
myC
−120
−100
−80
−60
−40
−20
0
h data set p2.dat
Recovery of K(u) = C Loam
30
25
20
15
10
5
0
Coefficient
K
myK
−120
−100
−80
−60
−40
−20
0
h using state nodes −3.00 −30.00 −110.00

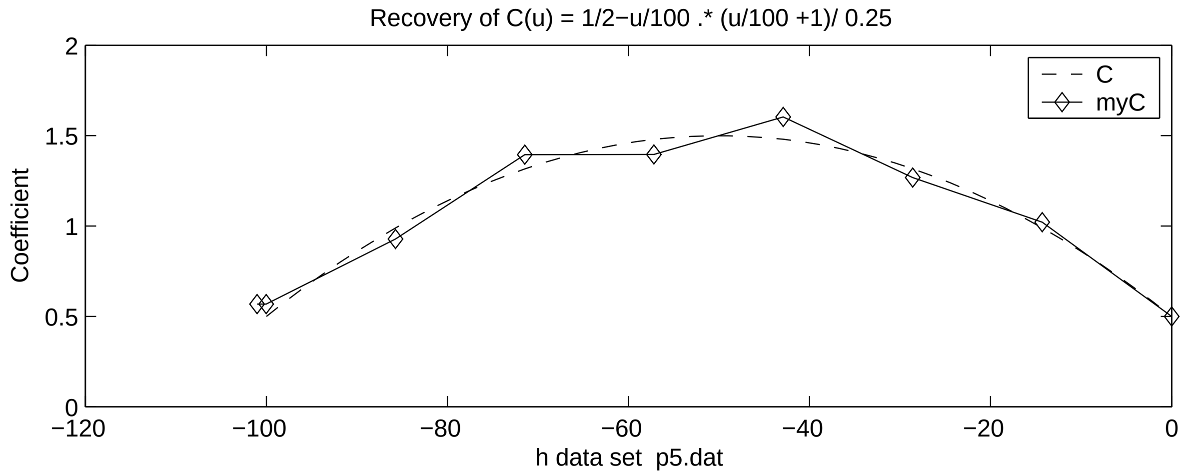
Finally, we added some noise...

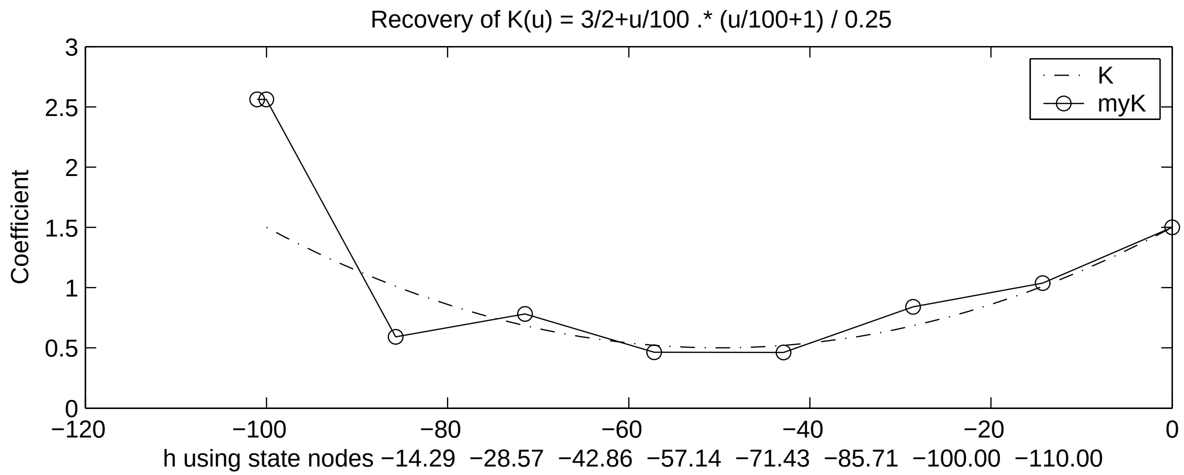
First the data



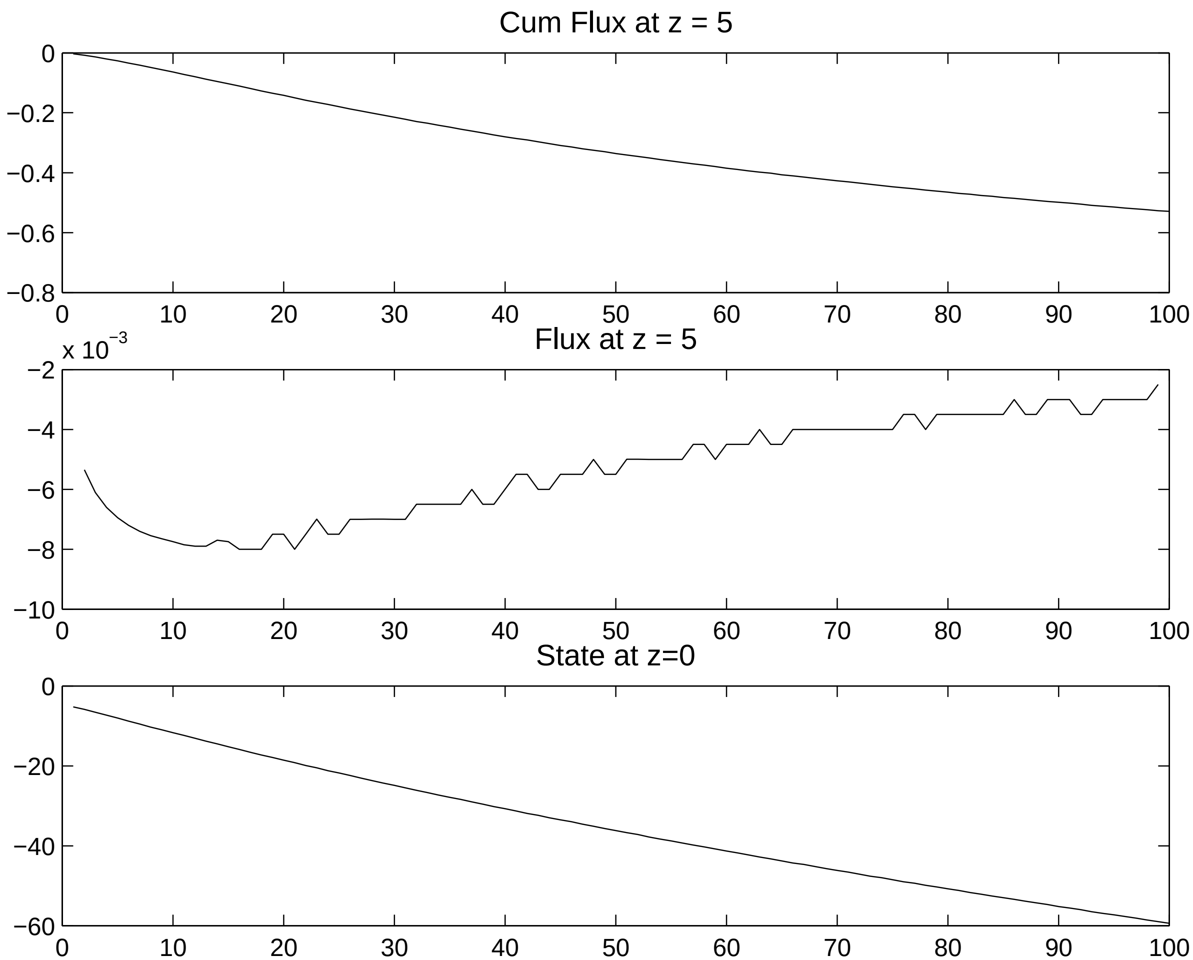


and then the recovery





Hydrus data


Cum Flux at z = 5
0
−0.2
−0.4
−0.6
−0.8
0 10 20 30 40 50 60 70 80 90 100
Flux at z = 5
x 10−3
−2
−4
−6
−8
−10
0 10 20 30 40 50 60 70 80 90 100
State at z=0
0
−20
−40
−60
0 10 20 30 40 50 60 70 80 90 100

Conclusions

- Adjoint methods allow effective means of exploring input/output mapping
- Provides both analytic and numeric information

  In addition, it appears
- robust under noise
- identification possible with more work

Future Work

- Apply to lab data
- Explore experimental parameters
- Cumulative flux formulation
- Nonlinear scaling possibilities
- Examine van Genuchten coupling
- Adaptive nodal basis
- OLS comparison