Markov Model in Time for Transport in Porous Media

Amir Hossein Delgoshaie, Yang Wong
amirdel@stanford.edu, zhiyangw@stanford.edu

Introduction

Applications for flow in porous media include:

- Groundwater pollution
- Fracking and flow in fracture networks
- Contaminant transport

Key challenges:

- **High uncertainty in physical data**
- Monte Carlo framework with multiple realizations to approximate uncertainty
- Full physics simulations in Monte Carlo framework is too expensive for practical problems

Project aims

- Machine learning approach to replace the full physics simulations in Monte Carlo framework for computational speed up

Problem Description

Physics-based simulations to obtain raw attributes:

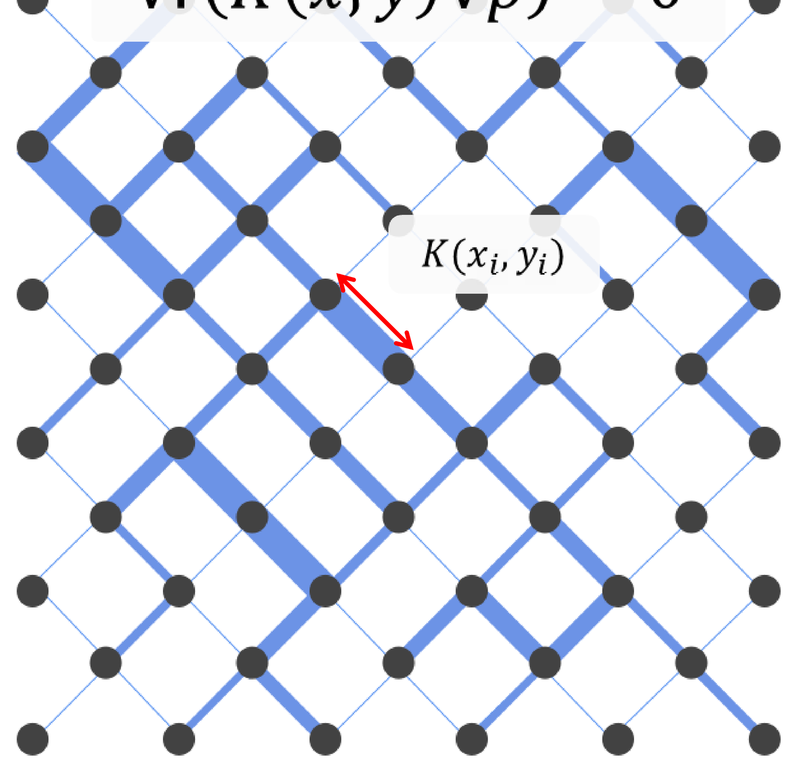
- $x_n^{(i)}, y_n^{(i)}, t_n^{(i)}$

Transformed position and times to get:

- Velocity in polar coordinates $(\mathbf{v_n^{(i)}} = \left|\bar{\mathbf{v}}_n^{(i)}\right|, \theta)$

Time averaged velocities by factor of 10 to obtain more speedup

$$\nabla.(K(x,y)\nabla p) = 0$$



Methodology

The key learning in this problem is to train the transition probabilities.

$$P\left(\mathbf{v_{n+1}^{(i)}} \middle| \mathbf{v_n^{(i)}}\right)$$

We modeled these transition probabilities in two ways:

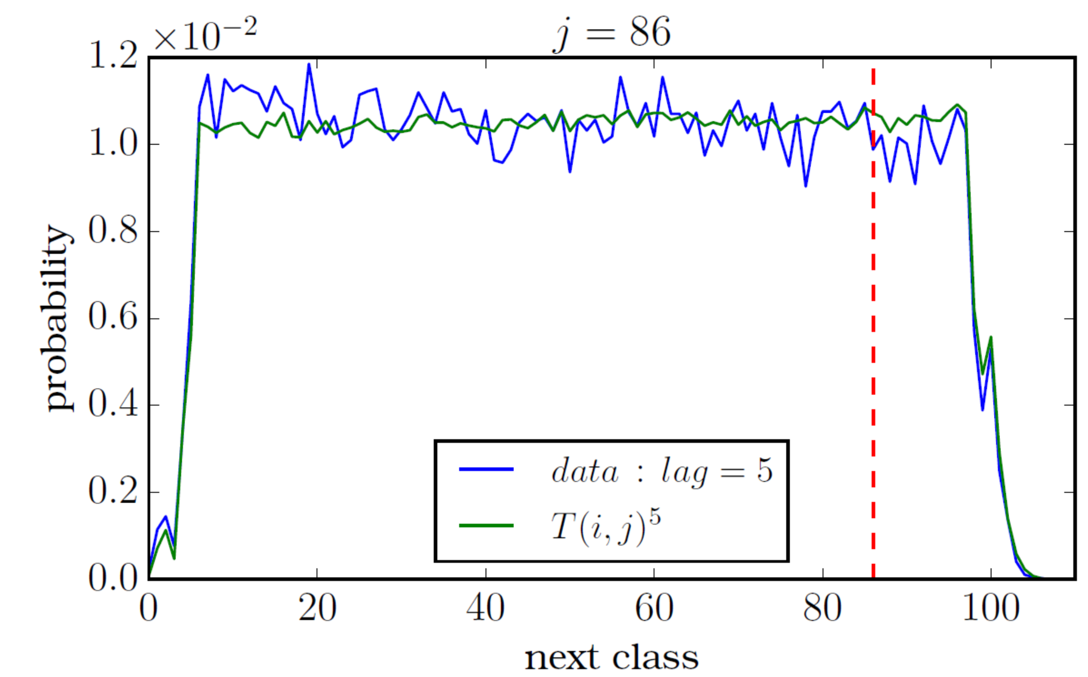
1. Discrete transition matrix

Splitting the velocity states into discrete bins

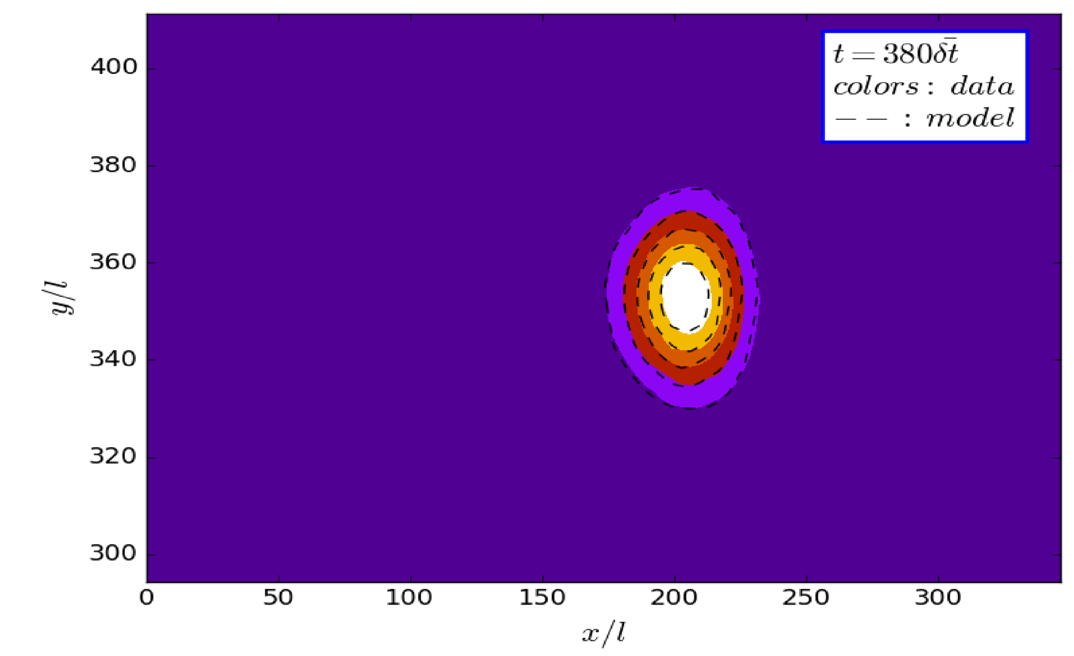
2. Kernel density estimation

fastkde Python package, doesn't require defining discrete bins.

Chapman-Kolmogorov equation was verified for lag five transitions, to test Markov assumption of velocity



By combining many velocity transitions (time averaging) results in 10 times speed up compared to physics simulation

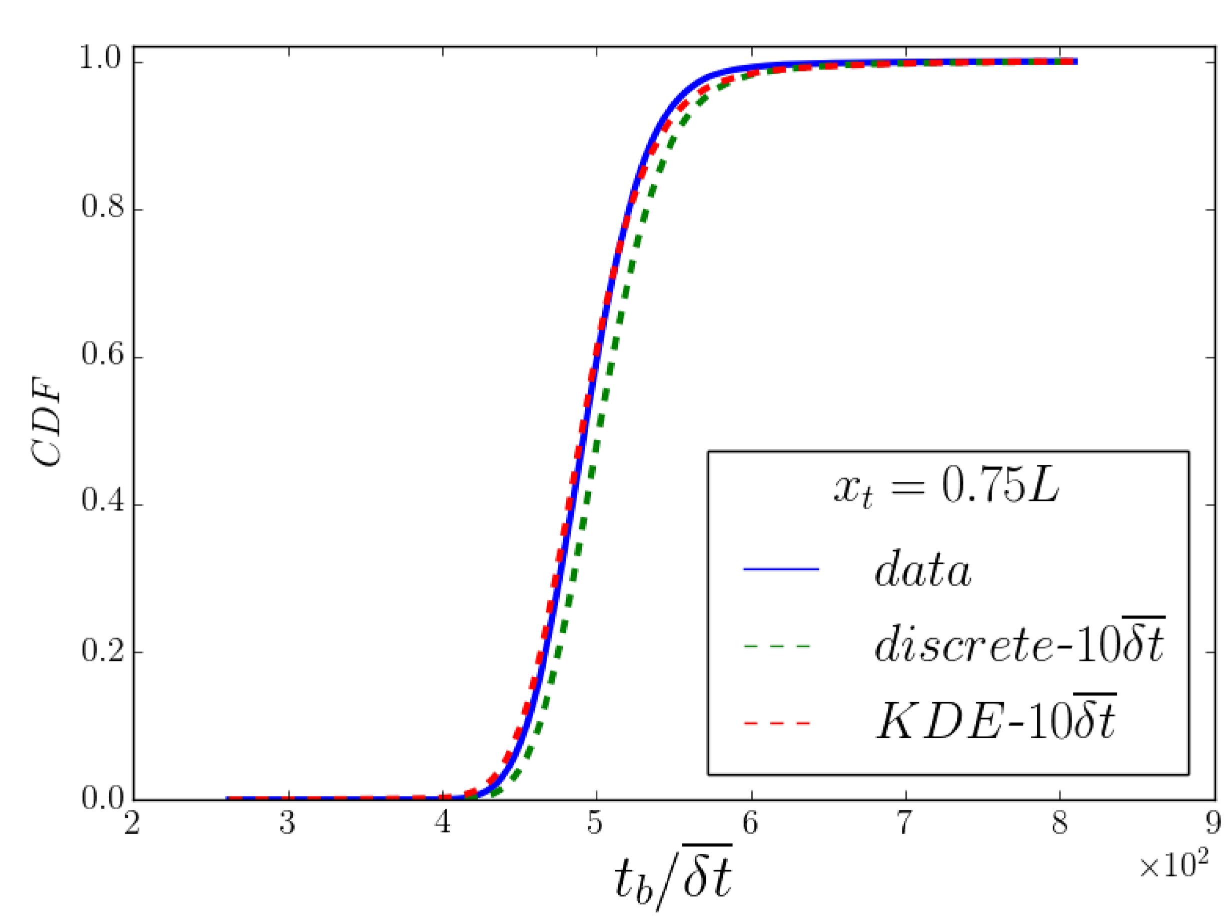


Experimental Results

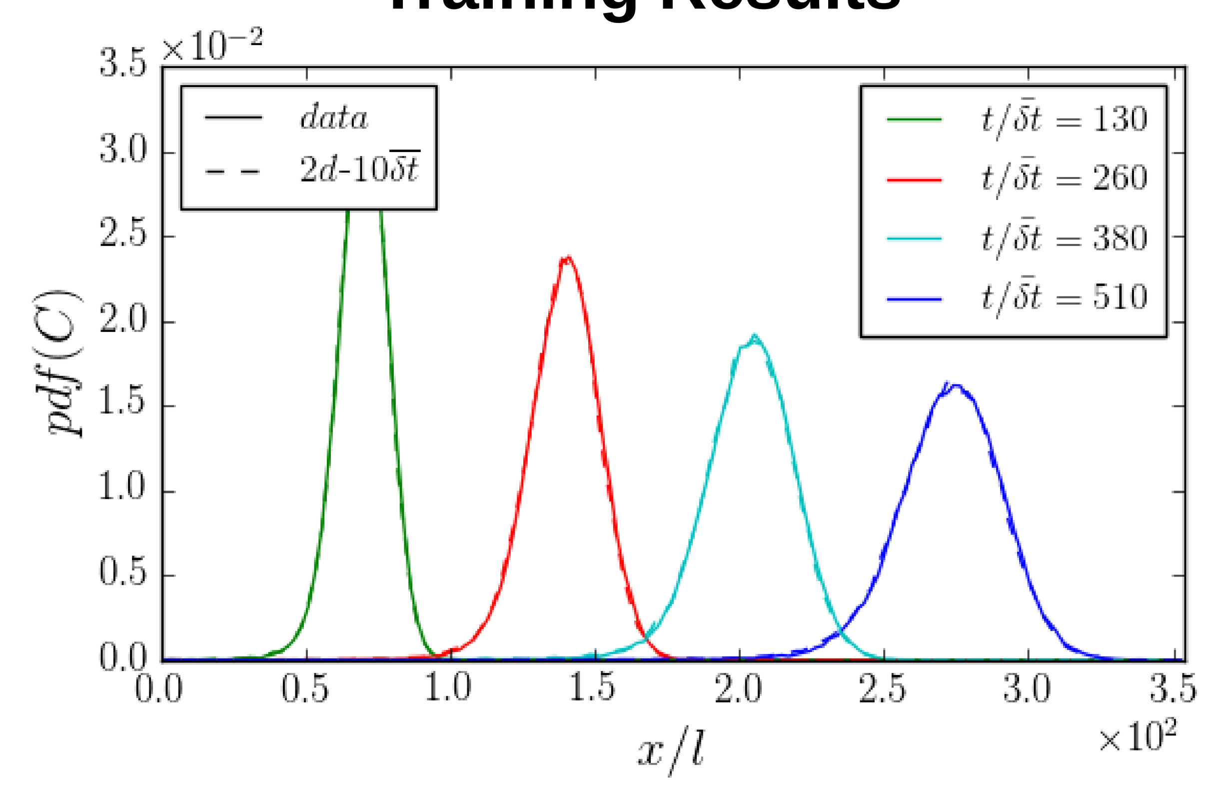
To test the potential of this Markov representation, we trained on 20% of the velocity time series and made predictions for the entire time domain. Two typical measures examined.

1. Breakthrough time at a certain location
2. Concentration map at a specific time.

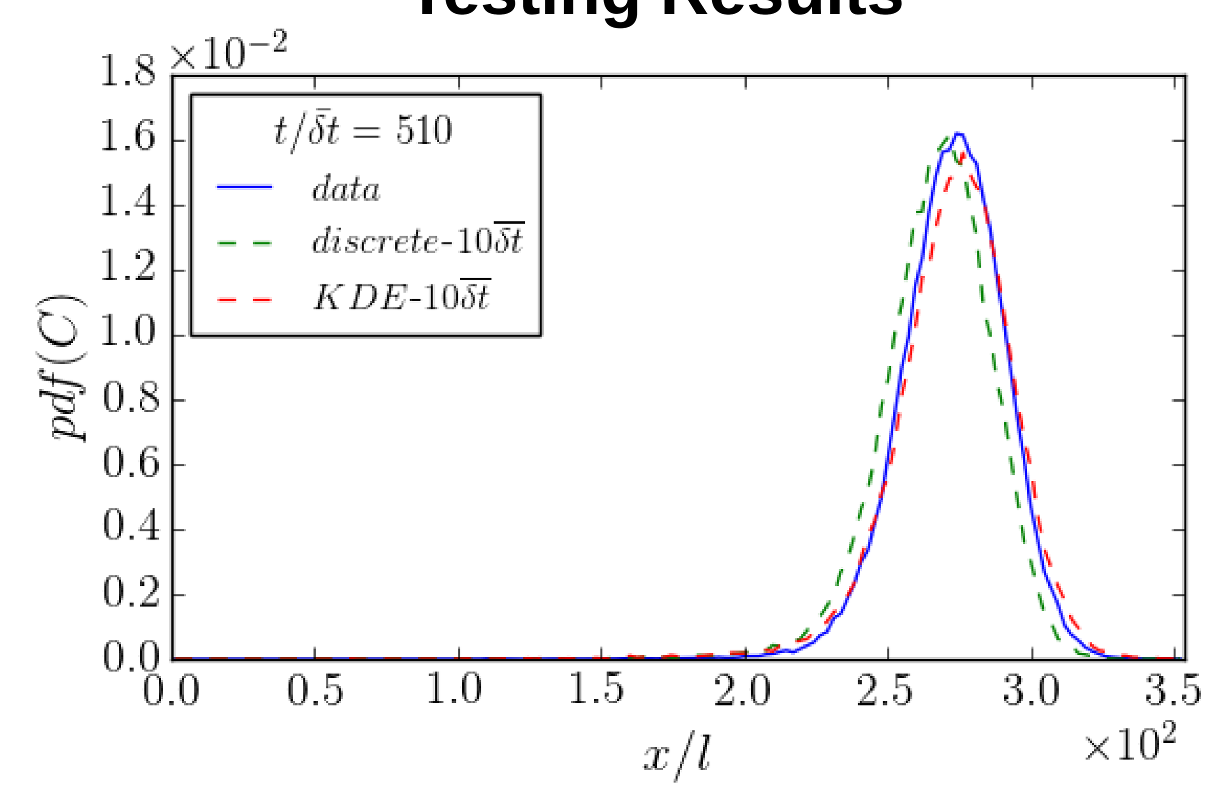
Breakthrough CDF



Training Results



Testing Results



Discussion

$\max\lvert error_i \rvert$	Training Error	Test Error
Discrete	0.042	0.206
KDE	0.128	0.106

When training with 20% of the time domain, we are able to get very good predictions for larger times. This means that we could train our Markov Model for full physics simulations at a smaller computationally feasible scale. Then use the Markov model to make predictions for a larger time domain that would be physically meaningful.

Future Work

A future extension would be to use Recurrent Neural Networks to represent the velocity transitions