Bayesian Structural Field Analysis of Large Eddy Turbulent Flow Simulation Using Probabilistic Graphical Modeling

Dr. Nicholas Scott, Modeling Scientist
Riverside Research
Open Innovation Center

Anthony Yao, Summer Intern
Ohio State University
Department of Physics

Dr. Zhen Cheng, Post-doctoral Scientist
Woods Hole Oceanographic Institution
Applied Ocean Physics and Engineering

RIVERSIDE RESEARCH
Technical Excellence. Trusted Solutions.

2019 Bayesialab Conference

Motivation: Preamble

- Commander: If I am to invest in BBN software, demonstrate to me what it can do for me first.

1) **State Estimation** - How can it help me **characterize** adversarial behavior?

2) **Optimal Learning** - How can it help me **decide** on what to do based on known goals?

- Request: Use easily understandable models!

- Answer: Use **Crowd Turbulence - Fluid Turbulence Analogy**

Outline

1) **Introduction:** Traditional Geo-Intelligence Problems

2) **DPF System Characterization**
   A) Why use DPF for system characterization?
   B) Why use DPF data for system modeling?
   C) Image Particle Dynamics Phenomenology

3) **Modeling Methodology**
   A) Global Two-Tier Processing
   i) Feature Extraction
   ii) Hidden Markov Model Parameter Learning

   B) Physical Interpretation of Emission Matrices

   C) Knowledge Gradient Policy Information Ranking

4) **Conclusions**

Introduction: Traditional Geo-Intelligence Problems

1) Military geo-intelligence electro-optical remote sensing platforms are often tasked with monitoring complex (including human) systems which **change over time**

   a) **Navy**: Radar remote sensing of riverine and ocean waters for underwater **mine detection**

   b) **Air Force**: Multidimensional imagery remote sensing of land processes for comprehending **adversarial motion**

   c) **Homeland Security:** Panchromatic remote sensing of **crowd turbulence** for adversarial surveillance

2) Traditionally, linear **optimal Bayesian estimators** have been used as state estimators to address these sorts of problems

Why use DPF for System Characterization?

- Real world systems = highly **nonlinear** and **probabilistic**
- Knowledge of the dynamical system model does **not** exist
- State estimation initially requires **model learning** or **system characterization**
- Dynamic particle fields (DPF) obtained from large eddy simulations (LES)

Optimal **temporal-based Bayesian system characterization**

- **Parameterized system model** aids in **future state assessment** and decision making

Why use DPF Data for System Modeling?

- Turbulent particle fields have strong similitude to both **marine** and **human many-body** systems of military interest

- DPF equations emanate from **turbulent fluid mechanics**

- DPF data possesses **both** particle **imagery motion** and the underlying **driving force** behind the motion

- **Both** variables necessary for robust probabilistic **system modeling**
- **Not readily available in open source data sets**

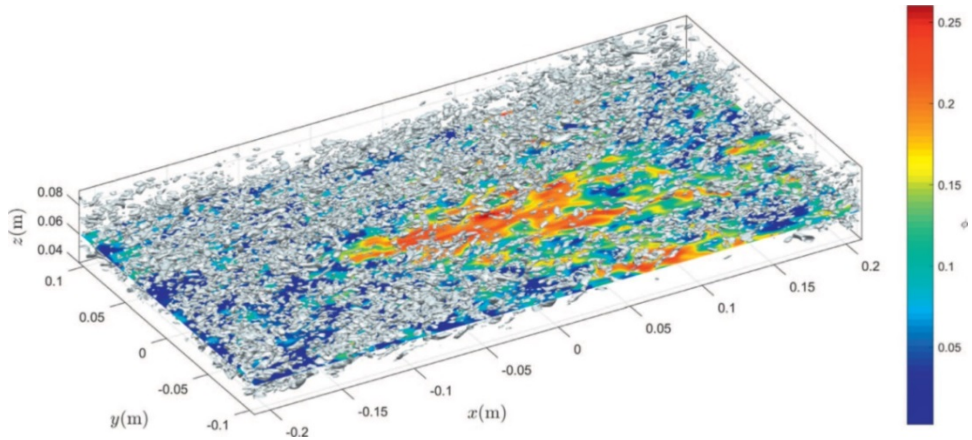
- DPF data = **noiseless and seemingly random**

  allows for **pure algorithmic exploration**

DPF System Characterization:
Imaged Particle Dynamics Phenomenology

1) Dynamic particles = point tracers *representing* different phenomena
(E.g. people or objects)

2) Modeling point: Though ***seemingly*** random and unpredictable, statistical structure exists as particles move through space/time.



3) Particle patterns emerge as particles **coagulate** into groups and **disperse**

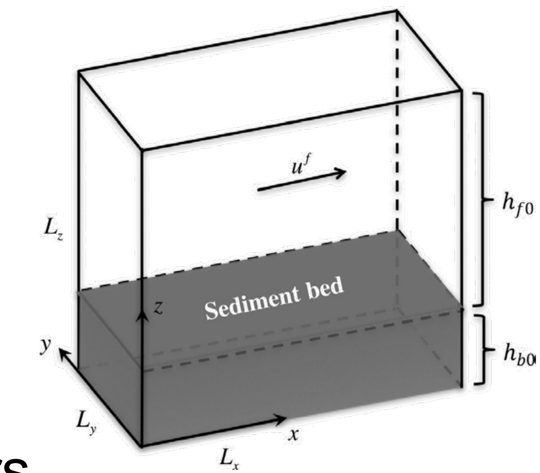
4) DPF dynamics mimics how **chaotic state** of geo-intelligence processes **with an organized or pattern-like** quality

E.g. Organized adversarial motion, coherent wakes cause by mines

Modeling Methodology: Two Tier Processing

- Objective of DPF data **system modeling:** to employ **machine learning algorithms** to create a probabilistic graphical model

- DPF data processing employs **two tiers**



A) Feature Extraction

- Data domain split horizontally into 2 layers
- Each **dimensionally reduced** to **single values**

- **Bottom layer** -> characteristic latent causal **states**
  Top layer -> characteristic surface particle **observations**

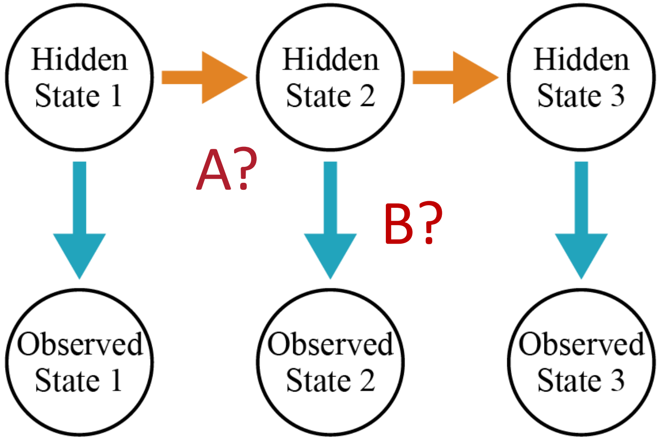
Modeling Methodology: Two Tier Processing

B) Hidden Markov (Bayesian) Model Parameter Learning

Statistical relationship between **observations (upper layer structure)** and **states (bottom layer structure)** can be **learned!**

Assumptions?

1) State- Markovian
2) Observations - independent



- **Instance counting** can be used to estimate the **transition probabilities, A and emission probabilities, B**

- Parameterized HMM allows for **system characterization** of relationship between surface and bottom

Feature Extraction from DPF Data

Global Methodology:

1) Decompose surface and sub-surface DPF using **feature extraction**

Machine ⇓ Learning

2) 9-D Feature time series array of surface and subsurface values!

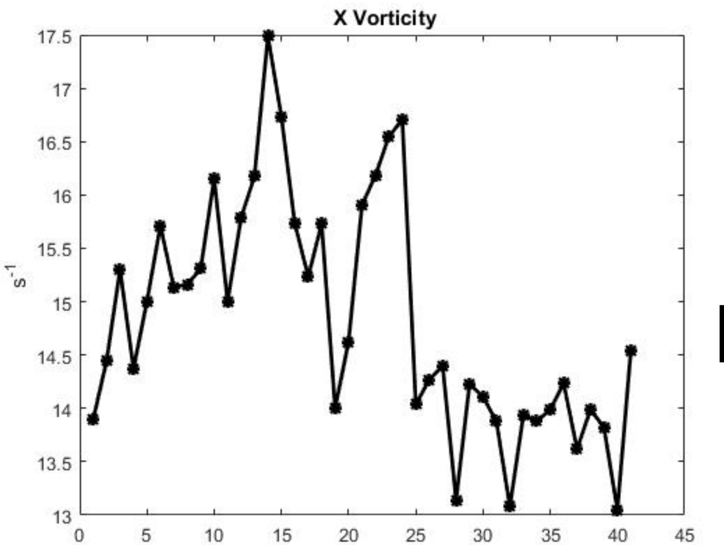
7 Surface Features (Effects)

- **a) Cross flow spatial scale**
- **b) Characteristic PCA based Concentration spatial scale**
- **c) Characteristic PCA based W velocity spatial scale**
- **d) Characteristics ICA based Concentration spatial scale**
- **e) RMS PCA based V velocity**
- **f) RMS PCA based W velocity**
- **g) RMS PCA based concentration**

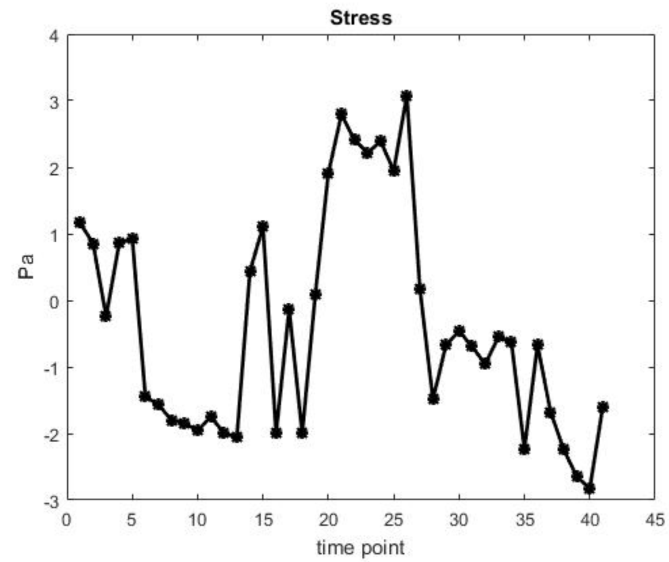
2 Sub-surface (Cause)

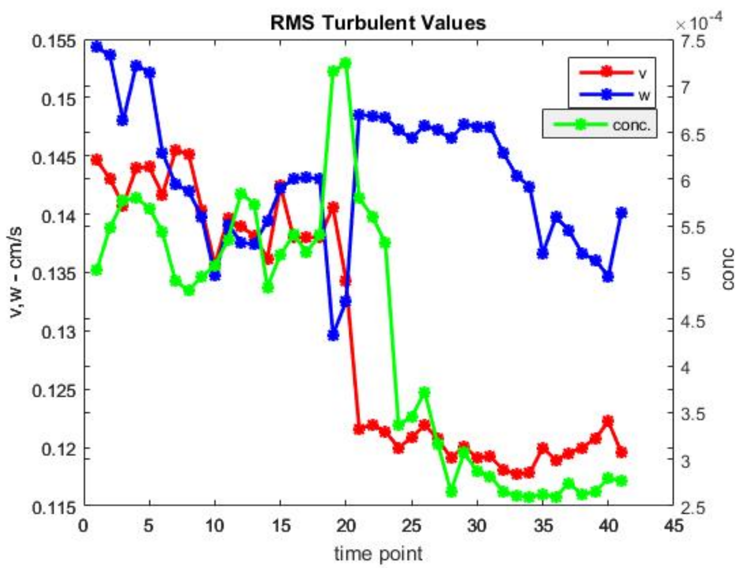
- **a) Vorticity**
- **b) Stress**

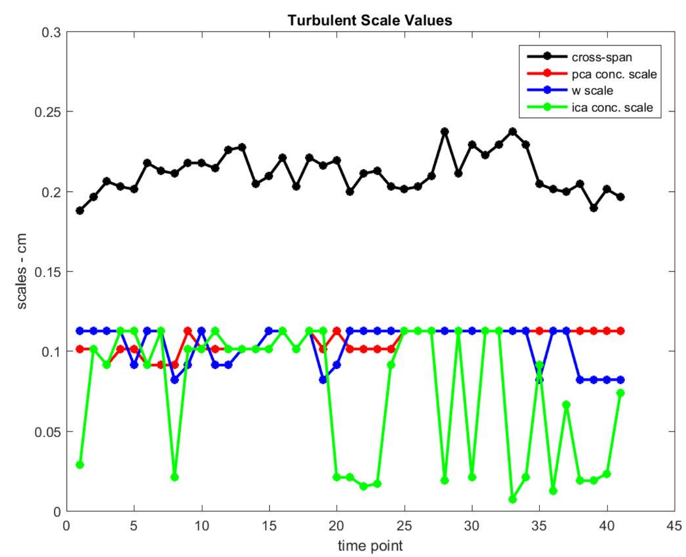
Sub-surface and Surface Feature Time Series



Data Paucity!





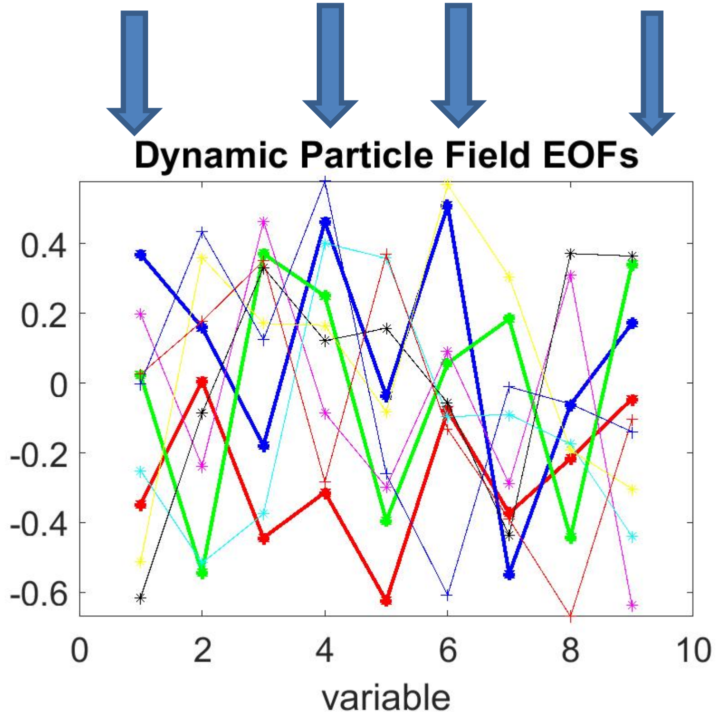


Feature Extraction from DPF Data

Global Methodology:

3) Perform **EOF analysis** to determine worthy candidates for HMM parameter estimation.

Choose surface and sub-surface variable subset



a) Vorticity (cause)
b) RMS PCA based V velocity (effect)
c) RMS PCA based concentration (effect)
d) ICA based concentration scales (effect)

9 -> 4 Features

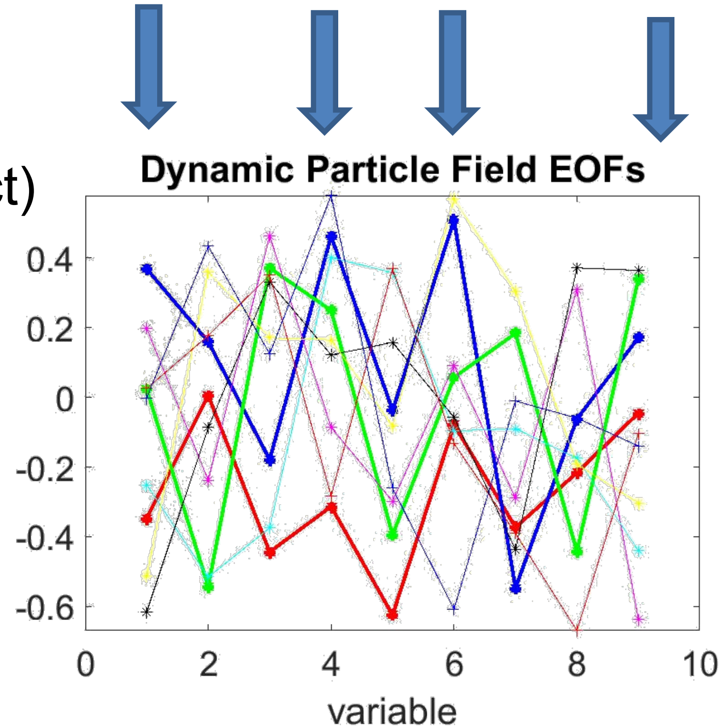
Hidden Markov Model Parameter Learning

Global Methodology:

3) Perform EOF analysis to determine worthy candidates for HMM parameter estimation. Choose surface and sub-surface variable subset

a) Vorticity (cause)
b) RMS PCA based V velocity (effect)
c) RMS PCA based concentration (effect)
d) ICA based concentration scales (effect)

4) **Use 4 features in HMM parameter estimation to estimate transition and emission matrices (1 cause, 3 effects)**



Physical Interpretation of Emission Matrix 1
Vorticity and RMS V Velocity

	Low rms v velocity 0.115 – 0.125 (m/s)	Medium rms v velocity 0.125 – 0.14 (m/s)	High rms v velocity 0.14 – 0.15 (m/s)
Low vorticity 13.0 – 14.5 (s^{-1})	0.6000	0	0.4000
Medium vorticity 14.5 – 16.0 (s^{-1})	0.0769	0.4615	0.4615
High vorticity 16.0 – 17.5 (s^{-1})	0.4286	0.4286	0.1429

Vorticity and RMS V Velocity are close **dynamic cousins**!
We expect strict **proportionality**.

Physical Interpretation of Emission Matrix 1
Vorticity and RMS V Velocity

- **Complex** emission matrix due to **non-linearity relationship** between vorticity and rms surface v velocity

Low vorticity 13.0 – 14.5 (s^{-1})	Low vel. 0.6000	Medium vel. 0	Large vel. 0.4000

- At low vorticity levels -> low rms v vel. (60%) **AND high rms v vel.(40%)**

Q: Why? **A: Flow nonlinearity**

- Low vorticity (cause) -> low stress (effect)
  -> low sediment injection into the water column
  -> flow field 'adjusts' to sediment modulation
  -> turbulent fluctuations increase

- Low vorticity levels -> high rms v turbulent vel.

Physical Interpretation of Emission Matrix 1
Vorticity and RMS V Velocity

- At medium vorticity levels -> strong % at medium and high rms v velocity levels
- At high vorticity levels -> strong % at medium and **low rms v velocity levels**
- **Q : Why?**

High vorticity 16.0 – 17.5 (s^{-1})	Low vel.	Medium vel.	High vel.
	0.4286	0.4286	0.1429

- A: High vorticity levels -> sediment flux to the surface boundary layer
  -> **dampens surface v velocity**

High vorticity levels -> low rms v velocity

Physical Interpretation of Emission Matrix 3
Vorticity and Sediment Concentration Spatial Scales

	Small Conc. Scales 0 – 0.08 (cm)	Medium Conc. Scales 0.08 – 0.11 (cm)	Large Conc. Scales 0.11 – 0.12 (cm)
Low Vorticity 13.0 – 14.5 (s^{-1})	0.5	0.1000	0.4000
Medium Vorticity 14.5 – 16.0 (s^{-1})	0.2857	0.3571	0.3571
High Vorticity 16.0 – 17.5 (s^{-1})	0.2857	0.7143	0

Vorticity and sediment conc. spatial scales **distant kinematic cousins**!
We expect a strict **inverse** proportionality.

Physical Interpretation of Emission Matrix 3
Vorticity and Sediment Concentration Spatial Scales

- Low vorticity levels -> small (50%) and **large** sediment spatial scales (40%)

  Q: Why?

  A: Low vorticity (low shear) allows for **large scale sediment amalgamation** and **small scale residuals**

Low Vorticity 13.0 – 14.5 (s^{-1})	Small Scale 0.5	Med. Scale 0.1000	Large Scale 0.4000

- Medium vorticity levels -> **small, medium, and large spatial scales** supported (~30% for all)

- Q: Why?
- A: Fluid vorticity and stress levels increase -> large spatial scales **break downs into all scales**

Physical Interpretation of Emission Matrix 3

Vorticity and Sediment Concentration Spatial Scales

- High vorticity levels -> 71% **medium** spatial scales
  -> 0% **large** spatial scales

High Vorticity 16.0 – 17.5 (s^{-1})	Small scale	Med. Scale	Large Scale
	0.2857	0.7143	0

- Q: Why?
- A: High vorticity levels -> high shear -> **destroys** large spatial scales

- **Low Vorticity –> small and large spatial scales**
- **High Vorticity –> medium and small spatial scales**

a) Weak enough to **support** large spatial scales
b) Strong enough to **destroy** large spatial scales

- Applicable **geo-intelligence systems describing adversarial behavior**
- E.g. Human systems feel 'stress' (cause) and coagulate and disperse (effect) in complex ways!

Knowledge Gradient Policy Information Ranking

Consider the problem:

1) 9 measurement (7 surface and 2 sub-surface) array

2) Leadership **projects future values** for 9 variable state

3) Leadership has **limited amount of resources** to take data

Q: What **order** should the variables be sampled over time to reach the projected goal state?

Question of **HOW** to collect information efficiently.

A: Knowledge gradient policy (KGP) processing = optimal learning

Knowledge Gradient Policy Information Ranking

- Estimate **path** towards the mean goal state S^* $(\theta^*, \sum^*)$

  1) Assume a prior mean vector θ^0 and covariance matrix $\sum^0$
     Mean goal state vector = [# # # # # # # # #]

  2) S^n $(\theta^n, \sum^n)$ = Bayesian belief state at **time n**

  3) **Learn or approach** goal state by

     a) **sampling** the data mean turbulent feature values (information sources)

     and

     b) **choosing** 1 variable out of the 9 at every n

- The criterion or policy used in choosing = **knowledge gradient policy**

Knowledge Gradient Policy Information Ranking

- Knowledge gradient =

  i) amount by which the state improves if feature $x' = x_M$ from M=9 array is selected

  ii) **marginal value** of a measurement in terms of **information value** gained

- **Information value** measured via utility function $X^{\pi,n}(S)$

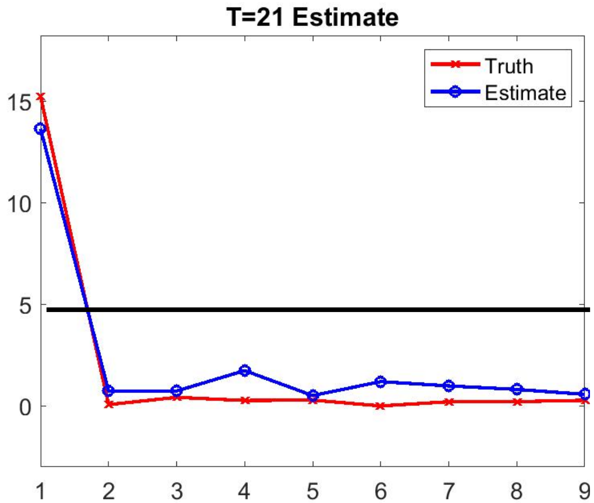
- Optimal decision choice = choice that causes largest **change** in $X^{\pi,n}(S)$
  -> maximizes **expected reward**
  -> minimizes opportunity cost

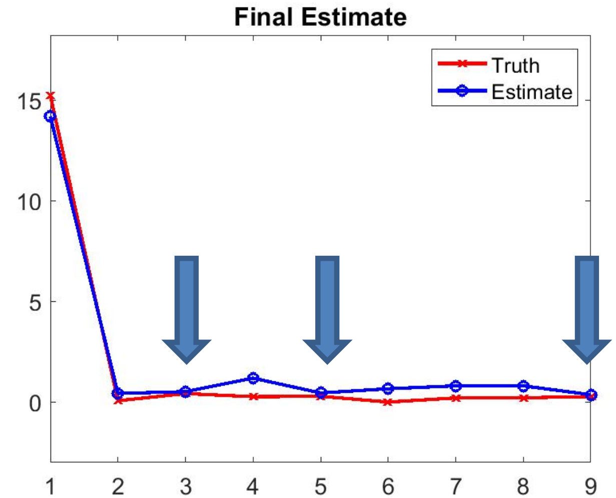
- Updated Bayesian state produces an **optimal state path through time**

~ **Method of steepest descent** $\mathbf{x}_{k+1} = \mathbf{x}_k - t_k \nabla f(\mathbf{x}_k), \quad k = 0, 1, 2, \ldots,$

Knowledge Gradient Policy Information Ranking

Prior = 5





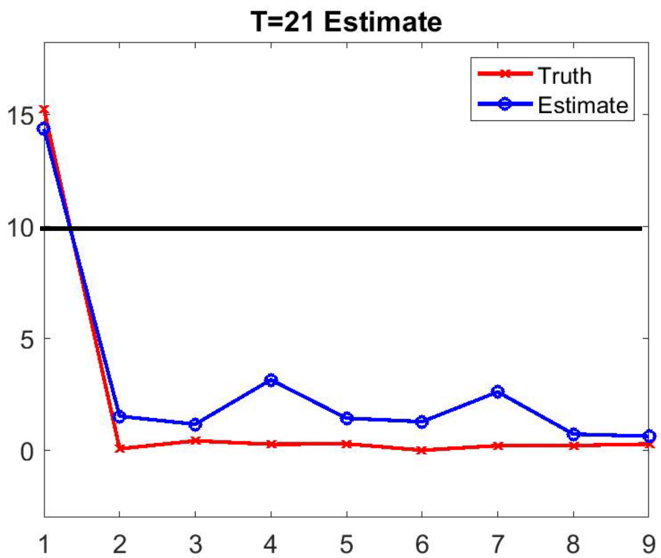
Truth = red
Estimate = blue

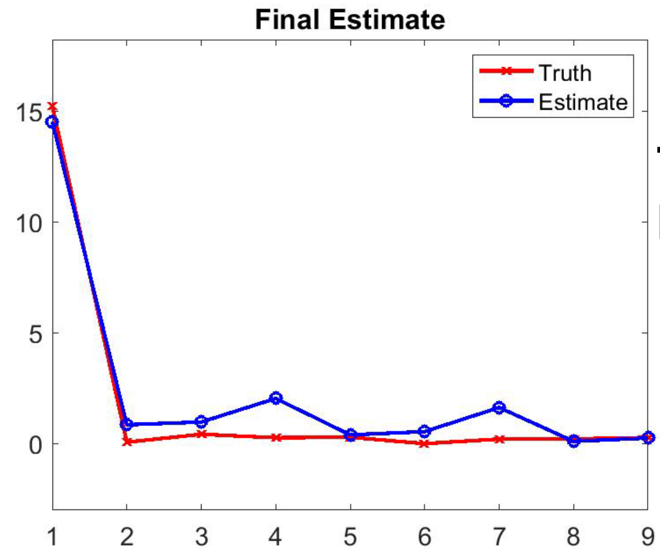
Var 1 = true mean value **Var 2-9 = 2 X true mean value**

- T = 21 estimate **less accurate** than T = 41 estimate
- KGP algorithm needs time to approach to truth
- Surface variable 3, 5, and 9 **converge first** (cross flow spatial scales, RMS w velocity, ICA based concentration spatial scales)
- **Some variables may be more significant in Bayesian goal state march!**

Knowledge Gradient Policy Information Ranking

Prior=10



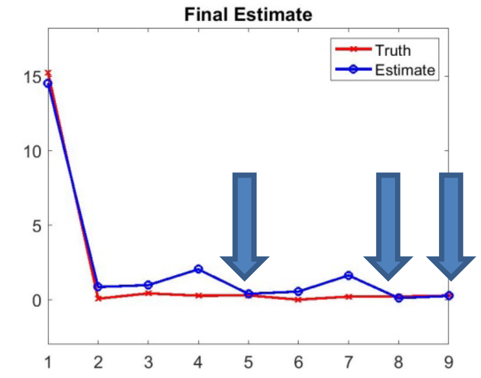


Truth = red
Estimate = blue

- Both estimates are overall **less accurate** than previous Prior = 5 case
- Vorticity is estimated very well compared to the Prior = 5 case
- **Lack of overall variable convergence** suggests that **Prior = 10 is too high** and more convergence time is needed

Knowledge Gradient Policy Information Ranking

- RMS surface w **velocity (5)**
- PCA-based surface **w velocity scales (8)**
- ICA-based surface **concentration scales (9)**



converge first !

Different set of variables to consider when seeking to attain goal state (because prior =10)

<u>Overall KGP Conclusions:</u>

- Goal state convergence time T **varies** depending on:
  1) variable correlation **(covariance)** 2) **priors assumed**.

- **What we believe affects how we obtain goals!**

Conclusions

- **Bayesian algorithms** applied to DPF data can be used to rationally understand simulated turbulent shear flow structure

1) **HMM models** captures the dynamic, sediment-induced **nonlinear flow dampening** in a sparsely sampled fluid flow.

2) **KGP algorithm** provides rational, resource saving guidance as to **how** to attain a goal state based on **Bayesian learning** 'powered' by information source covariance.

- **Not just data interpolation** but a crude way to inject a rudimentary sense of 'mind' using a functional policy for data paucity problems

- Developed algorithms possibly applicable to crowd turbulence

- **Presently seeking ways to utilize BNN software to automate and ease calculations.**