Optimization and Uncertainty Quantification Based on the Four-Dimensional Variational Method

Hiromichi NAGAO

Earthquake Research Institute, The University of Tokyo

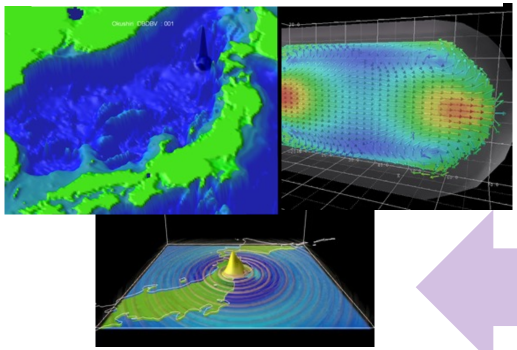
International Workshop on the Integration of
(Simulation + Data + Learning):
Towards Society 5.0 by h3-Open-BDEC

December 3, 2021

Data Assimilation

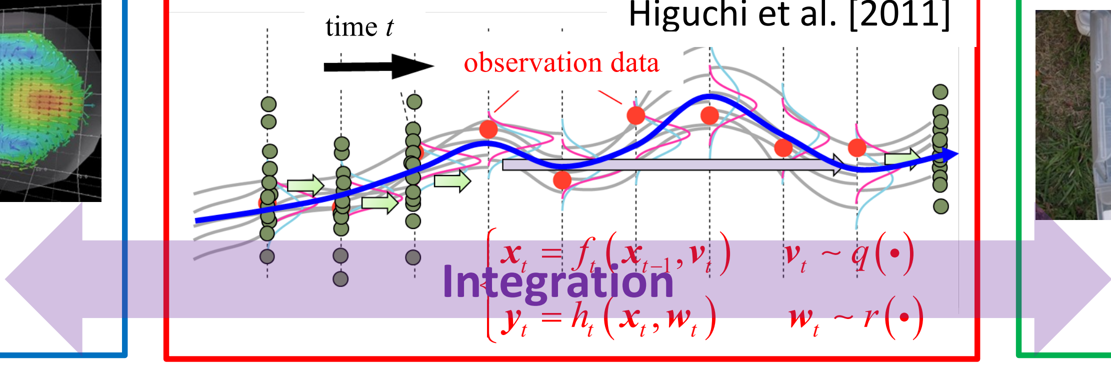
Integration of numerical simulations and observational data based on Bayesian statistics

Numerical Simulations



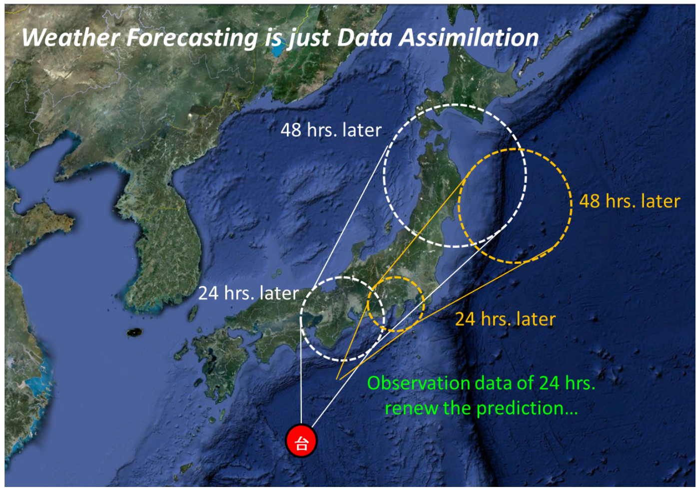
Bayesian Statistics, State Space Model

Higuchi et al. [2011]

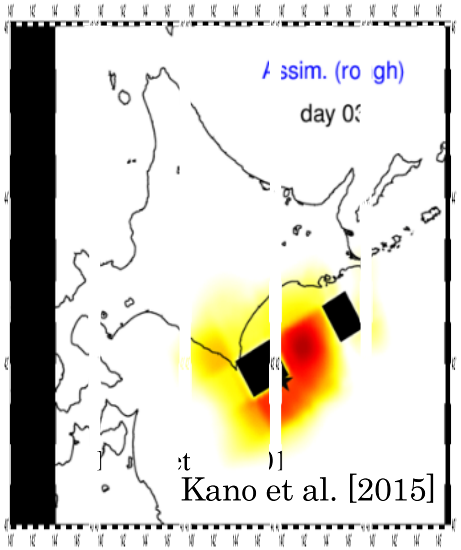


Observation Data

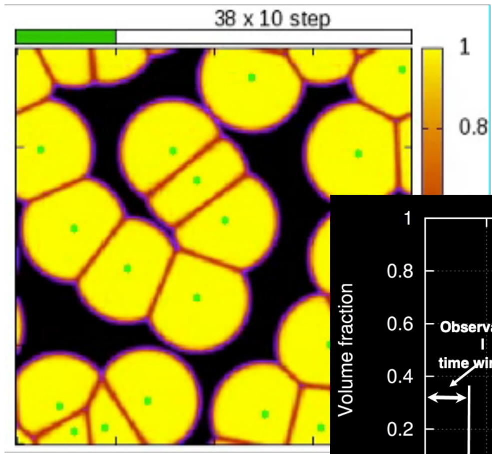
Weather Forecasting



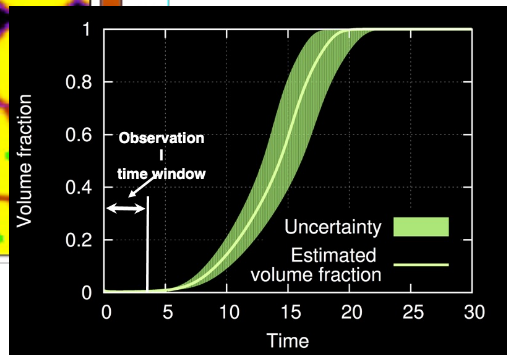
Seismology



Materials Science

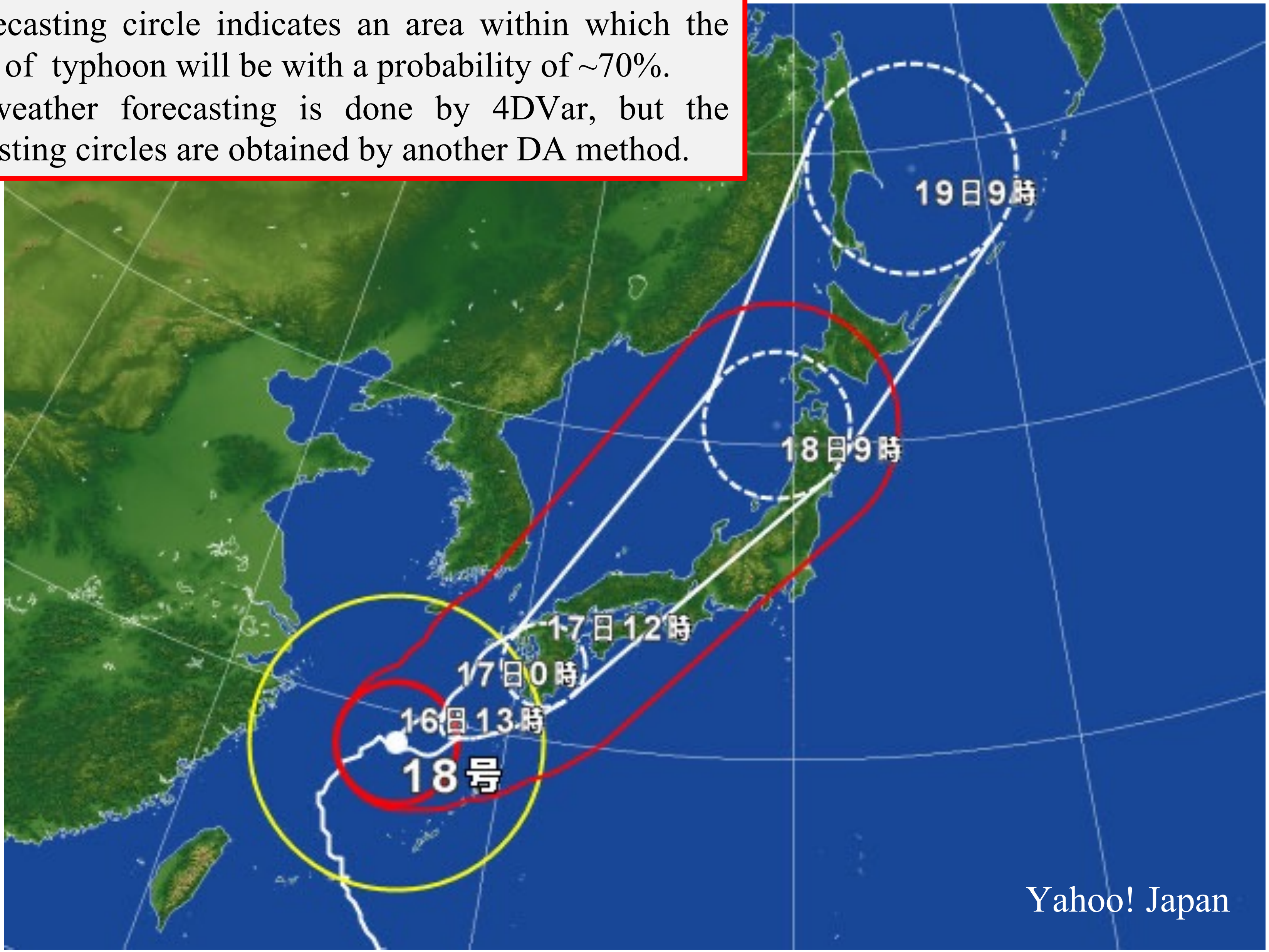


Ito et al. [2017]



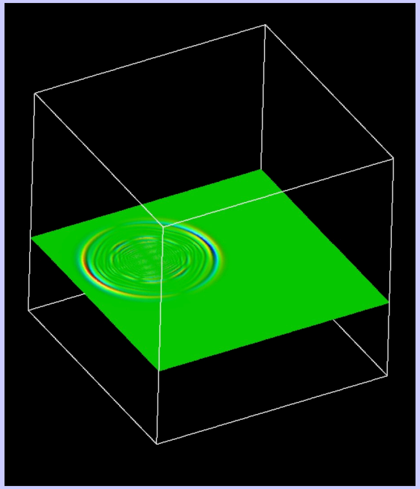
Weather Forecasting is Data Assimilation

- ✓ A forecasting circle indicates an area within which the center of typhoon will be with a probability of ~70%.
- ✓ The weather forecasting is done by 4DVar, but the forecasting circles are obtained by another DA method.



Four-Dimensional Variational Method (4DVar)

Massive simulations



Needs massive computational costs due to fine grids

Sequential Bayesian filters (e.g., ensemble Kalman filter)

Computational time $e^{O(N)}$ (N: degree of freedom)

Searches randomly parameter space ➡ Can estimate optimum and its uncertainty

4DVar (adjoint method)

Computational time $O(N)$

Estimates only optimum that maximizes posterior ➡ Cannot evaluate uncertainty of optimum

⬇

We develop a DA method that can estimate not only optimum but also its uncertainty even in the case of a system having large degrees of freedom

4DVar (adjoint method)

Simulation model

$$\frac{\partial \boldsymbol{\theta}_t}{\partial t} = \boldsymbol{F}(\boldsymbol{\theta}_t)$$

Observation model

$$D_t = h(\boldsymbol{\theta}_t) + \omega$$

Search $\boldsymbol{\Theta} = \boldsymbol{\theta}_0$ that best matches observation data

Bayes' theorem

$$\underbrace{p(\boldsymbol{\Theta}|\mathbf{D})}_{\text{posterior}} \propto \underbrace{p(\boldsymbol{\Theta})}_{\text{prior}}\underbrace{p(\mathbf{D}|\boldsymbol{\Theta})}_{\text{likelihood}}$$

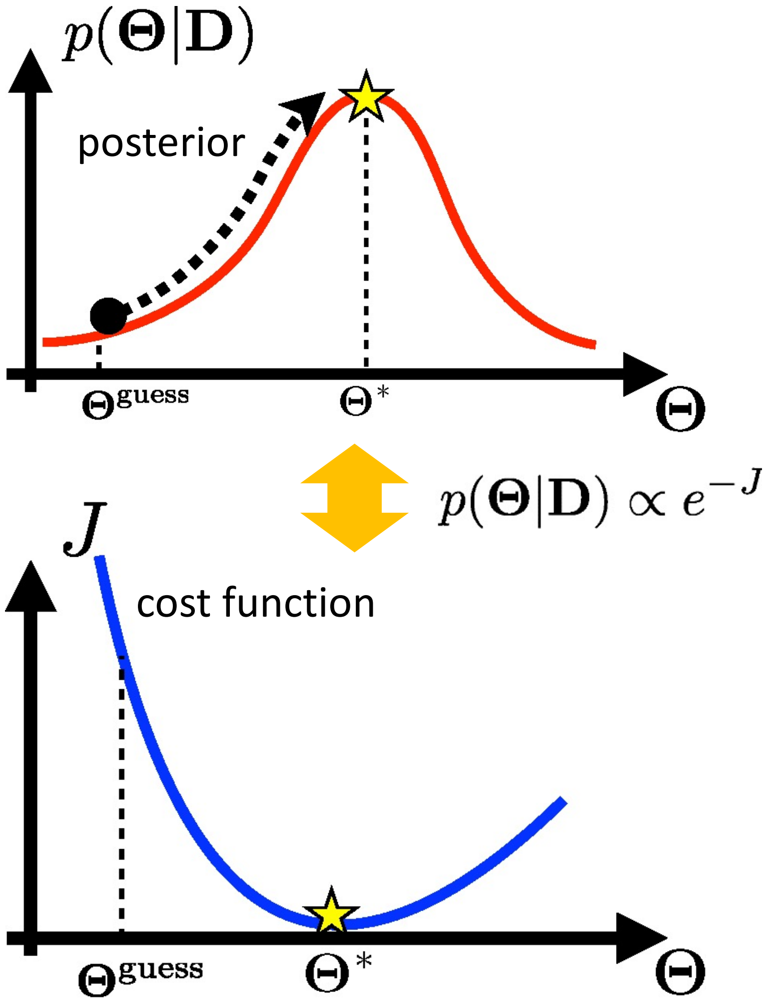
Cost function

$$J = -\log p(\boldsymbol{\Theta}) - \sum_{t \in \mathcal{T}} \log q\left(D_t - h(\boldsymbol{\theta}_t)\right)$$

Adjoint equation

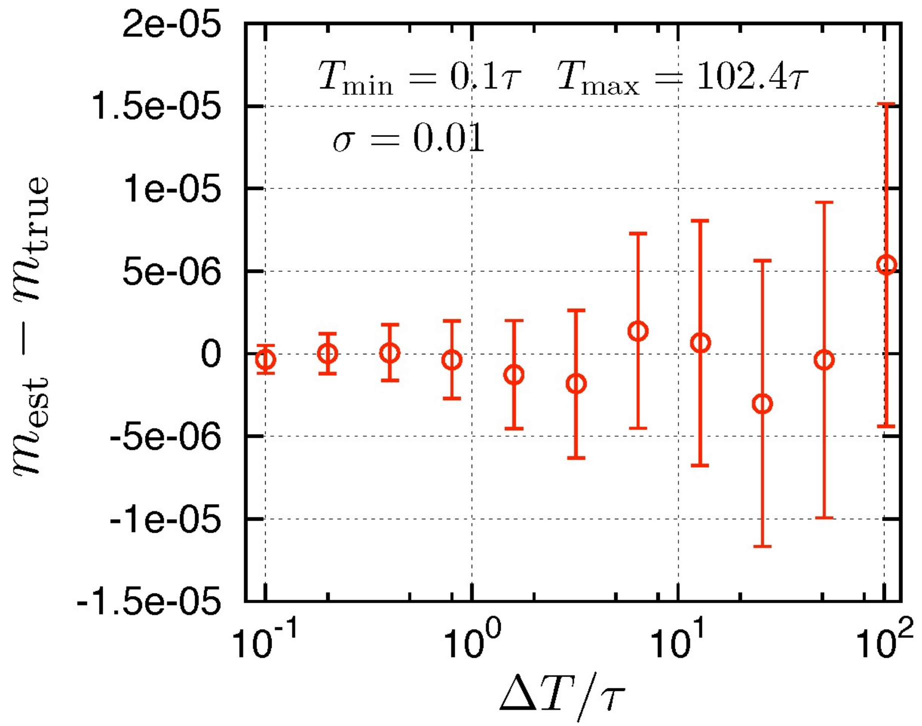
$$\frac{\partial \boldsymbol{\lambda}_t}{\partial t} + \left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{\theta}_t}\right)^{\top} \boldsymbol{\lambda}_t = \frac{\partial \mathcal{J}}{\partial \boldsymbol{\theta}_t}$$

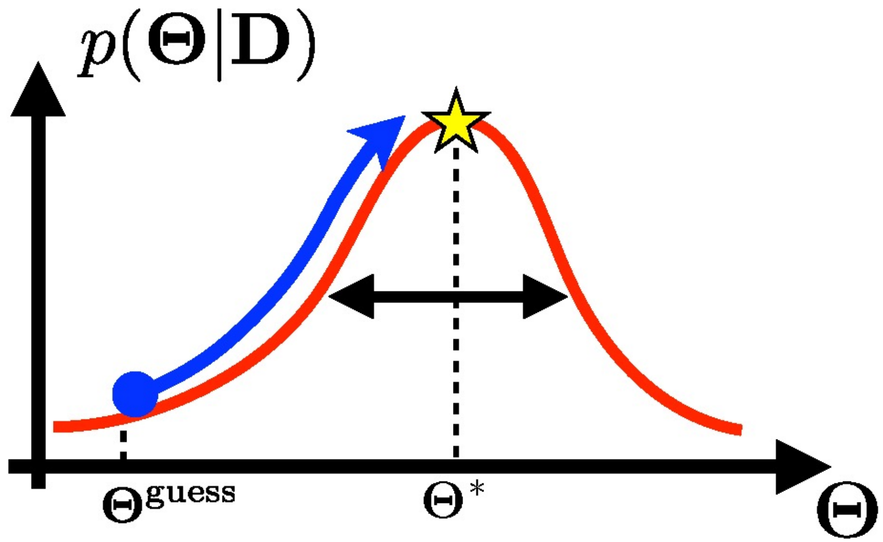
$$\boldsymbol{\lambda}_T = 0 \qquad \boldsymbol{\lambda}_0 = -\frac{\partial J}{\partial \boldsymbol{\Theta}}$$



Uncertainty Quantification (UQ)

We have established a methodology of uncertainty quantification using second-order adjoint method



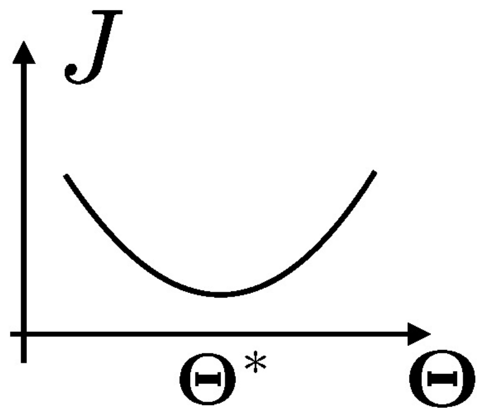


Enables us to estimate optimum and evaluate its uncertainty

Gives feedback to experimental design!

Ito, S., H. Nagao, A. Yamanaka, Y. Tsukada, T. Koyama, M. Kano, and J. Inoue, Data assimilation for massive autonomous systems based on a second-order adjoint method, Phys. Rev. E, 94, 043307, doi:10.1103/PhysRevE.94.043307, 2016.

Laplace Approximation of Posterior



Laplace approximation

Cost function can be approximated as a second-order polynomial in the neighborhood of the optimum Θ^*

$$J(\Theta) \sim J(\Theta^*) + \frac{1}{2}(\Theta - \Theta^*)^\top H (\Theta - \Theta^*)$$

$$p(\Theta|D) \sim N(\Theta^*, H^{-1})$$

H^{-1} : inverse of the Hessian matrix $H = \left.\frac{\partial^2 J}{\partial \Theta^2}\right|_{\Theta=\Theta^*}$

Direct computation of H^{-1} requires unpractical computation time $O(N^3)$.

But, what we need are only a limited number diagonal elements of H^{-1}

$$p(\Theta_k|D) = \int d\Theta_{-k}\, p(\Theta|D) = N\left(\Theta_k^*, (H^{-1})_{k,k}\right)$$

Second-Order Adjoint Method

We want to obtain only the k-th diagonal element in H^{-1} without explicitly computing H^{-1}

Solve $H\boldsymbol{r} = \boldsymbol{b}$ using an iterative method, where $\boldsymbol{b} = (0, \cdots, 0, 1, 0, \cdots, 0)^{\mathrm{T}}$

⬇

Needs a method to compute Hessian-vector product $H\boldsymbol{\alpha}$

Second-order adjoint method

Forward: Tangent linear model

$$\frac{\partial \boldsymbol{\xi}}{\partial t} = \left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{\theta}}\right) \cdot \boldsymbol{\xi} \qquad \boldsymbol{\xi}(0) = \underline{\boldsymbol{r}}$$

input

Backward: 2nd-order adjoint model

$$\frac{\partial \zeta}{\partial t} + \left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{\theta}}\right)^{\top} \cdot \zeta = \left(\frac{\partial^2 \boldsymbol{F}}{\partial \boldsymbol{\theta}^2} \cdot \boldsymbol{\xi}\right)^{\top} \cdot \boldsymbol{\lambda} - \frac{\partial^2 \mathcal{J}}{\partial \boldsymbol{\theta}^2} \cdot \boldsymbol{\xi}$$

$$\zeta(0) = \underline{\frac{\partial^2 J}{\partial \boldsymbol{\Theta}^2} \cdot \boldsymbol{r}} = H\boldsymbol{r}$$

output

$$\zeta(T) = 0$$

Procedure of UQ using Second-Order Adjoint Method

1. Estimate an optimum Θ^* minimizing J based on the adjoint method and a gradient method (we adopt here limited-memory BFGS method)

2. Evaluate the uncertainty of Θ^* based on the second-order adjoint method and a gradient method (we adopt here the conjugate residual method)

Remarks:

1. An array having size $O(N^2)$ is not needed.
2. Optimum estimation and UQ can be achieved with $O(K)$ computation (K: computation cost needed for a forward computation).

The proposed method is the only one that can estimate both optimum state and its uncertainty even with a system having large degrees of freedom

Kobayashi's Phase-Field Model

Phase-field model describing interface migration Kobayashi [1993]

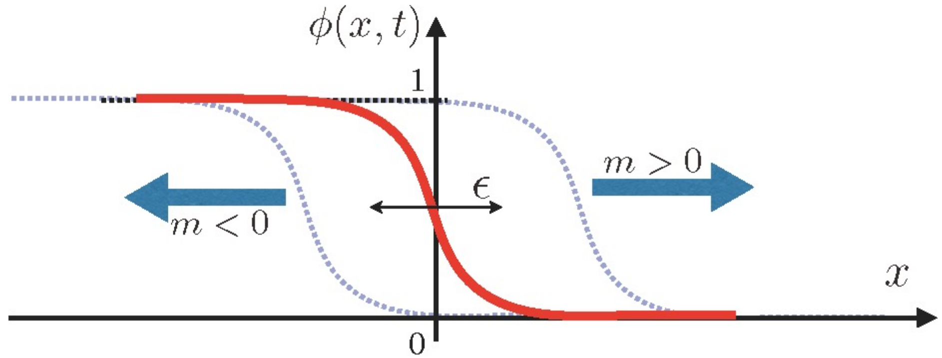
$$\tau \frac{\partial \phi}{\partial t} = \epsilon^2 \triangle \phi + \phi(1-\phi)\left(\phi - \frac{1}{2} + m\right) \qquad |m| < \frac{1}{2}$$

$\tau,\ \epsilon,\ m$ are assumed to be constants in time and space

1D

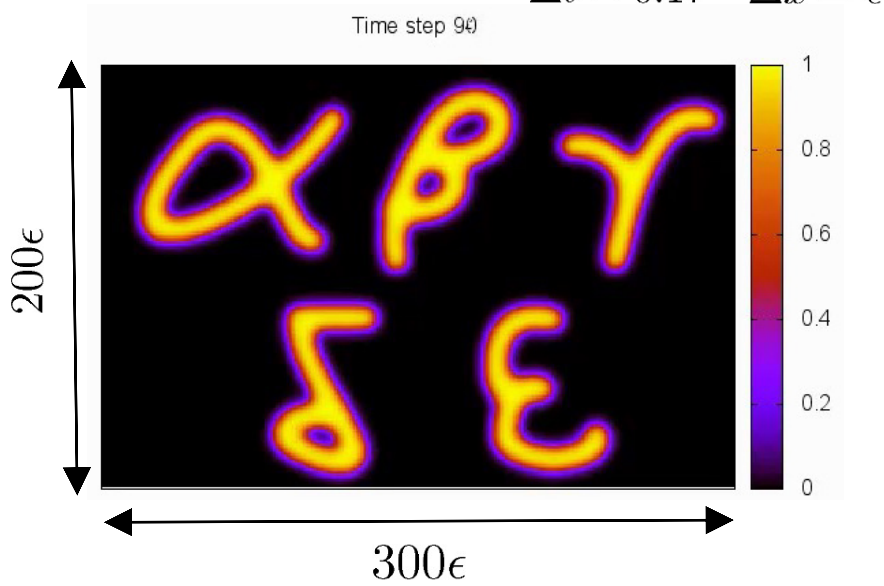
if $\begin{cases} m = \text{const.} \\ \phi(-\infty,t) = 1,\ \phi(\infty,t) = 0 \end{cases}$

$$\phi(x,t) = \frac{1}{2}\left[1 - \tanh\left(\frac{x}{2\sqrt{2}\varepsilon} - \frac{mt}{2\tau}\right)\right]$$



2D

$\Delta t = 0.1\tau \quad \Delta x = \epsilon$



First-Order (Conventional) Adjoint of Phase-Field Model

Forward

$$\tau \frac{\partial \phi_i}{\partial t} = \epsilon^2 \triangle_i \phi_i + \phi_i (1 - \phi_i) \left(\phi_i + m - \frac{1}{2} \right)$$

$$\frac{\partial m}{\partial t} = 0$$

Backward

$$-\tau \frac{\partial \tilde{\phi}_i}{\partial t} = \epsilon^2 \triangle_i \tilde{\phi}_i + \left\{ -3\phi_i^2 + (3 - 2m)\phi_i + m - \frac{1}{2} \right\} \tilde{\phi}_i - \frac{\partial \mathcal{J}}{\partial \phi_i} \sum_{k \in \mathcal{K}} \delta(t - t_k)$$

$$-\tau \frac{\partial \tilde{m}}{\partial t} = \sum_j \phi_j (1 - \phi_j) \tilde{\phi}_j - \frac{\partial \mathcal{J}}{\partial m} \sum_{k \in \mathcal{K}} \delta(t - t_k)$$

Second-Order Adjoint of Phase-Field Model

Forward

$$\tau \frac{\partial \hat{\phi}_i}{\partial t} = \epsilon^2 \triangle_i \hat{\phi}_i + \left\{ -3\phi_i^2 + (3-2m)\phi_i + m - \frac{1}{2} \right\} \hat{\phi}_i + \phi_i (1-\phi_i) \hat{m}$$

$$\frac{\partial \hat{m}}{\partial t} = 0$$

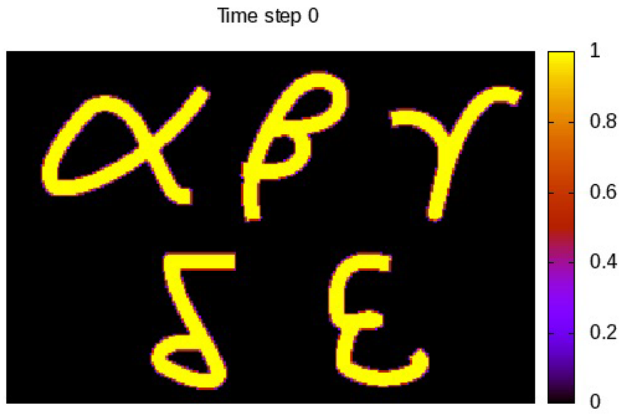
Backward

$$-\tau \frac{\partial \check{\phi}_i}{\partial t} = \epsilon^2 \triangle_i \check{\phi}_i + \left\{ -3\phi_i^2 + (3-2m)\phi_i + m - \frac{1}{2} \right\} \check{\phi}_i + (6\phi_i + 2m - 3)\, \tilde{\phi}_i \check{\phi}_i + (2\phi_i - 1)\, \tilde{\phi}_i \check{m} + \left[\sum_j \frac{\partial^2 \mathcal{J}}{\partial \phi_i \partial \phi_j} \hat{\phi}_j + \frac{\partial^2 \mathcal{J}}{\partial \phi_i \partial m} \hat{m} \right] \sum_{k \in \mathcal{K}} \delta(t - t_k)$$

$$-\tau \frac{\partial \check{m}}{\partial t} = \sum_j \left[\phi_j (1-\phi_j) \check{\phi}_j + (2\phi_j - 1)\, \tilde{\phi}_j \hat{\phi}_j \right] + \left[\sum_j \frac{\partial^2 \mathcal{J}}{\partial m \partial \phi_j} \hat{\phi}_j + \frac{\partial^2 \mathcal{J}}{\partial m^2} \hat{m} \right] \sum_{k \in \mathcal{K}} \delta(t - t_k)$$

Setup of Numerical Tests

Can the proposed method correctly reproduce the true parameter and true initial state from synthetic data that are generated by using the true parameter and initial state?



True initial state

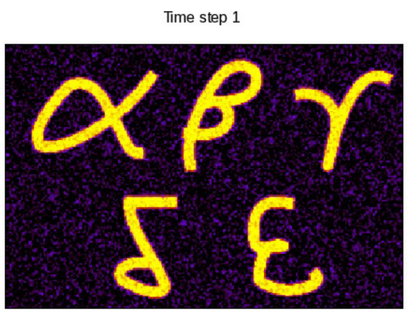
$m_{\mathrm{true}} = 0.1$

Synthetic observation data

True phase field obtained with time interval ΔT

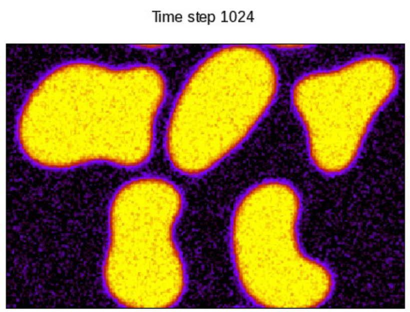
+

Gaussian noise $N\left(0, \sigma^2\right)$



$T_{\min} = 0.1\tau$

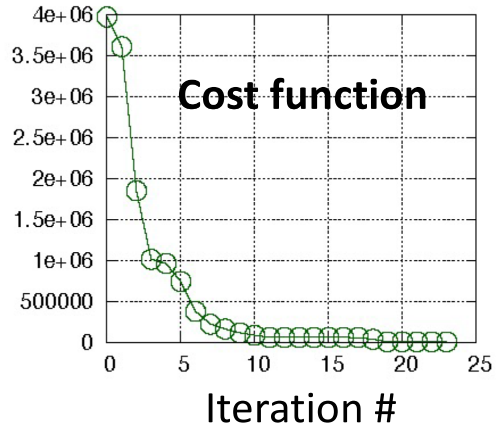
…

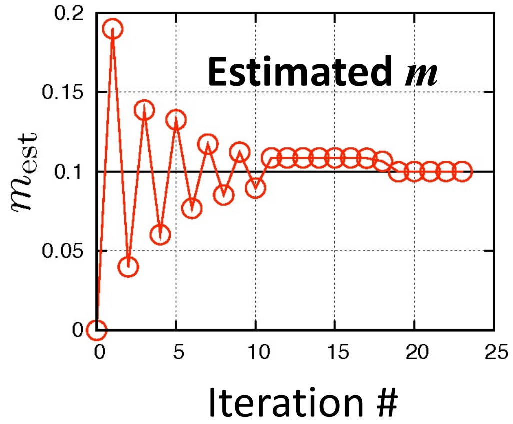


$T_{\max} = 102.4\tau$

$\sigma^2 = 0.01, \quad \Delta T = 0.1\tau$

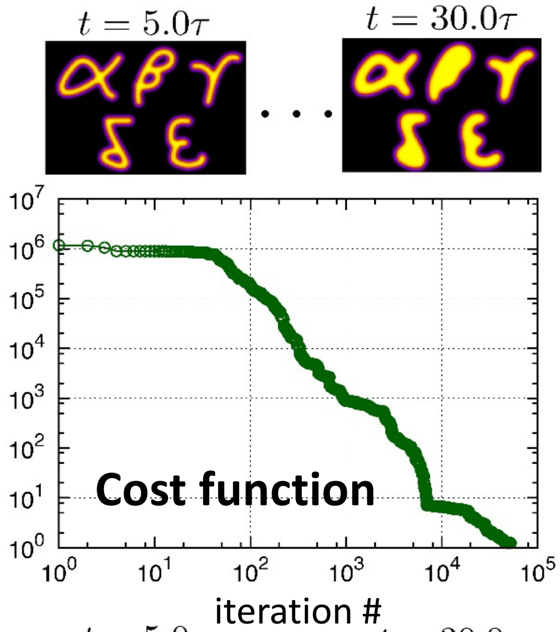
Example of twin experiment



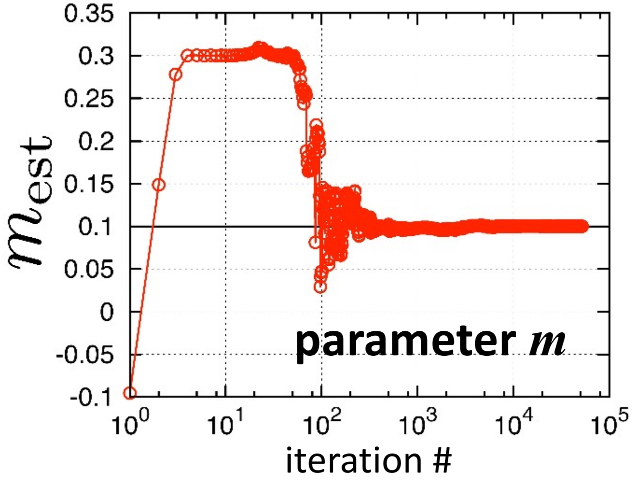


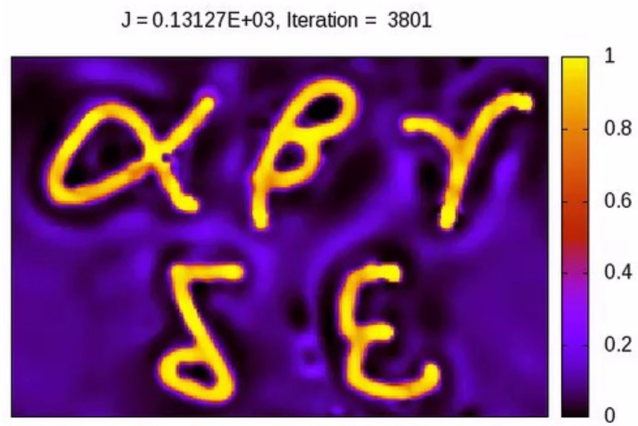
Results: Estimated Parameter & Initial State

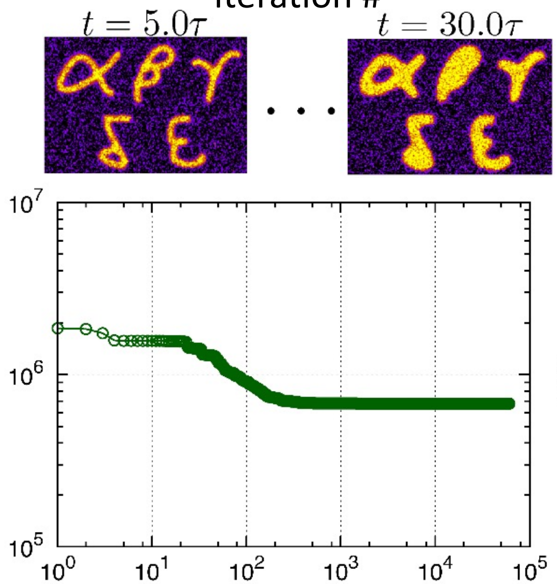
true



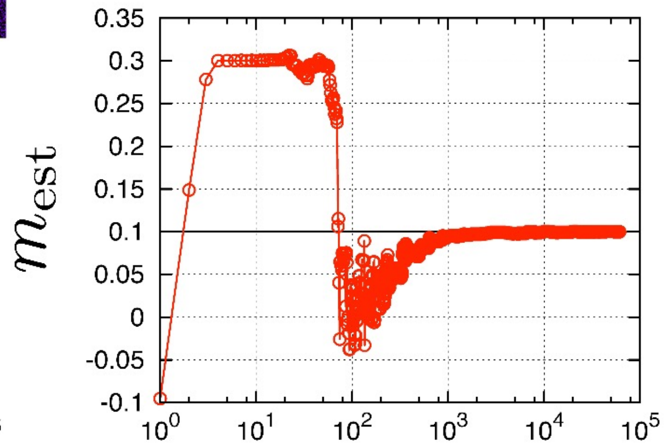
small noise: $\sigma = 1\times10^{-4}$

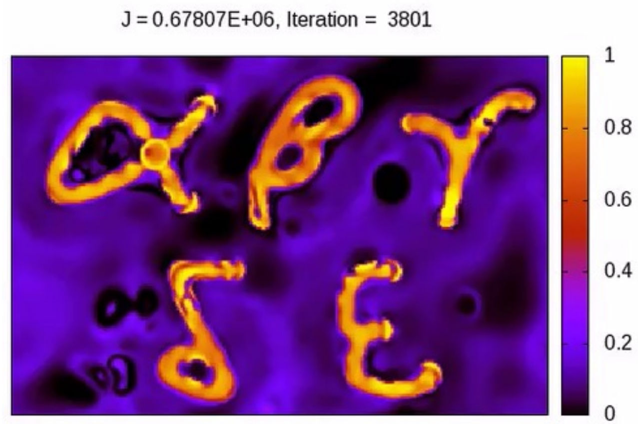




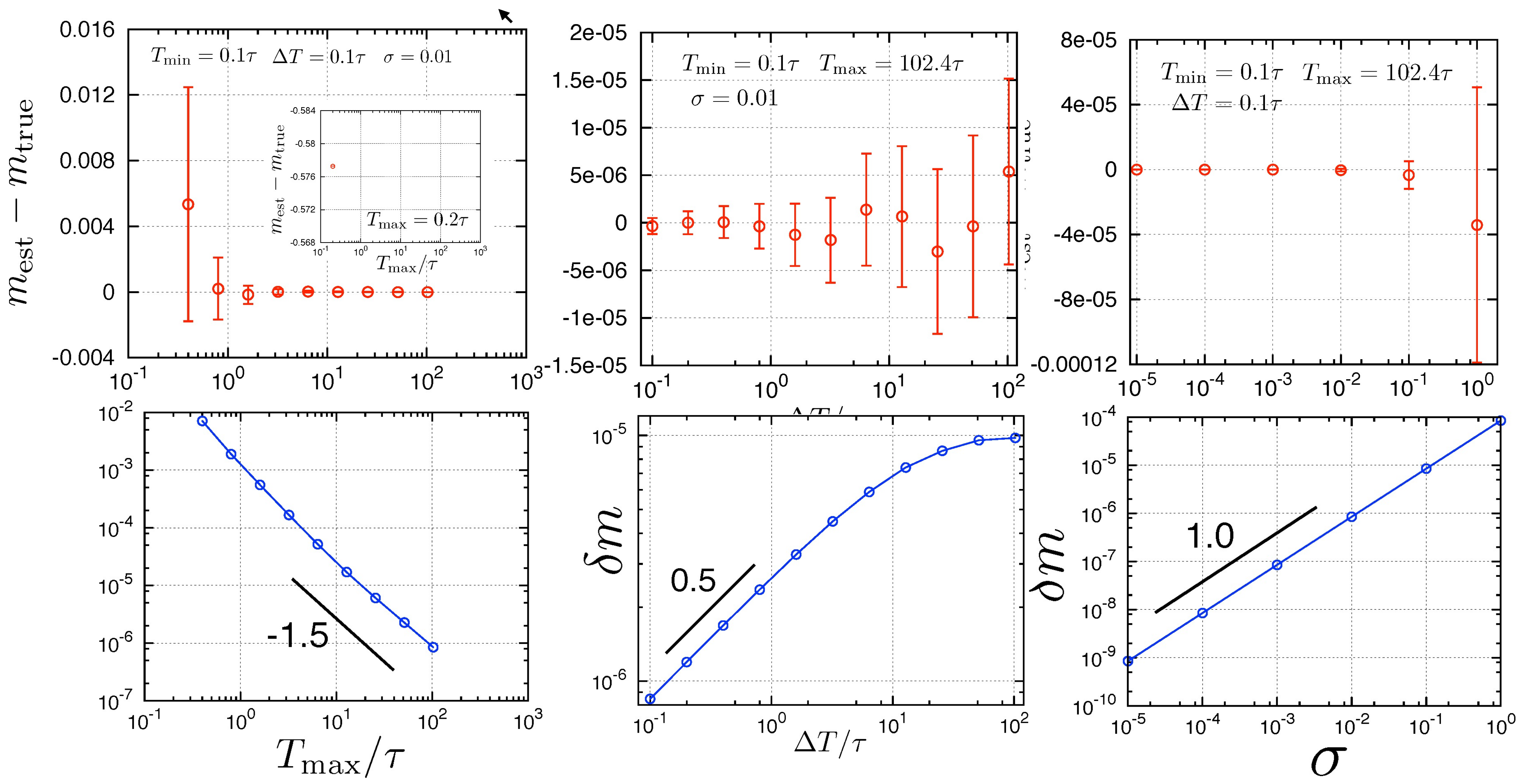


large noise: $\sigma = 0.3$





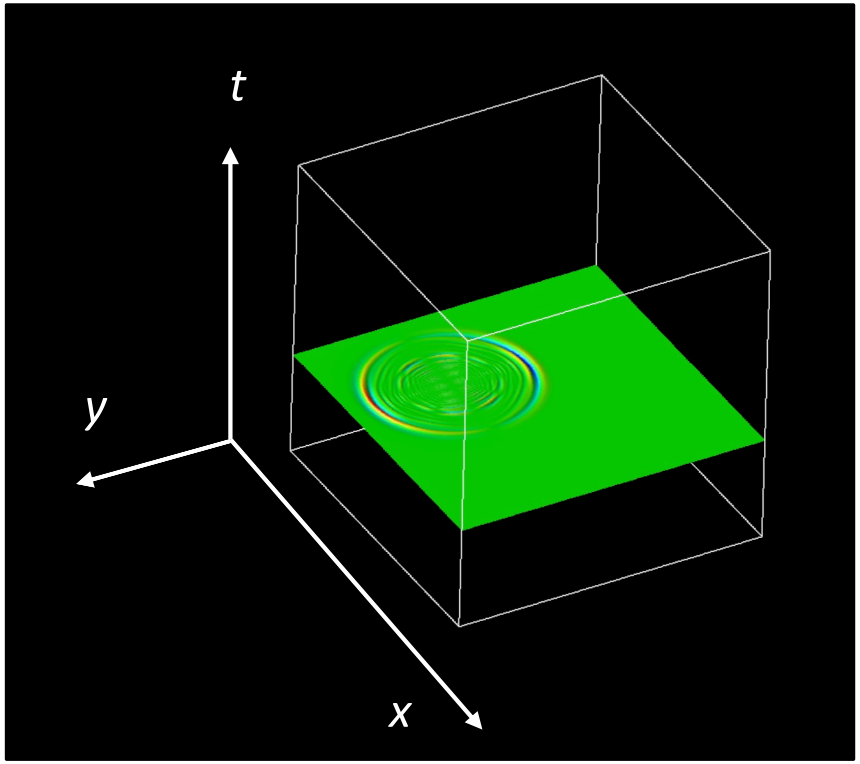
Parameter Estimation with Uncertainty Quantification



Estimation of parameter and its uncertainty depending on quality and quantity of data

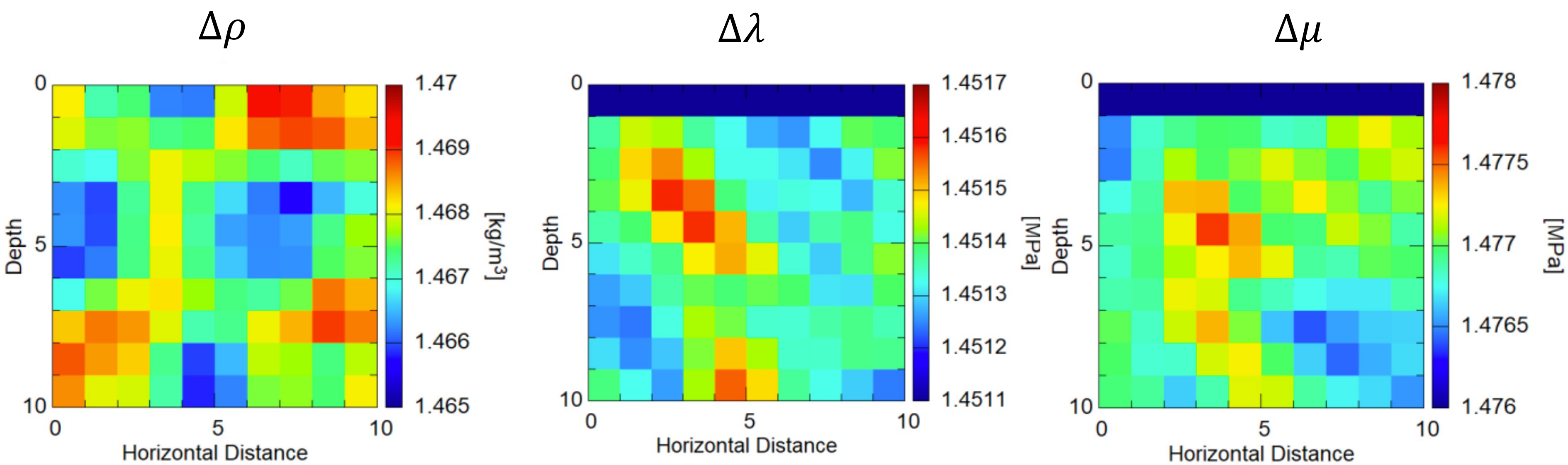
Feedback to experimental design

Application Trial to Seismic Wavefield Propagation



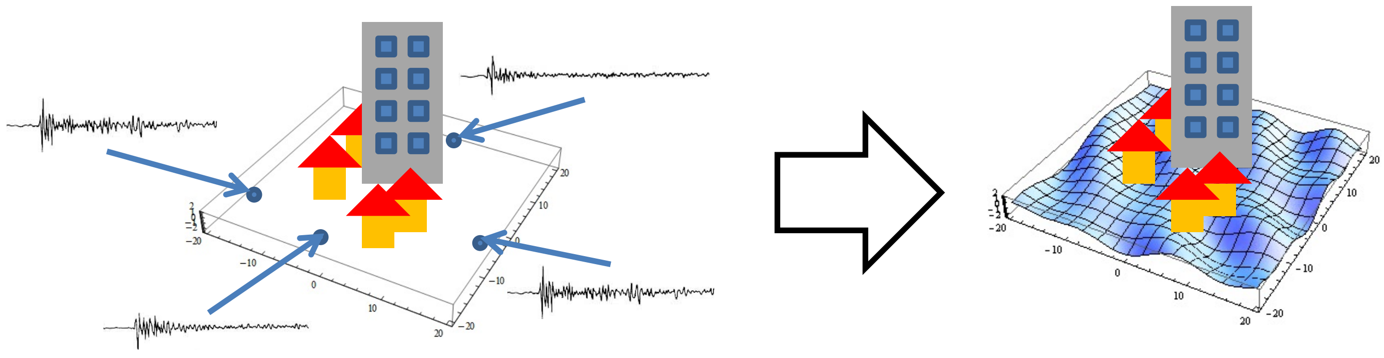
$$
\begin{cases}
\dot{u}_i = v_i \qquad (i, j, l = x, y) \\
\rho \dot{v}_i = \nabla_j \sigma_{ij} \\
\sigma_{ij} = \lambda \epsilon_{ll} \delta_{ij} + 2\mu \epsilon_{ij} \\
\epsilon_{ij} = \frac{1}{2} \left(\partial_i u_j + \partial_j u_i \right)
\end{cases}
$$

displacement u, velocity v
moduli of elasticity λ, μ
strain ϵ, stress σ, density ρ

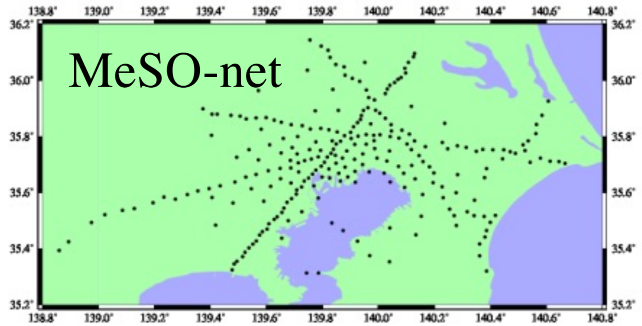


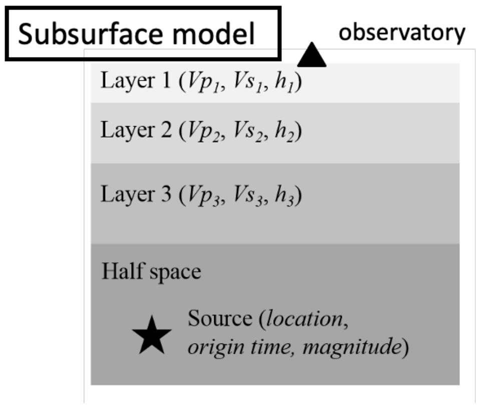
Seismic Wavefield & Underground Imaging Based on Learning

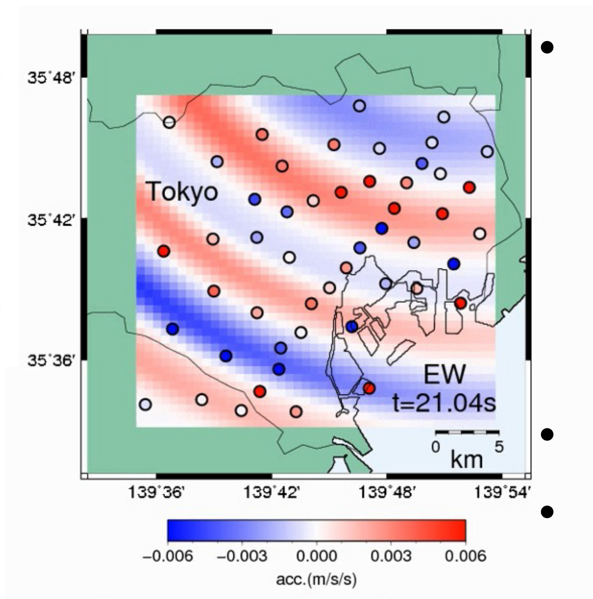
Rapid reconstruction of seismic wavefield with spatiotemporally-high resolution



Needs a combination of physical model and observation data to model wavefield <1Hz







Kano et al. (2017)

- Wavenumber integration method (Hisada, 1995) to simulate seismic waveforms with 1D underground structure (collaborating with Profs. Nakajima & Kawai for acceleration and parallelization)
- Replica exchange MCMC
- Construction of underground structure models in various areas in Japan based on transfer learning

Summary

- We have established a new 4DVar that enables us to estimate optimum state and parameters but also quantify their uncertainties based on the second-order adjoint method, which is applicable to a system having large degrees of freedom. Such uncertainty quantification could give feedback to designs of observations/experiments.
- We are dedicating to implement transfer learning techniques to apply seismic wavefield imaging method to efficiently estimate underground structures, accelerating and parallelizing the forward code that computes seismic waveforms.