Multi-level ensemble based data assimilation

Kristian Fossum, Trond Mannseth, Andreas Stordal

EnKF Workshop 2017

13/06/2017

Introduction

Coarse scale and multilevel ensemble based data assimilation

In reservoir simulation models

$$\text{state} = g(m) = (Saturation, Pressure)$$

obtained at a high computational cost

Introduction

Coarse scale and multilevel ensemble based data assimilation

In reservoir simulation models

$$\text{state} = g(m) = (Saturation, Pressure)$$

obtained at a high computational cost

Ensemble based data assimilation approximates

$$p(m|d) = \frac{p(d|m)p(m)}{\int (p(d|m)p(m))}$$

by Monte-Carlo estimation

Quality of estimation relies on the ensemble size

Introduction

Coarse scale and multilevel ensemble based data assimilation

In reservoir simulation models

$$\text{state} = g(m) = (Saturation, Pressure)$$

obtained at a high computational cost

Ensemble based data assimilation approximates

$$p(m|d) = \frac{p(d|m)p(m)}{\int (p(d|m)p(m))}$$

by Monte-Carlo estimation

Quality of estimation relies on the ensemble size

Finite computational resources $\rightarrow$ ensemble size not optimal

Introduction

Coarse scale and multilevel ensemble based data assimilation

The following assumptions are made throughout:

- Proxy model with adjustable accuracy:
  $d_l = f(\text{state}) = f(g_l(m))$ with $l = 0, 1, \ldots, L$

Introduction

Coarse scale and multilevel ensemble based data assimilation

The following assumptions are made throughout:

- Proxy model with adjustable accuracy:
  $d_l = f(\text{state}) = f(g_l(m))$ with $l = 0, 1, \ldots, L$
- For $l = L$ the proxy model equals the original model

Introduction

Coarse scale and multilevel ensemble based data assimilation

The following assumptions are made throughout:

- Proxy model with adjustable accuracy:
  $d_l = f(\text{state}) = f(g_l(m))$ with $l = 0, 1, \ldots, L$
- For $l = L$ the proxy model equals the original model
- Numerical accuracy inversely proportional to computational cost

Introduction

Coarse scale and multilevel ensemble based data assimilation

The following assumptions are made throughout:

- Proxy model with adjustable accuracy:
  $d_l = f(\text{state}) = f(g_l(m))$ with $l = 0, 1, \ldots, L$
- For $l = L$ the proxy model equals the original model
- Numerical accuracy inversely proportional to computational cost
- For all accuracy levels: $d \in \mathbb{R}^d$ and $m \in \mathbb{R}^m$

Introduction

Coarse scale and multilevel ensemble based data assimilation

The following assumptions are made throughout:

- Proxy model with adjustable accuracy:
  $d_l = f(\text{state}) = f(g_l(m))$ with $l = 0, 1, \ldots, L$
- For $l = L$ the proxy model equals the original model
- Numerical accuracy inversely proportional to computational cost
- For all accuracy levels: $d \in \mathbb{R}^d$ and $m \in \mathbb{R}^m$

Earlier results: *Coarse scale ensemble based data assimilation*

A better balance between numerical and statistical accuracy results in improved DA results

Introduction

Coarse scale and multilevel ensemble based data assimilation

The following assumptions are made throughout:

- Proxy model with adjustable accuracy:
  $d_l = f(\text{state}) = f(g_l(m))$ with $l = 0, 1, \ldots, L$
- For $l = L$ the proxy model equals the original model
- Numerical accuracy inversely proportional to computational cost
- For all accuracy levels: $d \in \mathbb{R}^d$ and $m \in \mathbb{R}^m$

Earlier results: *Coarse scale ensemble based data assimilation*

A better balance between numerical and statistical accuracy results in improved DA results

However:
High proxy error $\rightarrow$ Poor estimation of $p(m|d)$

Introduction
Multilevel methods

Problem with coarse scale DA: select accurate proxy model

The multilevel approach removes this problem

Alternative approach $\rightarrow$

1. MLEnKF – unbiased
2. Bayesian model average – biased

Introduction
Multilevel methods

Problem with coarse scale DA: select accurate proxy model

The multilevel approach removes this problem

Alternative approach $\rightarrow$

1. MLEnKF – unbiased
2. Bayesian model average – biased

Firstly: Investigate Multilevel Monte Carlo (MLMC)

Multilevel Monte Carlo

Introduction

MLMC was introduced as an efficient alternative to standard MC estimation

Multilevel Monte Carlo

Introduction

MLMC was introduced as an efficient alternative to standard MC estimation

Given a sequence $P_0, \ldots, P_{L-1}$ which approximates P_L

- increasing accuracy

Multilevel Monte Carlo

Introduction

MLMC was introduced as an efficient alternative to standard MC estimation

Given a sequence $P_0, \ldots, P_{L-1}$ which approximates P_L

- increasing accuracy
- increasing cost

Multilevel Monte Carlo

Introduction

MLMC was introduced as an efficient alternative to standard MC estimation

Given a sequence $P_0, \ldots, P_{L-1}$ which approximates P_L

- increasing accuracy
- increasing cost

$$\mathbb{E}[P_L] = \mathbb{E}[P_0] + \sum_{l=1}^{L} \mathbb{E}[P_l - P_{l-1}]$$

which can be estimated as

$$(N_e)_0^{-1} \sum_{n=1}^{(N_e)_0} P_0^n + \sum_{l=1}^{L} (N_e)_l^{-1} \sum_{n=1}^{(N_e)_l} (P_l^n - P_{l-1}^n)$$

Multilevel Monte Carlo

Introduction

MLMC was introduced as an efficient alternative to standard MC estimation

Given a sequence $P_0, \ldots, P_{L-1}$ which approximates P_L

- increasing accuracy
- increasing cost

$$\mathbb{E}[P_L] = \mathbb{E}[P_0] + \sum_{l=1}^{L} \mathbb{E}[P_l - P_{l-1}]$$

which can be estimated as

$$(N_e)_0^{-1} \sum_{n=1}^{(N_e)_0} P_0^n + \sum_{l=1}^{L} (N_e)_l^{-1} \sum_{n=1}^{(N_e)_l} (P_l^n - P_{l-1}^n)$$

With unlimited computational resources $\rightarrow$ the ML method is more efficient than the standard MC method

Multilevel Monte Carlo

Application with restricted computational resources

MSE of MLMC for Euler discretisation of simple SDE [1]

$$c_2 h_L^2 + \sum_{l=0}^{L} c_1 (N_e)_l^{-1} h_l$$

where h_l is grid size at level l, c_1 and c_2 weights bias and variance

[1] Giles, M. B. Multi-level Monte Carlo path simulation. Oper. Res. 56.

Multilevel Monte Carlo

Application with restricted computational resources

MSE of MLMC for Euler discretisation of simple SDE [1]

$$c_2 h_L^2 + \sum_{l=0}^{L} c_1 (N_e)_l^{-1} h_l$$

where h_l is grid size at level l, c_1 and c_2 weights bias and variance
Assume that similar relationship is valid for subsurface flow

[1] Giles, M. B. Multi-level Monte Carlo path simulation. Oper. Res. 56.

Multilevel Monte Carlo

Application with restricted computational resources

MSE of MLMC for Euler discretisation of simple SDE [1]

$$c_2 h_L^2 + \sum_{l=0}^{L} c_1 (N_e)_l^{-1} h_l$$

where h_l is grid size at level l, c_1 and c_2 weights bias and variance
Assume that similar relationship is valid for subsurface flow

MSE value depends on
- True values of c_1 & c_2
- Total computational resources

[1] Giles, M. B. Multi-level Monte Carlo path simulation. Oper. Res. 56.

Multilevel Monte Carlo

Application with restricted computational resources

MSE of MLMC for Euler discretisation of simple SDE [1]

$$c_2 h_L^2 + \sum_{l=0}^{L} c_1 (N_e)_l^{-1} h_l$$

where h_l is grid size at level l, c_1 and c_2 weights bias and variance
Assume that similar relationship is valid for subsurface flow

MSE value depends on
- True values of c_1 & c_2
- Total computational resources

MLMC fails when
- There are restricted computational resources, and $c_1 \gg c_2$

[1] Giles, M. B. Multi-level Monte Carlo path simulation. Oper. Res. 56.

Multilevel Monte Carlo

Application with restricted computational resources

MSE of MLMC for Euler discretisation of simple SDE [1]

$$c_2 h_L^2 + \sum_{l=0}^{L} c_1 (N_e)_l^{-1} h_l$$

where h_l is grid size at level l, c_1 and c_2 weights bias and variance
Assume that similar relationship is valid for subsurface flow

MSE value depends on
- True values of c_1 & c_2
- Total computational resources

MLMC fails when
- There are restricted computational resources, and $c_1 \gg c_2$

In general c_1 and c_2 are unknown

---

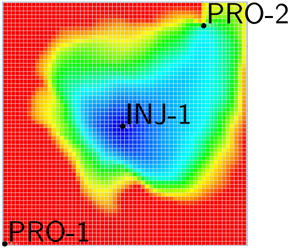
[1] Giles, M. B. Multi-level Monte Carlo path simulation. Oper. Res. 56.

Two phase flow test

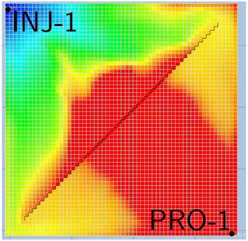
Investigate two test cases

- 60×60 grid-cells
- 80 assimilation time steps
- Proxy via uniform upscaling
- Model 1: No fault
- Model 2: Dominant impermeable fault

- Model 1 & 2: computational resources = 10 full runs



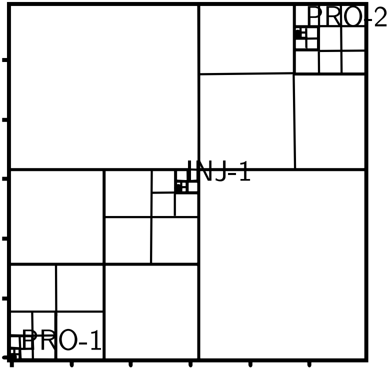
Model 1



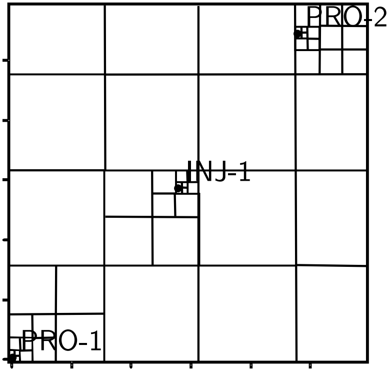
Model 2

Uniform upscaling

Grid model 1



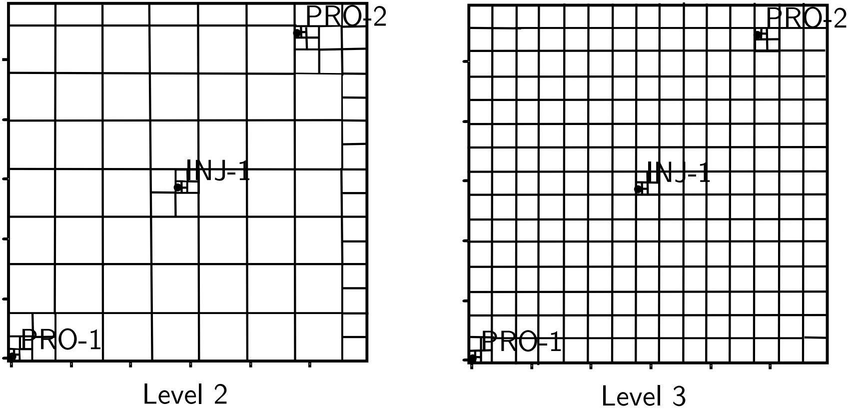
Level 0



Level 1

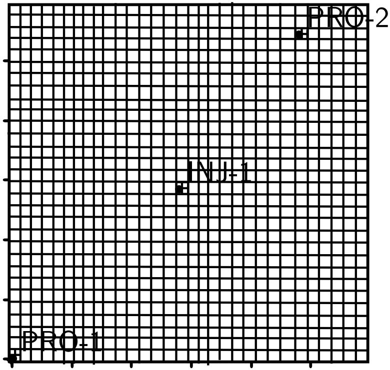
Uniform upscaling

Grid model 1



Uniform upscaling

Grid model 1



Level 4

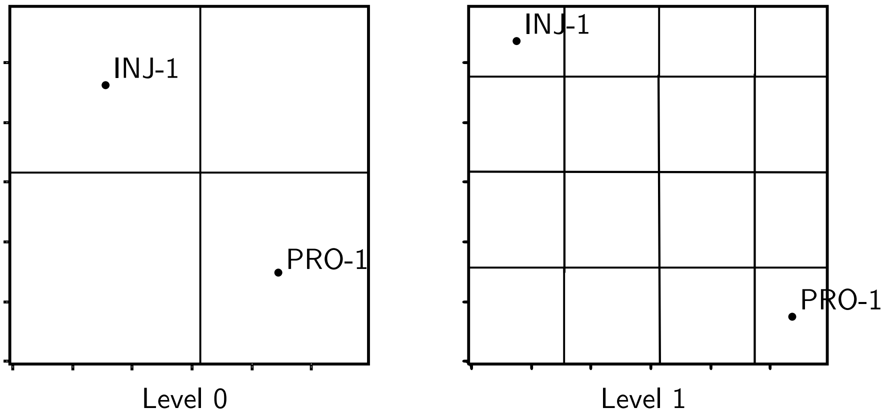
Grid-size for the various levels

Level 0	47
Level 1	53
Level 2	98
Level 3	243
Level 4	909

Original grid: 3600 grid-cells

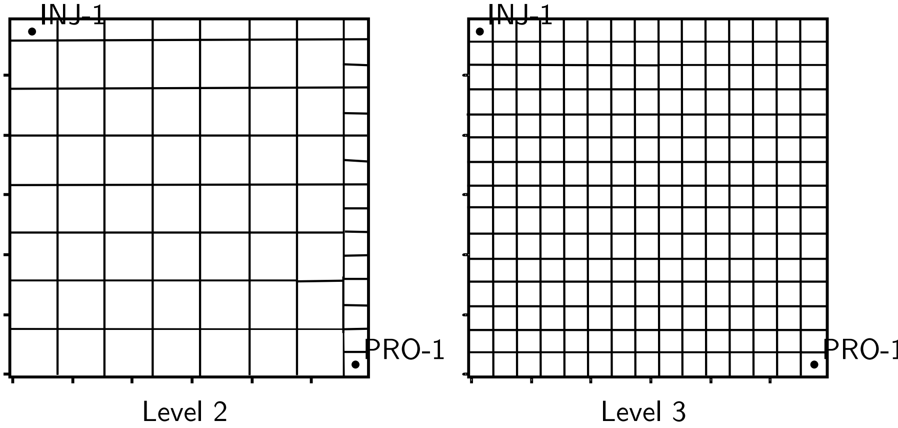
Uniform upscaling

Grid model 2



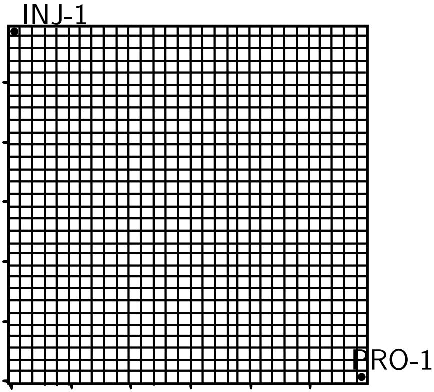
Uniform upscaling

Grid model 2



Uniform upscaling

Grid model 2



Level 4

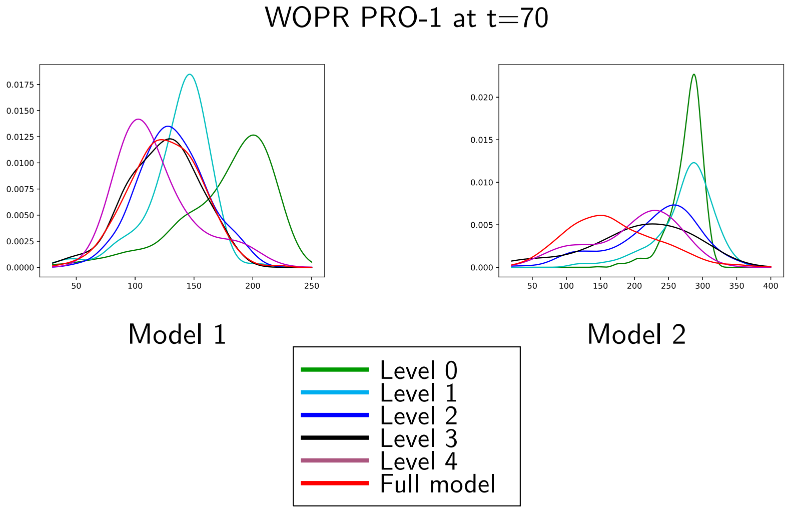
Grid-size for the various levels

Level 0	4
Level 1	16
Level 2	71
Level 3	225
Level 4	900

Original grid: 3600 grid-cells

Uniform upscaling

Kernel density estimate of simulator output



Bias and variance of C_{mg}

Estimation by bootstrapping

For data assimilation applications estimate of C_{mg} is important

Bias and variance of C_{mg}

Estimation by bootstrapping

For data assimilation applications estimate of C_{mg} is important

Evaluate the bias and variance for $(C_{mg})_l$ for $l = 0, 1, 2, 3, 4$ by bootstrapping

Bias and variance of C_{mg}

Estimation by bootstrapping

For data assimilation applications estimate of C_{mg} is important

Evaluate the bias and variance for $(C_{mg})_l$ for $l = 0, 1, 2, 3, 4$ by bootstrapping

Define $(N_e)_l = (N_e^{\text{full}})_l \times \frac{N_g^{\text{full}}}{(N_g)_l}$
where $\sum_{l=0}^{4} (N_e^{\text{full}})_l = 10$

Bias and variance of C_{mg}

Estimation by bootstrapping

For data assimilation applications estimate of C_{mg} is important

Evaluate the bias and variance for $(C_{mg})_l$ for $l = 0, 1, 2, 3, 4$ by bootstrapping

Define $(N_e)_l = (N_e^{\text{full}})_l \times \frac{N_g^{\text{full}}}{(N_g)_l}$
where $\sum_{l=0}^{4} (N_e^{\text{full}})_l = 10$

Ensemble size for the various levels

	Level 0	Level 1	Level 2	Level 3	Level 4
Model 1	454	209	123	43	9
Model 2	900	500	102	32	12

Bias and variance of C_{mg}

Estimation by bootstrapping

For data assimilation applications estimate of C_{mg} is important

Evaluate the bias and variance for $(C_{mg})_l$ for $l = 0, 1, 2, 3, 4$ by bootstrapping

Define $(N_e)_l = (N_e^{\text{full}})_l \times \frac{N_g^{\text{full}}}{(N_g)_l}$
where $\sum_{l=0}^{4} (N_e^{\text{full}})_l = 10$

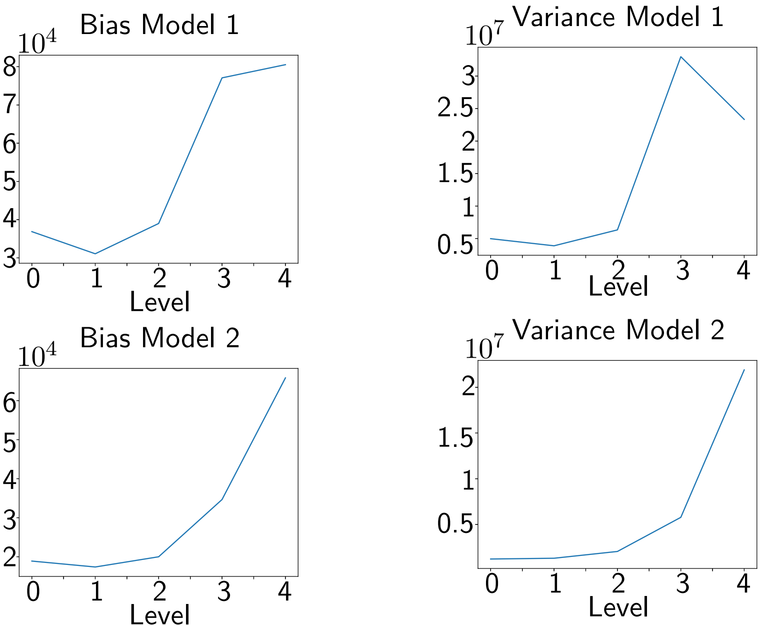
Ensemble size for the various levels

	Level 0	Level 1	Level 2	Level 3	Level 4
Model 1	454	209	123	43	9
Model 2	900	500	102	32	12

Calculate bias and variance of $(C_{mg})_l$ by 2000 replications

Element wise bias and variance of C_{mg}

Frobenius norm



Bias and variance

Estimation by bootstrapping

Keeping the total computational resources fixed on each level we observe that for both models

– Variance is the dominant factor
– Bias *increases* with accuracy due to MC error

Bias and variance

Estimation by bootstrapping

Keeping the total computational resources fixed on each level we observe that for both models

– Variance is the dominant factor
– Bias *increases* with accuracy due to MC error

We cannot calculate bootstrapped estimates of bias and variance for large cases
– Realistic cases have lower value of $^{\text{computational resources}}/_{\text{grid-cells}}$

Bias and variance

Estimation by bootstrapping

Keeping the total computational resources fixed on each level we observe that for both models

– Variance is the dominant factor
– Bias *increases* with accuracy due to MC error

We cannot calculate bootstrapped estimates of bias and variance for large cases
– Realistic cases have lower value of $^{\text{computational resources}}/_{\text{grid-cells}}$

Analysis of Model 1 & 2 with limited computational resources
– Resources best spent to reduce variance

Bias and variance

Estimation by bootstrapping

Keeping the total computational resources fixed on each level we observe that for both models

– Variance is the dominant factor
– Bias *increases* with accuracy due to MC error

We cannot calculate bootstrapped estimates of bias and variance for large cases
– Realistic cases have lower value of $^{\text{computational resources}}/_{\text{grid-cells}}$

Analysis of Model 1 & 2 with limited computational resources
– Resources best spent to reduce variance

$\rightarrow$ Evaluate two different multilevel algorithms

Multilevel methods

Multilevel EnKF

MLMC can be extended to the EnKF framework

$$C_{ML} = C_0 + \sum_{l=1}^{L} (C_l - C_{l-1})$$

Multilevel methods

Multilevel EnKF

MLMC can be extended to the EnKF framework

$$C_{ML} = C_0 + \sum_{l=1}^{L} (C_l - C_{l-1})$$

the MLEnKF analysis is then

$$m^a = m^f + C_{ML} H^T (H C_{ML} H^T + C_d)^{-1} (d - g(m))$$

Multilevel methods

Multilevel EnKF

MLMC can be extended to the EnKF framework

$$C_{ML} = C_0 + \sum_{l=1}^{L} (C_l - C_{l-1})$$

the MLEnKF analysis is then

$$m^a = m^f + C_{ML}H^T(HC_{ML}H^T + C_d)^{-1}(d - g(m))$$

– Converges to the KF solution

Multilevel methods

Bayesian model average

Let the forecast density be defined by Bayesian model averaging

$$p(Y|d) \propto p(d|Y)p(Y) = p(d|Y)\sum_{l=0}^{L} p(Y|M_l)p(M_l)$$

Each model M_l represents an accuracy level of the proxy

Assume that all densities are Gaussian

Multilevel methods

Bayesian model average

Let the forecast density be defined by Bayesian model averaging

$$p(Y|d) \propto p(d|Y)p(Y) = p(d|Y)\sum_{l=0}^{L} p(Y|M_l)p(M_l)$$

Each model M_l represents an accuracy level of the proxy

Assume that all densities are Gaussian

Bayesian model averaging utilize all proxy models

– Bias-variance tradeoff adjusted through the weights, $p(M_l)$

Multilevel methods

Bayesian model average

New total empirical forecast covariance given by law of total covariance

- $C_{tot} = \sum_{l=0}^{L} p(M_l) C_l + \sum_{l=0}^{L} p(M_l)(\mu_l - \overline{\mu})(\mu_l - \overline{\mu})^T$

How to select $p(M_l)$?

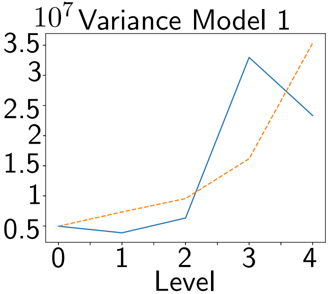
Multilevel methods

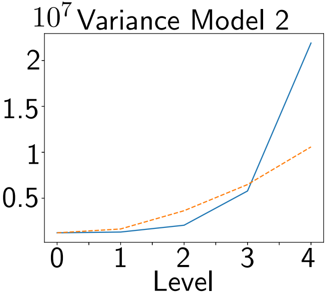
Bayesian model average

New total empirical forecast covariance given by law of total covariance

- $C_{tot} = \sum_{l=0}^{L} p(M_l) C_l + \sum_{l=0}^{L} p(M_l)(\mu_l - \overline{\mu})(\mu_l - \overline{\mu})^T$

How to select $p(M_l)$?





Dashed line: $\frac{c}{\sqrt{(N_e)_l}}$

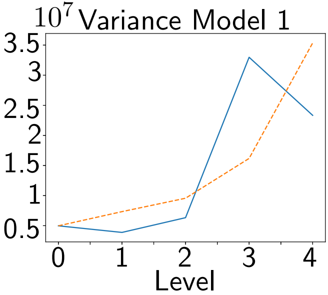
Multilevel methods

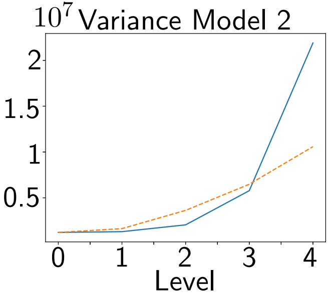
Bayesian model average

New total empirical forecast covariance given by law of total covariance

- $C_{tot} = \sum_{l=0}^{L} p(M_l) C_l + \sum_{l=0}^{L} p(M_l)(\mu_l - \overline{\mu})(\mu_l - \overline{\mu})^T$

How to select $p(M_l)$?





Dashed line: $\frac{c}{\sqrt{(N_e)_l}}$

Select the weights as $p(M_l) \propto \sqrt{(N_e)_l}$
– constraints $\sum_{l=0}^{L} p(M_l) = 1$

Multilevel methods

Bayesian model average – analysis step

Each level has a different forecast bias $\rightarrow$ update each level uniquely

Multilevel methods

Bayesian model average – analysis step

Each level has a different forecast bias $\rightarrow$ update each level uniquely

$$m_l^a = m_l^f + C_{mg}^{tot}(C_{gg}^{tot} + \alpha_l C_d)^{-1}(d_l - g_l(m_l^f))$$

Multilevel methods

Bayesian model average – analysis step

Each level has a different forecast bias $\rightarrow$ update each level uniquely

$$m_l^a = m_l^f + C_{mg}^{tot}(C_{gg}^{tot} + \alpha_l C_d)^{-1}(d_l - g_l(m_l^f))$$

α_l is selected adaptively

$$\alpha_l = \frac{\|d_{true} - \overline{g_l(m_l^f)}\|}{\|d_{true} - \overline{g_L(m_L^f)}\|}$$

Multilevel methods

Bayesian model average – analysis step

Each level has a different forecast bias $\rightarrow$ update each level uniquely

$$m_l^a = m_l^f + C_{mg}^{tot}(C_{gg}^{tot} + \alpha_l C_d)^{-1}(d_l - g_l(m_l^f))$$

α_l is selected adaptively

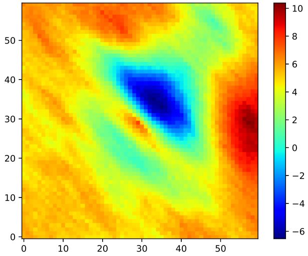
$$\alpha_l = \frac{\|d_{true} - \overline{g_l(m_l^f)}\|}{\|d_{true} - \overline{g_L(m_L^f)}\|}$$

$\rightarrow$ if all models have same accuracy $\alpha_l = 1 \quad \forall l$

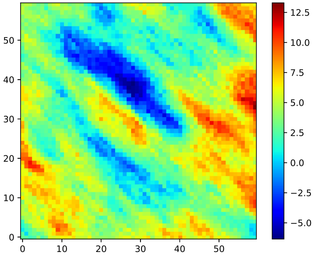
In the Gaussian case with linear dynamic models:
– Equally accurate models $\rightarrow$ Converge to KF

BMA

Mean example 1



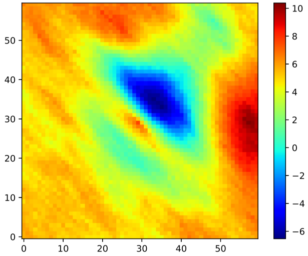
ES with large ensemble



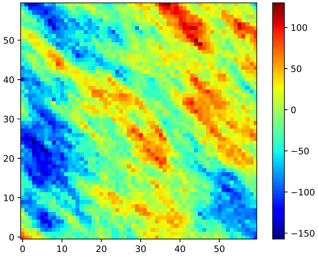
BMA

MLEnKF

Mean example 1



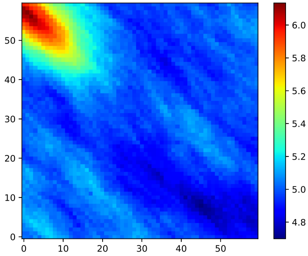
ES with large ensemble



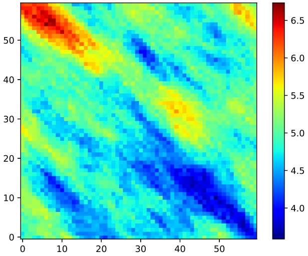
MLEnKF

BMA

Mean example 2



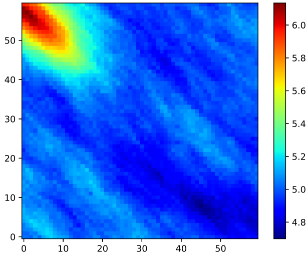
ES with large ensemble



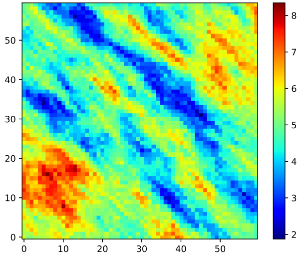
BMA

MLEnKF

Mean example 2



ES with large ensemble



MLEnKF

Summary and conclusion

We have investigated multilevel methods for ensemble based data assimilation

Summary and conclusion

We have investigated multilevel methods for ensemble based data assimilation

MLMC methods are optimal for cases that are not dominated by variance, and with unlimited computational resources

Summary and conclusion

We have investigated multilevel methods for ensemble based data assimilation

MLMC methods are optimal for cases that are not dominated by variance, and with unlimited computational resources

Analysis of two simple two-phase flow problems indicates that MLMC methods are not optimal for these cases

Summary and conclusion

We have investigated multilevel methods for ensemble based data assimilation

MLMC methods are optimal for cases that are not dominated by variance, and with unlimited computational resources

Analysis of two simple two-phase flow problems indicates that MLMC methods are not optimal for these cases

An alternative multilevel method, based on BMA is introduced to handle such cases
– Method aims to reduce variance
– Method is biased

Summary and conclusion

We have investigated multilevel methods for ensemble based data assimilation

MLMC methods are optimal for cases that are not dominated by variance, and with unlimited computational resources

Analysis of two simple two-phase flow problems indicates that MLMC methods are not optimal for these cases

An alternative multilevel method, based on BMA is introduced to handle such cases
– Method aims to reduce variance
– Method is biased

Numerical DA experiments shows that
– MLEnKF fails to estimate the mean
– The alternative method gives good estimates of the mean

Acknowledgements

The authors are grateful for the financial support provided by the CIPR/IRIS cooperative research project “4D Seismic History Matching” which is funded by industry partners Eni, Petrobras, and Total, as well as the Research Council of Norway (PETROMAKS 2)