The University of Tulsa
Petroleum Reservoir Exploitation Projects

Modified ES-MDA Algorithms for Data Assimilation and Uncertainty Quantification

Javad Rafiee and Al Reynolds

12th EnKF Workshop
June 14, 2017

Outline

- Ensemble Smoother with Multiple Data Assimilation (ES-MDA)
- Discrepancy principle and choice of inflation factors in ES-MDA
- Convergence (after Geir Evensen)

ES-MDA

- Define

$$\Delta M^{f,i} = \frac{1}{\sqrt{N_e - 1}} \left[m_1^{f,i} - \bar{m}^{f,i}, \ldots, m_{N_e}^{f,i} - \bar{m}^{f,i} \right], \tag{1}$$

and

$$\Delta D^{f,i} = \frac{1}{\sqrt{N_e - 1}} \left[d_1^{f,i} - \bar{d}^{f,i}, \ldots, d_{N_e}^{f,i} - \bar{d}^{f,i} \right], \tag{2}$$

where $\bar{d}^{f,i} = (1/N_e) \sum_j d_j^{f,i}$ and $\bar{m}^{f,i} = (1/N_e) \sum_j m_j^{f,i}$.

ES-MDA Algorithm

1. Choose the number of data assimilations, N_a, and the coefficients, α_i for $i = 1, ..., N_a$.
2. Generate initial ensemble $\{m_j^{f,1}\}_{j=1}^{N_e}$
3. For $i = 1, ..., N_a$:
   (a) Run the ensemble from time zero,
   (b) For each ensemble member, perturb the observation vector with the inflated measurement error covariance matrix, i.e., $d_{uc,j}^{i} \sim \mathcal{N}(d_{\text{obs}}, \alpha_i C_D)$.
   (c) Use the update equation to update the ensemble.

$$m_j^{a,i} = m_j^{f,i} + \Delta M^{f,i} (\Delta D^{f,i})^T \left[\Delta D^{f,i} (\Delta D^{f,i})^T + \alpha_i C_D \right]^{-1} \left(d_{\text{uc,j}}^{i} - d_j^{f,i} \right)$$

$$m_j^{f,i+1} = m_j^{a,i}$$

- Comment: Requires $\sum_{k=1}^{N_a} \frac{1}{\alpha_k} = 1$.

Dimensionless Sensitivity

- The dimensionless sensitivities control the change in model parameters that occurs when assimilating data (Zhang et al., 2003; Tavakoli and Reynolds, 2010). The standard dimensionless sensitivity is defined as

$$\widehat{G}_D^i \equiv C_D^{-1/2} G(\bar{m}^{f,i}) C_M^{1/2}, \tag{3}$$

where $G(m)$ is the sensitivity matrix for $d^f(m)$ where

$$\widehat{g}_{i,j} = \frac{\partial d_i^f(m)}{\partial m_j}. \tag{4}$$

- Dimensionless sensitivity matrix components are

$$g_{i,j} = \frac{\sigma_{m,j}}{\sigma_{d,i}} \frac{\partial d_i^f}{\partial m_j}. \tag{5}$$

- The direct analogue of the standard dimensionless sensitivity matrix in ensemble based methods is given by

$$G_D^i \equiv C_D^{-1/2} \Delta D^{f,i} \approx C_D^{-1/2} G(\bar{m}^{f,i}) \Delta M^{f,i}. \tag{6}$$

ES-MDA Update Equation

Recall the ES-MDA update equation

$$m_j^{a,i} = m_j^{f,i} + \Delta M^{f,i}(\Delta D^{f,i})^T \left[\Delta D^{f,i}(\Delta D^{f,i})^T + \alpha_i C_D\right]^{-1} \left(d_{\mathrm{uc,j}}^i - d_j^{f,i}\right) \quad (7)$$

Using the definition of the dimensionless sensitivity ($G_D^i \equiv C_D^{-1/2} \Delta D^i$), we can write ES-MDA update equation as

$$m_j^{a,i} = m_j^{f,i} + \Delta M^{f,i}(G_D^i)^T \left[G_D^i (G_D^i)^T + \alpha_i I_{N_\mathrm{d}}\right]^{-1} C_D^{-1/2} \left(d_{uc,j}^i - d_j^{f,i}\right). \quad (8)$$

for $j = 1, \ldots, N_e$.

Why do we need damping?

- ES similar to doing one GN iteration with full step using the same average sensitivity coefficient to update each ensemble method with the forecast as the initial guess.
- $$O(m) = \frac{1}{2} \parallel m - \bar{m} \parallel^2_{C_M^{-1}} + \frac{1}{2} \parallel d^f(m) - d_{\text{obs}} \parallel^2_{C_D^{-1}}$$

  GN based on approximating $O(m)$ by a quadratic but far from a minimum quadratic approximation good only in small region around current model. TR better than line search.
- Proof of convergence of GN requires the possibility of taking a full (unit) step.
- Juris Rommelsee, PhD thesis, TU Delft (2009).

Least Squares Problem

Similar to Eq. 8, one can update the mean of m directly as

$$\bar{m}^{a,i} = \bar{m}^{f,i} + \Delta M^{f,i}(G_D^i)^T \left[G_D^i (G_D^i)^T + \alpha_i I_{N_\mathrm{d}} \right]^{-1} C_D^{-1/2} \left(d_{\mathrm{obs}} - \bar{d}^{f,i} \right). \quad (9)$$

It is easy to show that $\bar{m}^{a,i}$ is the solution of the regularized least squares problem given by

$$x^{a,i} = \arg\min_x \left\{ \frac{1}{2} \left\| G_D^i x - y \right\|^2 + \frac{\alpha_i}{2} \|x\|^2 \right\}, \quad (10)$$

where

$$x = (\Delta M^{f,i})^+ \left(m - \bar{m}^{f,i} \right), \quad (11)$$

$$y = C_D^{-1/2} \left(d_{\mathrm{obs}} - \bar{d}^{f,i} \right), \quad (12)$$

where $(\Delta M^{f,i})^+$ is the pseudo-inverse of $\Delta M^{f,i}$.

Discrepancy Principle

- Assume
$$\|y\| = \|C_D^{-1/2}\left(d_{\text{obs}} - \bar{d}^{f,i}\right)\| > \eta, \tag{13}$$
where η is the noise level given by
$$\eta^2 = \|C_D^{-1/2}\left(d_{\text{obs}} - d^f(m_{\text{true}})\right)\|^2 \approx N_d. \tag{14}$$
- Based on the discrepancy principle the minimum regularization parameter, α_i, should be selected such that
$$\eta = \|G_D^i x^{a,i} - y\| = \| C_D^{-1/2}(\bar{d}^a - d_{\text{obs}}) \| . \tag{15}$$

Discrepancy Principle

- From Eqs. 13 and 15 we can write

$$\|C_D^{-1/2}\left(d_{\text{obs}} - \bar{d}^{f,i}\right)\| > \eta = \alpha_i \left\| \left[G_D^i (G_D^i)^T + \alpha_i I_{N_\text{d}} \right]^{-1} C_D^{-1/2} \left(d_{\text{obs}} - \bar{d}^{f,i}\right) \right\|. \quad (16)$$

Therefore, for some $\rho \in (0,1)$

$$\rho \|C_D^{-1/2}\left(d_{\text{obs}} - \bar{d}^{f,i}\right)\| = \alpha_i \left\| \left[G_D^i (G_D^i)^T + \alpha_i I_{N_\text{d}} \right]^{-1} C_D^{-1/2} \left(d_{\text{obs}} - \bar{d}^{f,i}\right) \right\|. \quad (17)$$

- Hanke (1997) proposed RLM:

$$\rho^2 \left\| C_D^{-1/2}\left(d_{\text{obs}} - \bar{d}^{f,i}\right) \right\|^2 \le \alpha_i^2 \left\| \left[G_D^i (G_D^i)^T + \alpha_i I_{N_\text{d}} \right]^{-1} C_D^{-1/2} \left(d_{\text{obs}} - \bar{d}^{f,i}\right) \right\|^2. \quad (18)$$

- Iglesias (2015) used Eq. 18 for choosing inflation factors in his version of ES-MDA (IR-ES).
- Le et al. (2015) used a much stricter condition based on Eq. 18 for choosing inflation factors in ES-MDA-RLM.

An Analytical Procedure for Calculation of Inflation Factors

Recall that

$$\rho^2 \left\| C_D^{-1/2} \left(d_{\text{obs}} - \bar{d}^{f,i} \right) \right\|^2 \leq \alpha_i^2 \left\| \left[G_D^i (G_D^i)^T + \alpha_i I_{N_\text{d}} \right]^{-1} C_D^{-1/2} \left(d_{\text{obs}} - \bar{d}^{f,i} \right) \right\|^2 . \tag{18}$$

Using the definitions of $y = C_D^{-1/2} \left(d_{\text{obs}} - \bar{d}^{f,i} \right)$ and $C \equiv G_D^i (G_D^i)^T + \alpha_i I_{N_\text{d}}$,

$$\rho^2 \leq \alpha_i^2 \frac{\left\| C^{-1} y \right\|^2}{\left\| y \right\|^2} . \tag{19}$$

$$\frac{\left\| C^{-1} y \right\|^2}{\left\| y \right\|^2} \geq \min_j \gamma_j^2 = \min_j \frac{1}{\left(\lambda_j^2 + \alpha_i \right)^2} = \frac{1}{\left(\lambda_1^2 + \alpha_i \right)^2} \tag{20}$$

where γ_j's are the eigenvalues of C^{-1} and λ_j's are the singular values of G_D^i.

An Approximate Method for Inflation Factors

Instead of enforcing

$$\rho^2 \le \alpha_i^2 \frac{1}{\left(\lambda_1^2 + \alpha_i\right)^2},$$

we use

$$\rho^2 \le \alpha_i^2 \frac{1}{\left(\overline{\lambda}^2 + \alpha_i\right)^2}, \tag{21}$$

$$\alpha_i = \frac{\rho}{1-\rho}\overline{\lambda}^2 \tag{22}$$

where $\overline{\lambda}$ is the average singular value of G_D^i given by

$$\overline{\lambda} = \frac{1}{N}\sum_{j=1}^{N}\lambda_j \quad \text{where} \quad N = \min\{N_d, N_e\}. \tag{23}$$

Motivation: Discrepancy principle overestimates the optimal inflation factor in the linear case.

We use $\rho = 0.5$, so $\alpha_i = \overline{\lambda}^2$.

ES-MDA with Geometric Inflation Factors

- Specify the number of data assimilation steps (N_a).
- Assume that the inflation factors form a monotonically decreasing geometric sequence:

$$\alpha_{i+1} = \beta^i \alpha_1, \tag{24}$$

- Determine

$$\alpha_1 = \overline{\lambda}^2 = \left(\frac{1}{N} \sum_{j=1}^{N} \lambda_j \right)^2 . \tag{25}$$

ES-MDA with Geometric Inflation Factors

- Recall that ES-MDA requires that

$$1 = \sum_{i=1}^{N_a} \frac{1}{\alpha_i} = \sum_{i=1}^{N_a} \frac{1}{\beta^{i-1}\alpha_1}$$

- Solve

$$\frac{1-(1/\beta)^{N_a-1}}{1-(1/\beta)} = \alpha_1, \tag{26}$$

  for β.
- We call the proposed method ES-MDA-GEO.

Comments on "Convergence" of ES-MDA

- Classifying ES-MDA as an iterative ES may be augmentable; stops when $\sum_{k=1}^{N_a} \frac{1}{\alpha_k} = 1$.
- Criterion based on ensuring methods samples correctly in the linear Gaussian case as ensemble size goes to infinity.
- Analogue of Hanke's suggestion for RLM, should terminate ES-MDA when

$$\frac{1}{N_d} \| C_D^{-1/2} \left(d_{\text{obs}} - \bar{d}^{f,i} \right) \|^2 < \tau^2$$

where $\tau > 1/\rho = 2$.
- This means, terminate when the normalized objective function is less that 4.
- GE: Does ES-MDA converge as $N_a \to \infty$? To what?

Numerical Examples

- The performance of ES-MDA-GEO is compared to IR-ES, ES-MDA-RLM and ES-MDA-EQL.
- To investigate the performance of the methods, we define the following measures:

$$\text{RMSE} = \frac{1}{N_e} \sum_{j=1}^{N_e} \left(\frac{1}{N_m} \sum_{k=1}^{N_m} (m_{\text{true},k} - m_{j,k})^2 \right)^{1/2}, \tag{27}$$

$$\overline{\sigma} = \frac{1}{N_m} \sum_{k=1}^{N_m} \sigma_k, \tag{28}$$

$$O_{Nd} = \frac{1}{\mathrm{N_e N_d}} \sum_{j=1}^{\mathrm{N_e}} (d_j^f - d_{\text{obs}})^T C_D^{-1} (d_j^f - d_{\text{obs}}). \tag{29}$$

Example 1: 2D Waterflooding

Two-dimensional waterflooding problem:

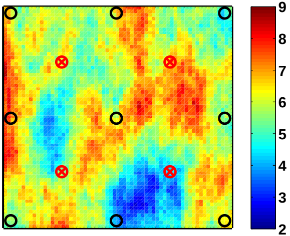
- 64×64×1 grid.
- 9 production wells (BHP control).
- 4 injection wells (BHP control).

Observed data:

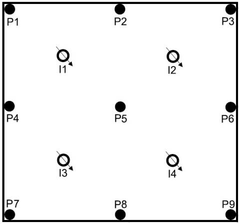
- Oil and water production rates and water injection rates.
- Standard deviations of measurement error: 3% of true data.
- Data from the first 36 months are history-matched and data for next 20 are used for prediction.

Model parameters:

- The gridblock log-permeabilities are considered as the model parameters.



True permeability field



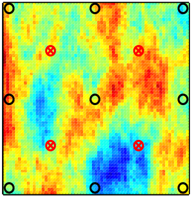
Well locations

Example 1: Results

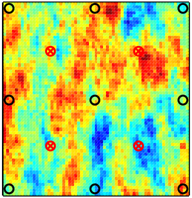
- An ensemble of 400 realizations is generated from the prior distribution.
- First inflation factor from DP is 1049.4; N_a of 4 and 6, respectively, give β equal to 0.102 and 0.264.
- Comment IR-ES with $\rho = 0.8$ did not converge after 200 iterations.

	Prior	ES-MDA-RLM	IR-ES	ES-MDA-EQL		ES-MDA-GEO	
		ρ=0.5	ρ=0.5	N_a =4	N_a =6	N_a =4	N_a =6
RMSE	2.23	0.613	0.902	1.45	1.09	0.586	0.633
$\overline{\sigma}$	0.995	0.334	0.517	0.258	0.255	0.380	0.362
O_{Nd}	16121	1.06	8.14	8.45	1.344	25.2	5.78
Iter	-	21	9	4	6	4	6

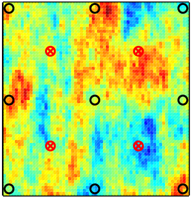
The posterior mean of the log-permeability



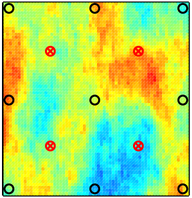
(a) True



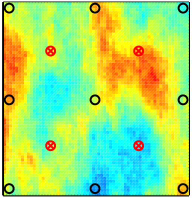
(b) ES-MDA-EQLx4



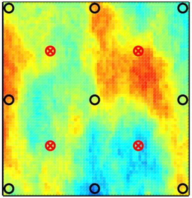
(c) ES-MDA-EQLx6



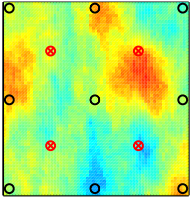
(d) ES-MDA-GEOx4



(e) ES-MDA-GEOx6



(f) ES-MDA-RLM 0.5



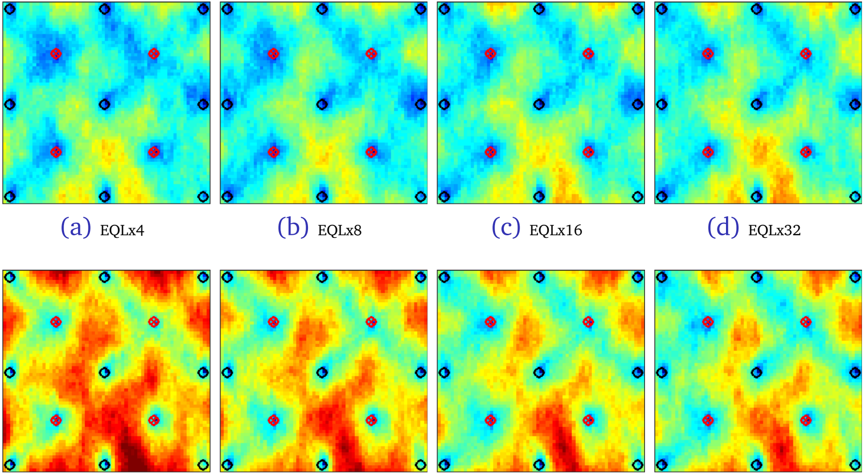
(g) IR-ES 0.5

“Convergence” Results

	Prior	ES-MDA-EQL					ES-MDA-GEO				
Iter	-	4	8	16	32	64	4	8	16	32	64
RMSE	2.23	1.451	0.977	0.969	0.838	0.732	0.586	0.537	0.553	0.560	0.585
$\overline{\sigma}$	0.995	0.258	0.257	0.267	0.275	0.284	0.380	0.351	0.329	0.317	0.312
O_{Nd}	16121	8.451	1.094	0.947	0.907	0.922	25.246	6.689	1.413	0.978	0.905

Table: Effect of number of iteration on ES-MDA

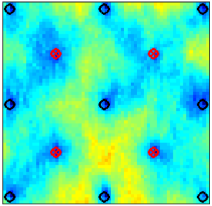
Posterior S.D. Versus N_a with 95% Truncation



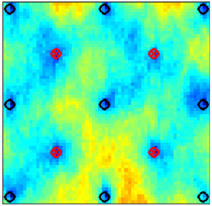
(a) EQLx4 (b) EQLx8 (c) EQLx16 (d) EQLx32

(e) GEOx4 (f) GEOx8 (g) GEOx16 (h) GEOx32

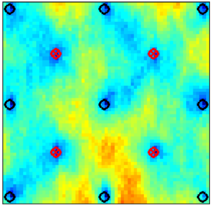
Posterior S.D. Versus N_a with 95% Truncation



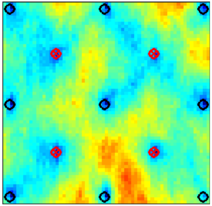
(a) EQLx8



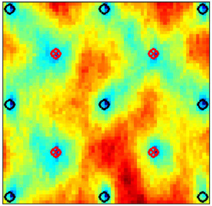
(b) EQLx16



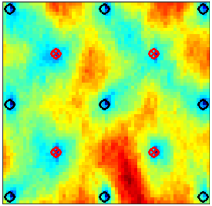
(c) EQLx32



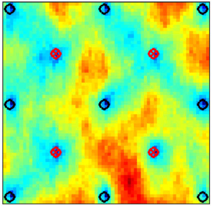
(d) EQLx64



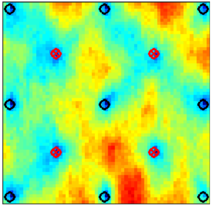
(e) GEOx8



(f) GEOx16

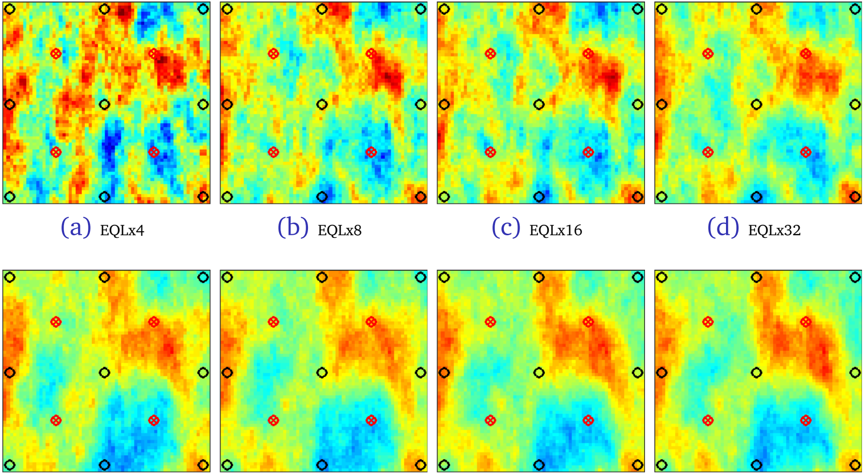


(g) GEOx32



(h) GEOx64

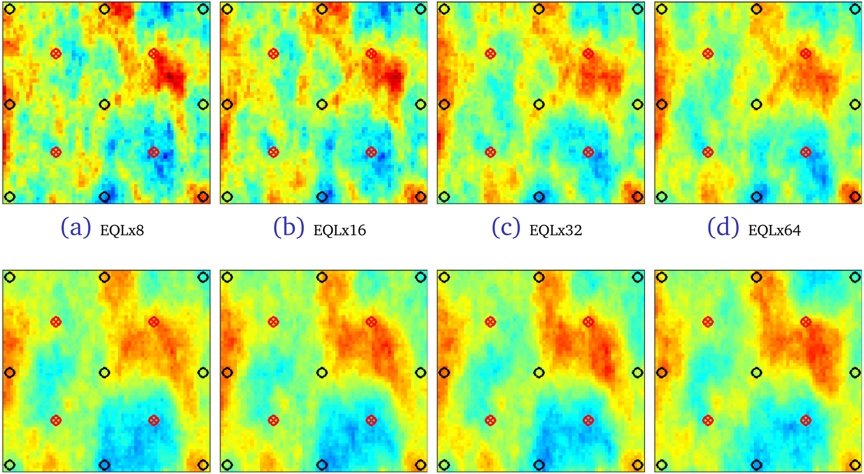
Posterior Mean Versus N_a with 95% Truncation



(a) EQLx4 (b) EQLx8 (c) EQLx16 (d) EQLx32

(e) GEOx4 (f) GEOx8 (g) GEOx16 (h) GEOx32

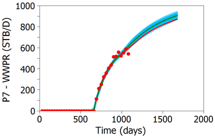
Posterior Mean Versus N_a with 95% Truncation



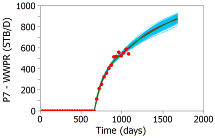
(a) EQLx8 (b) EQLx16 (c) EQLx32 (d) EQLx64

(e) GEOx8 (f) GEOx16 (g) GEOx32 (h) GEOx64

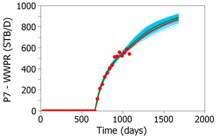
Data Match - P7 Water Rate



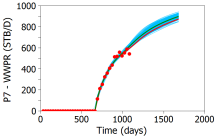
(a) EQLx8



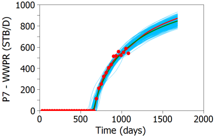
(b) EQLx16



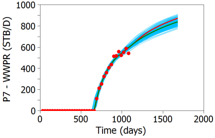
(c) EQLx32



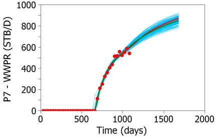
(d) EQLx64



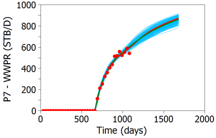
(e) GEOx8



(f) GEOx16

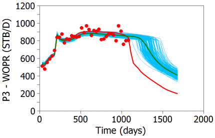


(g) GEOx32

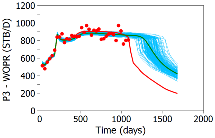


(h) GEOx64

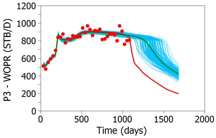
Data Match - P3 Oil Rate



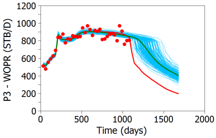
(a) EQLx8



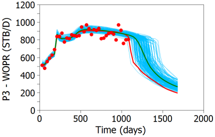
(b) EQLx16



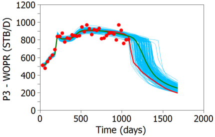
(c) EQLx32



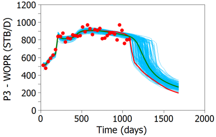
(d) EQLx64



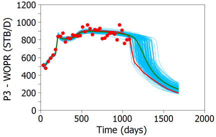
(e) GEOx8



(f) GEOx16

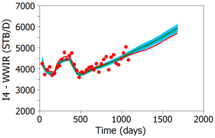


(g) GEOx32

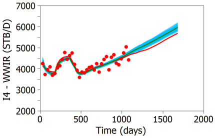


(h) GEOx64

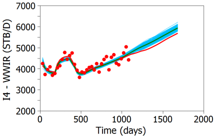
Data Match - I4 Injection Rate



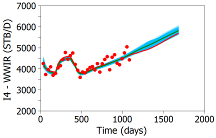
(a) EQLx8



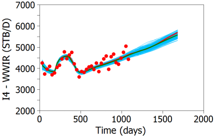
(b) EQLx16



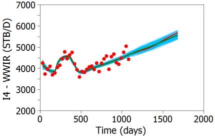
(c) EQLx32



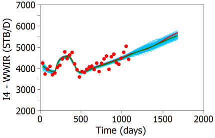
(d) EQLx64



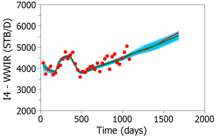
(e) GEOx8



(f) GEOx16



(g) GEOx32



(h) GEOx64

Results

	Prior	ES-MDA-EQL					ES-MDA-GEO				
Iter	-	4	8	16	32	64	4	8	16	32	64
RMSE	2.23	1.451	0.977	0.969	0.838	0.732	0.586	0.537	0.553	0.560	0.585
$\overline{\sigma}$	0.995	0.258	0.257	0.267	0.275	0.284	0.380	0.351	0.329	0.317	0.312
O_{Nd}	16121	8.451	1.094	0.947	0.907	0.922	25.246	6.689	1.413	0.978	0.905

Table: Effect of number of iteration on ES-MDA

Summary and Conclusions

- We presented analytical expression that enables the exact calculation of the minimum inflation factor that satisfies the inequality derived from the discrepancy principle that is the basis of IR-ES.
- The ES-MDA-GEO algorithm developed here is an efficient data assimilation method that allows the user to specify a priori the number of data assimilation step.
- ES-MDA-GEO is more robust than using the original ES-MDA algorithm with equal inflation factors.
- ES-MDA-GEO and ES-MDA-equal appear to converge to different distributions. Which is best?
- The performance of IR-ES highly depend on the parameters ρ, and IR-ES with $\rho = 0.8$ (suggested by the author) did not converge after 200 iterations.