My Decade-Long Journey Through the Field of Ensemble-Based Data Assimilation

Al Reynolds, Alexandre Emerick, Duc Le, Mei Han, Gaoming Li,
Reza Tavakoli, Kristian Tulin, Mohammad Zafari, Yong Zhao
The University of Tulsa
Petroleum Reservoir Exploitation Projects

23 June 2015
9th International EnKF Workshop

Introduction

- A Starting Point.
- Relationship to Gradient-Based Data Assimilation.
- Localization.
- Pseudo-Inverse and Subspace Methods.
- ES-MDA, Field Case.
- ES-MDA (Adaptive) with Illustration.
- Non-Gaussian Geology.

Starting Points

- My failed beginning: Naevdal et al., SPE 75235 (2002), Evensen, *J. Geophysical Research Research* (1994).
- Evensen, The ensemble Kalman filter: theoretical formulation and practical implementation *Ocean Dynamics* (2003), "The combined parameter and state estimation problem," (2005 manuscript).
- Additional reading: G. Evensen, *Data Assimilation: The Ensemble Kalman Filter* Springer, 2009. Two review papers: G. Evensen, *IEEE Control Systems Magazine* (2009); Aanonsen et al., *SPE Journal* 2009.

The Dawn

- EnKF is essentially equivalent to doing one Gauss-Newton iteration with a full-step using an average sensitivity coefficient to update each realization at each data assimilation time; SIAM Geoscience Conference 2005 (Avignon) and Reynolds et al. "Iterative Forms of the Ensemble Kalman Filter," ECMOR X (2006).
- Randomized maximum likelihood for parameter estimation/simulation, Oliver et al. ECMOR (1996), provides an approximate sampling of $f(m|d_{\text{obs}}) \propto \exp(-O(m))$

$$O(m) = \frac{1}{2}(m - m_{\text{prior}})^T C_M^{-1}(m - m_{\text{prior}}) + \frac{1}{2}(d^f(m) - d_{\text{obs}})^T C_D^{-1}(d^f(m) - d_{\text{obs}})$$

$$m_{\text{prior}} \leftarrow m_{uc,j} \sim N(m_{\text{prior}}, C_M),\ d_{\text{obs}} \leftarrow d_{uc,j} \sim N(d_{\text{obs}}, C_D),$$

RML

- Minimizing with Gauss-Newton gives

$$m_{c,j}^{\ell+1} = \alpha_\ell m_{uc,j} + (1-\alpha_\ell) m_j^\ell + \alpha_\ell C_M G_{\ell,j}^T (C_D + G_{\ell,j} C_M G_{\ell,j}^T)^{-1}$$
$$\times (d_{uc,j} - d^f(m_j^\ell) + G_{\ell,j}(m_j^\ell - m_{uc,j})) \quad \text{for} \quad j = 1, 2, \cdots N_e.$$

  $G_{\ell,j}$ is the sensitivity matrix evaluated at m_j^ℓ, the equation is Gauss-Newton iteration for minimizing

- Samples correctly in the linear ($d^f(m) = Gm$) Gaussian case. Note to obtain correct sampling in the linear-Gaussian case; it is necessary to perturb the data, i.e., d_{obs} is replaced by $d_{uc,j} \sim \mathcal{N}(d_{\text{obs}}, C_D)$; Oliver, *Mathematical Geology* (1996); Reynolds et al. AAPG Memoir 71 (1999); Burgers et al. *Monthly Weather Review* (1998).

RML

- One iteration ($\ell = 0$); initial guess equal to unconditional realization ($m_j^0 = m_{uc,j}$); full step ($\alpha_0 = 1$), all $G_{\ell,j}$ replaced by $\bar{G} = G(\bar{m}^0)$

$$m_{c,j}^{\ell+1} = \alpha_\ell m_{uc,j} + (1-\alpha_\ell) m_j^\ell + \alpha_\ell C_M G_{\ell,j}^T (C_D + G_{\ell,j} C_M G_{\ell,j}^T)^{-1}$$
$$\times (d_{uc,j} - d^f(m_j^\ell) + G_{\ell,j}(m_j^\ell - m_{uc,j})) \quad \text{for} \quad j = 1, 2, \cdots N_e$$

$$m_{c,j}^1 = m_{uc,j} + C_M \bar{G}^T (C_D + \bar{G} C_M \bar{G}^T)^{-1} (d_{uc,j} - d^f(m_{uc,j})).$$

or

$$m_j^a = m_j^f + C_M \bar{G}^T (C_D + \bar{G} C_M \bar{G}^T)^{-1} (d_{uc,j} - d_j^f). \qquad (1)$$

EnKF - **One Data Assimilation Step**, Reynolds et al. ECMOR (2006); SIAM Geosciences (2005);

$$m_j^a = m_j^f + \frac{1}{N_e - 1}\Delta M^f (\Delta D^f)^T \left(C_D + \frac{1}{N_e - 1}(\Delta D^f)(\Delta D^f)^T \right)^{-1} (d_{uc,j} - d_j^f)$$

$$\bar{m}^f = \frac{1}{N_e}\sum_{j+1}^{N_e} m_j^f \qquad \bar{d}^f = \frac{1}{N_e}\sum_{j+1}^{N_e} d_j^f$$

$$\Delta M = [\cdots m_j^f - \bar{m}^f \cdots] \qquad \Delta D^f = [\cdots d_j^f - \bar{d}^f \cdots]$$

$$\tilde{C}_{MD}^f \equiv \frac{1}{N_e - 1}\Delta M^f (\Delta D^f)^T = \frac{1}{N_e - 1}\sum_{j=1}^{N_e} (m_j^f - \bar{m}^f)(d_j^f - \bar{d}^f)^T$$

We prefer replacing $\bar{d}^f$ by $d^f(\bar{m}^f)$ although this second term is not necessarily a good approximation of the first. Then

$$(d_j^f - \bar{d}^f)^T = (d_j^f - d^f(\bar{m}^f))^T = (G(\bar{m}^f)(m_j^f - \bar{m}^f) + e)^T \approx (m_j^f - \bar{m}^f)^T \bar{G}^T$$

EnKF - **One Data Assimilation Step**

$$m_j^a = m_j^f + \frac{1}{N_e - 1}\Delta M^f (\Delta D^f)^T \left(C_D + \frac{1}{N_e - 1}(\Delta D^f)(\Delta D^f)^T \right)^{-1} (d_{uc,j} - d_j^f)$$

$$\tilde{C}_{MD}^f \equiv \frac{1}{N_e - 1}\Delta M^f (\Delta D^f)^T = \frac{1}{N_e - 1}\sum_{j=1}^{N_e}(m_j^f - \bar{m}^f)(d_j^f - \bar{d}^f)^T$$

$$\frac{1}{N_e - 1}\sum_{j=1}^{N_e}(m_j^f - \bar{m}^f)(m_j^f - \bar{m}^f)^T \bar{G}^T \approx \tilde{C}_M^f \bar{G}^T$$

Similarly,

$$C_{DD}^f \equiv \frac{1}{N_e - 1}(\Delta D^f)(\Delta D^f)^T = \bar{G}\tilde{C}_M^f \bar{G}^T + e_d$$

$$m_j^a = m_j^f + \tilde{C}_M^f \bar{G}^T \left(C_D + \bar{G}\tilde{C}_M^f \bar{G}^T \right)^{-1} (d_{uc,j} - d_j^f).$$

the same results as we had for one iteration of Gauss-Newton ...

Comments

- Iterative ensemble smoother methods: ES-MDA (adaptive) or Chen- Oliver, LevenbergâMarquardt-Iterative-ES, Computational Geosciences (2013) (essentially utilizes a truncated SVD of dimensionless sensitivity matrix).
- Suggests that we can improve performance of EnKF (at least the data match) by an iterative process that mimics Gauss-Newton iteration. Some of the proposed iterative EnKF schemes are compared for a simple reservoir problem in Emerick and Reynolds *Computational Geosciences*, (2013).
- Even with the prior regularization term, a full-step of Gauss-Newton often leads to overshooting and undershooting, i.e., extremely high or extremely low value of property fields so additional regularization is sometimes required especially if noise level is low.

Peripheral Questions

- How can updated (analyzed) simulation variables honor material balance? How can updated (analyzed) reservoir variables (parameters) be in any sense consistent with updated states (primary variables predicted from forward model)? For linear-Gaussian case, they are statistically consistent; Thulin et al. SPE 109975 (2007) (Computations in this paper are incorrect.) In highly nonlinear-case, inconsistency cannot be avoided unless we rerun updated models from time zero after some data assimilation step; this inconsistency issue can be avoided by using the ensemble smoother where all data are simulated at once and only reservoir parameters are estimated.
- Spurious correlations result from sampling error due to small ensemble size and must be dealt with by some form of covariance or Kalman gain localization.

Another Problem of Limited Ensemble Size

- Each updated vector of model parameters is a linear combination of the initial ensemble of models. Note for parameter vector, $m_j^{f,n} = m_j^{a,n-1}$.

$$m_j^{n,a} = m_j^{n,f} + \frac{1}{N_e - 1} \Delta M^{f,n} (\Delta D^{f,n})^T \left(C_D^n + \frac{1}{N_e - 1} \Delta D^{f,n} (\Delta D^{f,n})^T \right)$$

$$= m_j^{n-1,a} + \frac{1}{N_e - 1} \Delta M^{a,n-1} \left[(\Delta D^{f,n})^T \left(C_D^n + \frac{1}{N_e - 1} \Delta D^{f,n} (\Delta D^{f,n})^T \right) \right.$$

$$= m_j^{n-1,a} + \Delta M^{a,n-1} x_j = m_j^{n-1,a} + \sum_{i=1}^{N_e} (x^j)_i (m_i^{n-1,a} - \bar{m}^{n-1,a}).$$

- All analyzed reservoir models are in the subspace spanned by the ensemble of initial realizations; choose your initial realizations wisely; Oliver and Chen *Computational Geosciences* (2009); Dovera and Della Rossa *Computational Geosciences* (2012).

Limited Ensemble Size

- All analyzed reservoir models are in the subspace spanned by the ensemble of initial realizations. Assume $N_m > N_d > N_e$, $\mathscr{R}(\Delta M^{f,n}) \leq N_e - 1$, thus, we have only $N_e - 1$ degrees of freedom available to adjust data. May not be able to match data well. As we keep assimilating data, may diverge farther from true state.
- Lorenc Q. J. R. Meteorol. Soc. (2003) shows that a perfect observation (zero noise) results in a loss of one degree of freedom in the ensemble.

Rescaling and Pseudo-Inverse

- At each EnKF analysis step we must invert an $N_n \times N_n$ matrix C given by

$$C = HC_Y^f H^T + C_D = C_{DD}^f + C_D.$$

  If C_D is positive-definite C_{DD}^f is a real-symmetric positive semi-definite matrix, but may be poorly conditioned, hence truncated SVD (TSVD) is usally used for inversion. This can lead to loss of information when data measurement errors have significantly different scales.
- For example, the information leading to water-cut data can be lost (problem with computations in Thulin et al. paper mentioned earlier) so that water cut data cannot be matched; see Wang et al. *SPEJ*, 2009.
- $C_D = \text{diag}(\sigma_{d,i}^2)$.
- Rescale as

$$C = C_D^{1/2}\left[C_D^{-1/2} C_{DD}^f C_D^{-T/2} + I_{N_n}\right] C_D^{T/2}$$

Rescaling and Pseudo-Inverse

- $$C = C_D^{1/2}\left[C_D^{-1/2} C_{DD}^{f} C_D^{-T/2} + I_{N_n}\right] C_D^{T/2}$$
- Truncated SVD is now applied to the matrix in square brackets denoted by $\tilde{C}$, i.e.,
$$\widetilde{C} = \widetilde{U}_r \widetilde{\Lambda}_r \widetilde{U}_r^{T},$$
with the pseudo-inverse of C given by
$$C^{+} = C_D^{-T/2} \widetilde{U}_r \widetilde{\Lambda}_r^{-1} \widetilde{U}_r^{T} C_D^{-1/2}.$$
Truncate when
$$\frac{\sum_{i=1}^{N_r} \lambda_i}{\sum_{i=1}^{N_n} \lambda_i} \leq \xi = 0.999.$$

Comments on Scaling

- $$\widetilde{C} = I_{N_n} + C_{\mathrm{D}}^{-1/2} C_{\mathrm{DD}}^{f} C_{\mathrm{D}}^{-\mathrm{T}/2} = I_{N_n} + C_{\mathrm{D}}^{-1/2} \bar{G} C_M \bar{G}^{\,\mathrm{T}} C_{\mathrm{D}}^{-\mathrm{T}/2}$$
$$= I_{N_n} + C_{\mathrm{D}}^{-1/2} \bar{G} C_M^{1/2} C_M^{\mathrm{T}/2} \bar{G}^{\,\mathrm{T}} C_{\mathrm{D}}^{-\mathrm{T}/2}$$
$$= I_{N_n} + C_{\mathrm{D}}^{-1/2} \bar{G} C_M^{1/2} \left(C_{\mathrm{D}}^{-1/2} \bar{G} C_M^{1/2} \right)^T = I_{N_n} + G_D G_D^{\mathrm{T}}.$$

- One model, one datum,

$$g = \frac{\sigma_m}{\sigma_d} \frac{\partial d}{\partial m}$$

Comments on Scaling

- $\widetilde{C} = I_{N_n} + G_D G_D^{\mathsf{T}}$. For linear-Gaussian case, Tavakoli and Reynolds [*SPEJ* (2009), *Comp. Geo.* (2011) showed that the singular values of G_D govern the reduction in uncertainty in the model obtain by assimilating data. (Ideas of their methods go back to Vogel and Wade, "Iterative SVD-based methods for ill-posed problems," SIAM J. Sci. Comput. (1994).
- $$\frac{V'}{V} = \sqrt{\frac{\det C_{\mathrm{MAP}}}{\det C_M}} = \sqrt{\prod_{i=1}^{N_d} \frac{1}{1+\lambda_i^2}},$$

  where λ_i's are singular values of G_D.
- Applying TSVD to $\widetilde{C}$ small λ_i corresponds to eliminating small singular values of G_{D} which have the smallest influence on the reduction of uncertainty. In this sense, the rescaling procedure presented in this section is optimal. Chen-Oliver LM-IES uses TSVD of the ES analogue of G_D.

Subspace Inversion

- Evensen, *Ocean Dynamics* (2004) introduced a subspace inversion scheme which is more computationally more efficient that the pseudo-inversion when $N_e << N_n$; Also see Skjervheim et al *SPEJ* (2007). Useful for assimilation of large data sets, e.g., seismic data. With the Evensen subspace inversion, it may not be necessary to rescale because C_D is left intact. (A scaled version of it is given in Emerick and Reynolds, *Computational Geosciences* (2012) but it involves computing $C_D^{1/2}$.
- Subspace inversion uses the TSVD $(U_r W_r (V_r)^T)$ of the $N_d \times N_e$ matrix ΔD^f :

$$C = C_{DD}^f + C_D = \Delta D^f (\Delta D^f)^T + C_D = (U_r W_r (V_r)^T)(U_r W_r (V_r)^T)^T + C_D \approx U_r W_r \left[I_{N_r} + W_r^{-1} U_r^T C_D U_r W_r^{-1} \right] W_r U_r^T$$

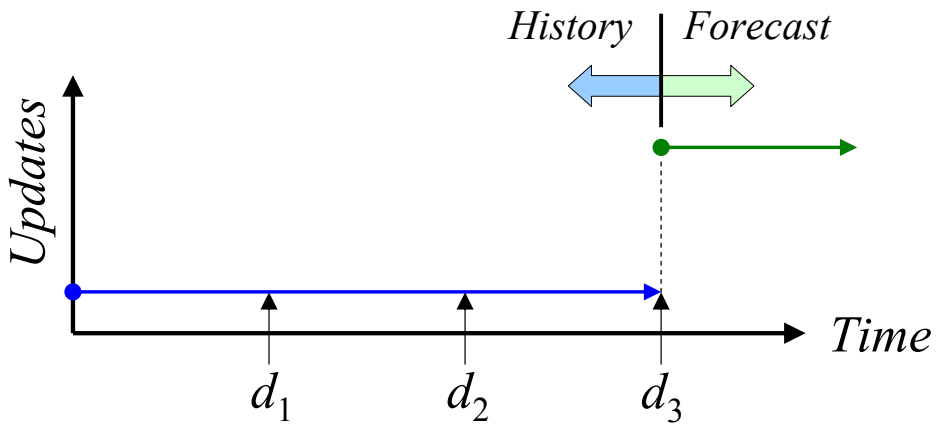
where when convenient, we have assumed $U_r U_r^T = I$.

Subspace Inversion

- $$C \approx U_r W_r \left[I_{N_r} + W_r^{-1} U_r^{\mathsf{T}} C_{\mathsf{D}} U_r W_r^{-1} \right] W_r U_r^{\mathsf{T}}.$$
- $$W_r^{-1} U_r^{\mathsf{T}} C_{\mathsf{D}} U_r W_r^{-1} = Z_r \Lambda_r Z_r^T$$
- $$C = C_{\mathsf{DD}}^f + C_{\mathsf{D}} = U_r W_r \left[I_{N_r} + Z_r \Lambda_r Z_r^T \right] W_r U_r^{\mathsf{T}}$$
- $$C^+ = (C_{\mathsf{DD}}^f + C_{\mathsf{D}})^+ = U_r W_r^{-1} Z_r \left[I_{N_r} + \Lambda_r \right]^{-1} Z_r^T W_r^{-1} U_r^T$$

Ensemble Smoother (ES)

- Similar to EnKF, but with a single update with all data available, i.e., no sequential data assimilation.



- Parameter-estimation problem.
- Faster and easier to implement than EnKF.
- Problem: ES often yields a data match significantly inferior to that obtained with EnKF.

EnKF (and ES) as Gauss-Newton

- Conjecture: Perhaps sequential assimilation is one of the reasons why EnKF gives acceptable results when assimilating production data that are fairly closely-spaced in time (sequential updates are similar to multiple GN updates).
- As noted previously, a single GN update, which may not be enough for conditioning the realizations to the observations, and can suffer from overcorrection.
- ES-MDA: Assimilate all data at once but assimilate it N_a times with inflated measurement error covariance matrix (Emerick and Reynolds, 4 papers). Motivated by 2009 PhD dissertation of Rommelse, TUDelft.

Multiple Data Assimilation

- Assimilate the same data N_a times with inflated measure error covariance matrix, $C_D \leftarrow \alpha_i C_D$. Provides regularization and avoids over correction.
- Single and multiple data assimilations are equivalent for the linear-Gaussian case provided that

$$\sum_{i=1}^{N_a} \frac{1}{\alpha_i} = 1^\star \quad \text{(ex. } \alpha_i = N_a \text{ for } i = 1, \ldots, N_a\text{)}.$$

- We replace a single (and potentially large correction) by N_a smaller corrections.
- MDA can be interpreted as applying the first iteration of the Levenberg-Marquardt algorithm N_a times (Emerick and Reynolds, 2012) and is very similar to ensemble-based version of regularizing Levenberg-Marquardt (Hanke, 1999, Iglesias and Dawson, (2013); Iglesias (2014); Bergemann and Reich, “A Mollified Ensemble Kalman Filter” (2010)

ES-MDA Procedure

1. Choose the number of data assimilations, N_a, and the coefficients α_i for $i = 1, \ldots, N_a$.
2. For $i = 1$ to N_a:
   1. Run the ensemble from time zero.
   2. For each ensemble member, perturb the observation vector using

$$d_{\text{uc},j} = d_{\text{obs}} + \sqrt{\alpha_i} C_{\text{D}}^{1/2} z_d,$$

   where $z_d \sim \mathcal{N}(0, I_{N_d})$.
   3. Update the ensemble using

$$m_j^a = m_j^f + \widetilde{C}_{\text{MD}}^f \left(\widetilde{C}_{\text{DD}}^f + \alpha_i C_{\text{D}} \right)^{-1} \left(d_{\text{uc},j} - d_j^f \right),$$

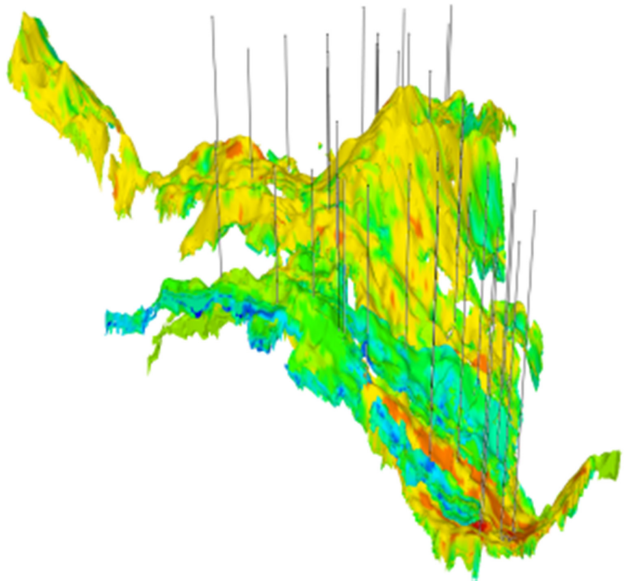
   for $j = 1, 2, \cdots, N_e$.

ES-MDA, Historical Notes

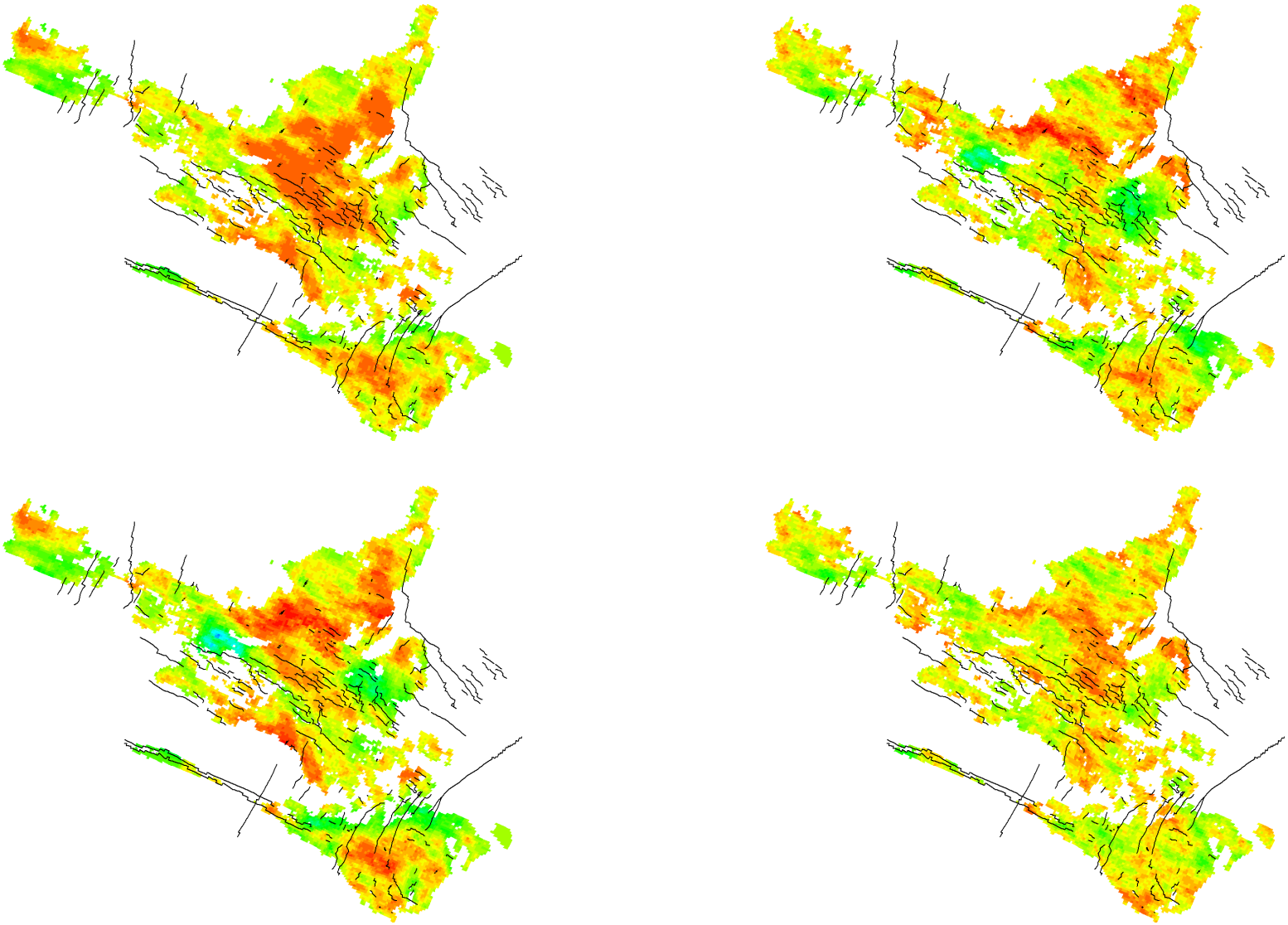
- For 1 parameter and 1D data vector, several people have shown this samples correctly in the linear-Gaussian case: Rommelse (2009), Oliver and Chen. " Improved initial sampling for the ensemble Kalman Filter" (2009); Bergemann and Reich, "A Mollified Ensemble Kalman Filter" (2010).
- Our proof of correct sampling in the linear Gaussian case was general and based on linear algebra. Henning Omre pointed out that the result is obvious it is based on simply factoring the likelihood function and using sequential updating; also see Bergemann and Reich.

Field Case 1

- Turbidite reservoir in Campus Basis.
- Observed data:
  - 20 producers: oil rate, water rate, GOR, bottom-hole pressure.
  - 10 water injection: bottom-hole pressure.
- Initial ensemble:
  - 200 models.
  - Porosity and permeability ($> 125,000$ active gridblocks).
  - Anisotropic ratio k_v/k_h.
- Data assimilation with ES-MDA $(4\times)$ with localization.

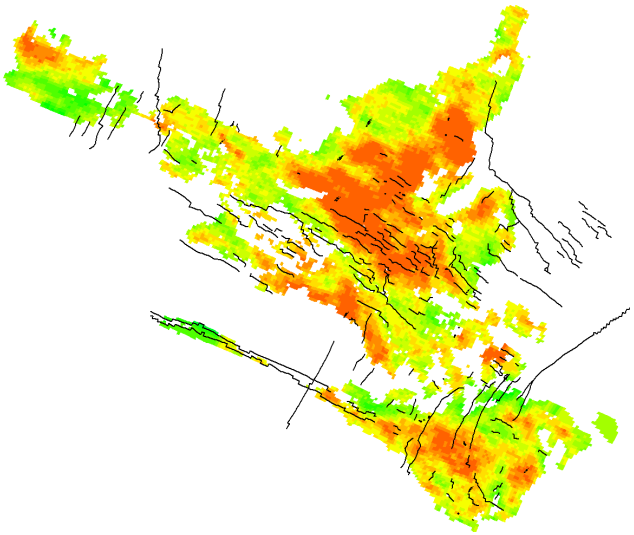


Field Case 1: Model Plausibility – Permeability

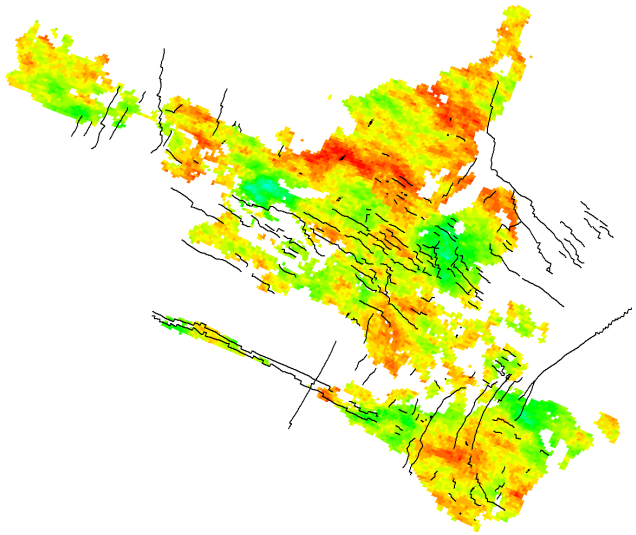


Field Case 1: Model Plausibility – Permeability

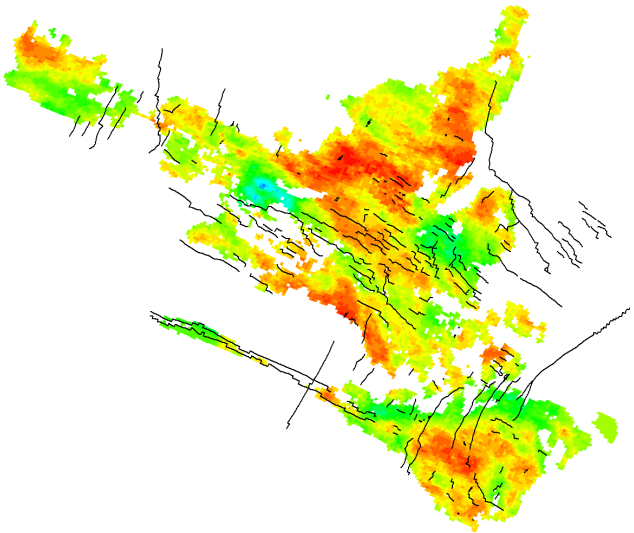
Prior # 200



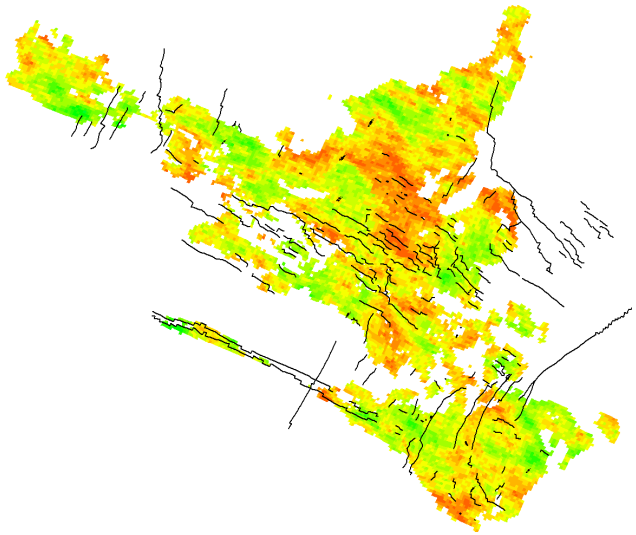
Post # 1



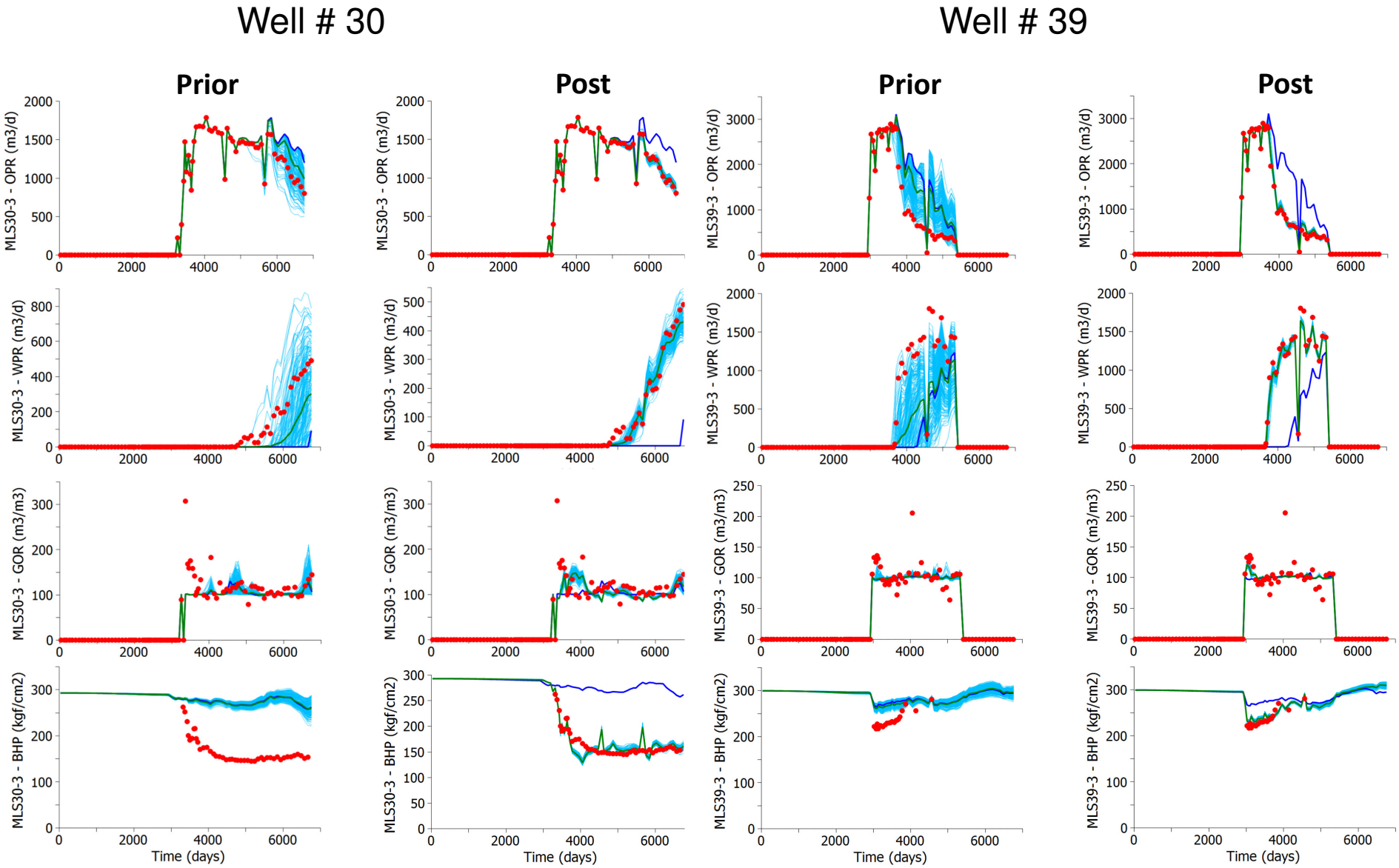
Post # 200



Prior # 1



Field Case 1: Well Data



Field Case 1: Time Spent in the Study

- Generation of the initial ensemble (PETREL) $\approx$ 2 days.
- Reservoir simulator conversion (ECLIPSE to IMEX) $\approx$ 2 weeks.
- File preparation $\approx$ 1 day.
- Test runs and sensitivity analysis $\approx$ 1 week.
- Data assimilation (ES-MDA) $\approx$ 2 days*.
- Total $\approx$ 4 weeks.

$\star$ Time for each reservoir simulation $\approx$ 1.5 hours.
Approximately 40 simultaneous reservoir simulations.
Total of 1,000 simulations ($4 \times 200 + 200$).

Adaptive ES-MDA

- Larger inflation factors at early iterations damp the change in model parameters and tend to prevent excessive roughness.
- We propose two methods to choose the inflation factors.
  - The 1st method is intuitive and works by limiting the maximum change of model parameters at each iteration.
  - The 2nd method is based on a theory on the regularization of least-squares inverse problems.

Method 1

At each iteration:

1. Run simulations for the ensemble.
2. Calculate the average normalized objective function

$$\overline{O}_{Nd} = \frac{1}{N_e}\frac{1}{N_d}\sum_{j=1}^{N_e}(d_j^f - d_{obs})^T C_D^{-1}(d_j^f - d_{obs})$$

3. Set $\alpha = 0.25 * \overline{O}_{Nd}$ as the initial guess for the inflation factor.
4. Calculate the new model parameters using the ES-MDA update equation.
5. Check all ensemble members to make sure that no model parameter is changed by more than 2 (prior) standard deviations. If violated, double α and return to step 4.

Regularizing Levenberg-Marquardt, Hanke (1997); Iglesias-Dawson (2013)

- $$m^{\ell+1} = m^{\ell} + C_M G_\ell^T (G_\ell C_M G_\ell^T + \alpha_\ell C_D)^{-1} (d_{obs} - g(m^\ell))$$

- The last equation is the same structure as ES-MDA:

$$m_j^a = m_j^f + \widetilde{C}_{\text{MD}}^f \left(\widetilde{C}_{\text{DD}}^f + \alpha_i C_{\text{D}} \right)^{-1} \left(d_{\text{uc},j} - d_j^f \right)$$

e.g.,

$$\widetilde{C}_{\text{MD}}^f \approx C_M \bar{G} \quad \text{and} \quad \widetilde{C}_{\text{DD}}^f = \bar{G} C_M \bar{G}^T$$

Regularizing Levenberg-Marquardt

- At each iteration, α_ℓ is chosen such that

$$\rho^2 ||C_D^{-1/2}(d_{\text{obs}} - g(m^\ell))||^2 \leq \alpha^2 ||C_D^{1/2}(G_\ell C_M G_\ell^T + \alpha C_D)^{-1}(d_{\text{obs}} - g(m^\ell))||^2,$$

  for some ρ with $0 < \rho < 1$. Larger ρ requires larger α_ℓ and more iterations but we do more damping at each iteration.

- By the simple analogy, we choose the α_i adaptively (Adaptive ES-MDA) by requiring at the ith data assimilation step:

$$\rho^2 ||C_D^{-1/2}(d_{uc,j} - d_j^f)||^2 \leq \alpha_i^2 ||C_D^{1/2}(C_{DD}^f + \alpha_i C_D)^{-1}(d_{uc,j} - d_j^f)||^2.$$

Adaptive ES-MDA

- Iglesias and Dawson also provide a stopping criteria to avoid overmatching data, but we have not found a need to do that. Instead we stop when the sequence of $1/\alpha_i$ sum to unity to ensure correct sampling in the linear Gaussian case.

Method 2

At each iteration:

1. Run simulations for the ensemble.
2. Calculate the average normalized objective function $\overline{O}_{Nd}$.
3. Set $\alpha = 0.25 * \overline{O}_{Nd}$ as the initial guess for the inflation factor.
4. Check the following conditions for all ensemble members:

$$\rho^2 ||C_D^{-1/2}(d_{uc,j} - d_j^f)||^2 \leq \alpha^2 ||C_D^{1/2}(C_{DD}^f + \alpha C_D)^{-1}(d_{uc,j} - d_j^f)||^2.$$

   If violated, double α and recheck.
5. Apply ES-MDA update to obtain new model parameters.
6. Stop when $\sum_i \frac{1}{\alpha_i} = 1$.

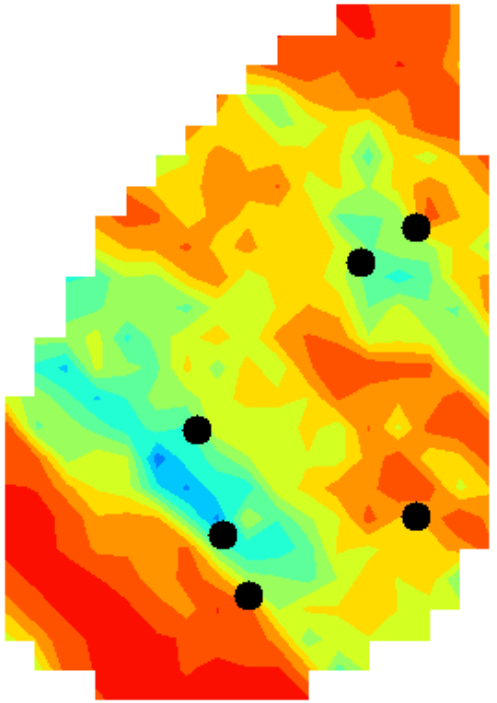
Case Description

Three-dimensional PUNQ-S3 problem:

- Six production wells under rate control.
- Analytical aquifer along the rim of the reservoir
- Three phase flow.

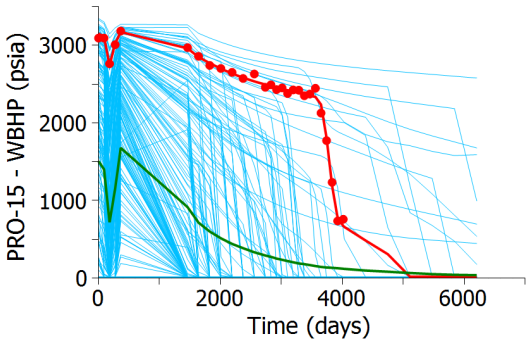
Observed data and Parameters:

- Bottom hole pressure, gas-oil ratio, water cut.
- Standard deviations of measurement error: 10 psi, 3%, 3%.
- Case difficult: 3 layers have channels with no hard data, large number of mixed parameters, low noise levels; there are 247 data but some of the pressure data are effectively lost.
- ϕ, k_h and k_v fields, power law rel. perm. paramters, initial depths of fluid contacts.

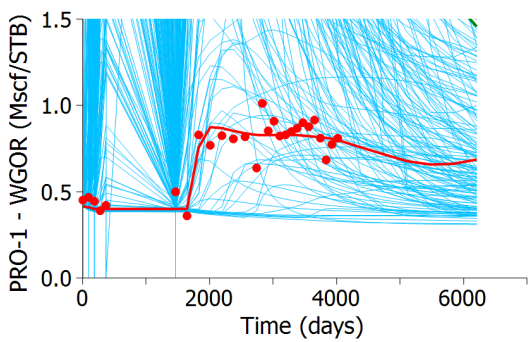


Well locations and the true horizontal permeability field, layer 3.

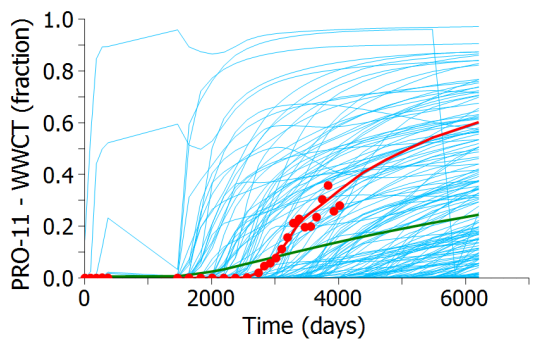
Prior Realizations



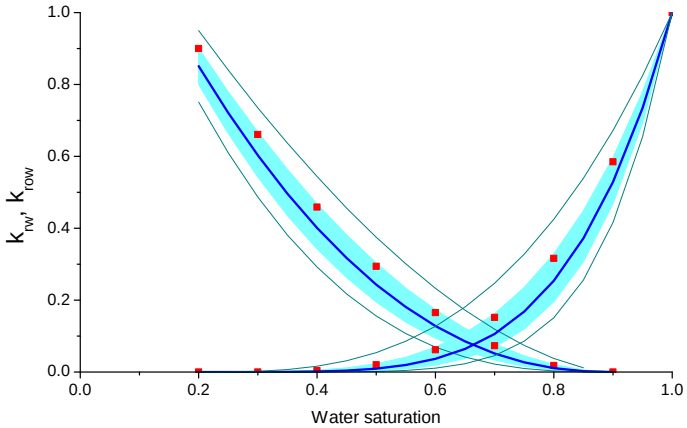
(a) PRO-15 BHP



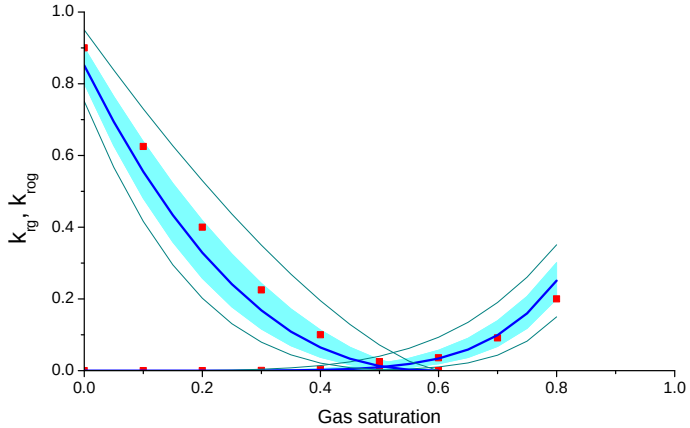
(b) PRO-1 GOR



(c) PRO-11 WCT

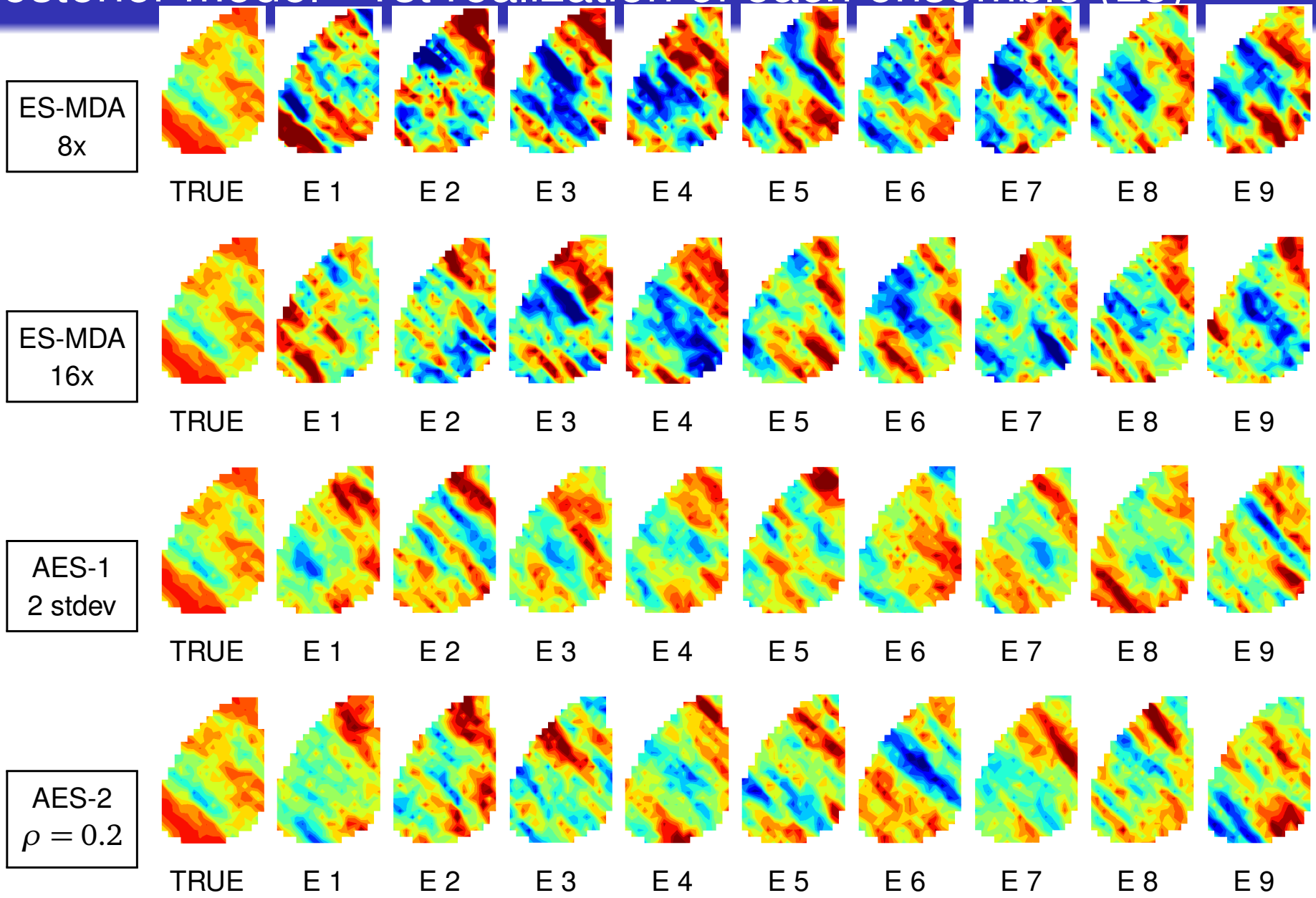


(a) k_{rw} and k_{row}

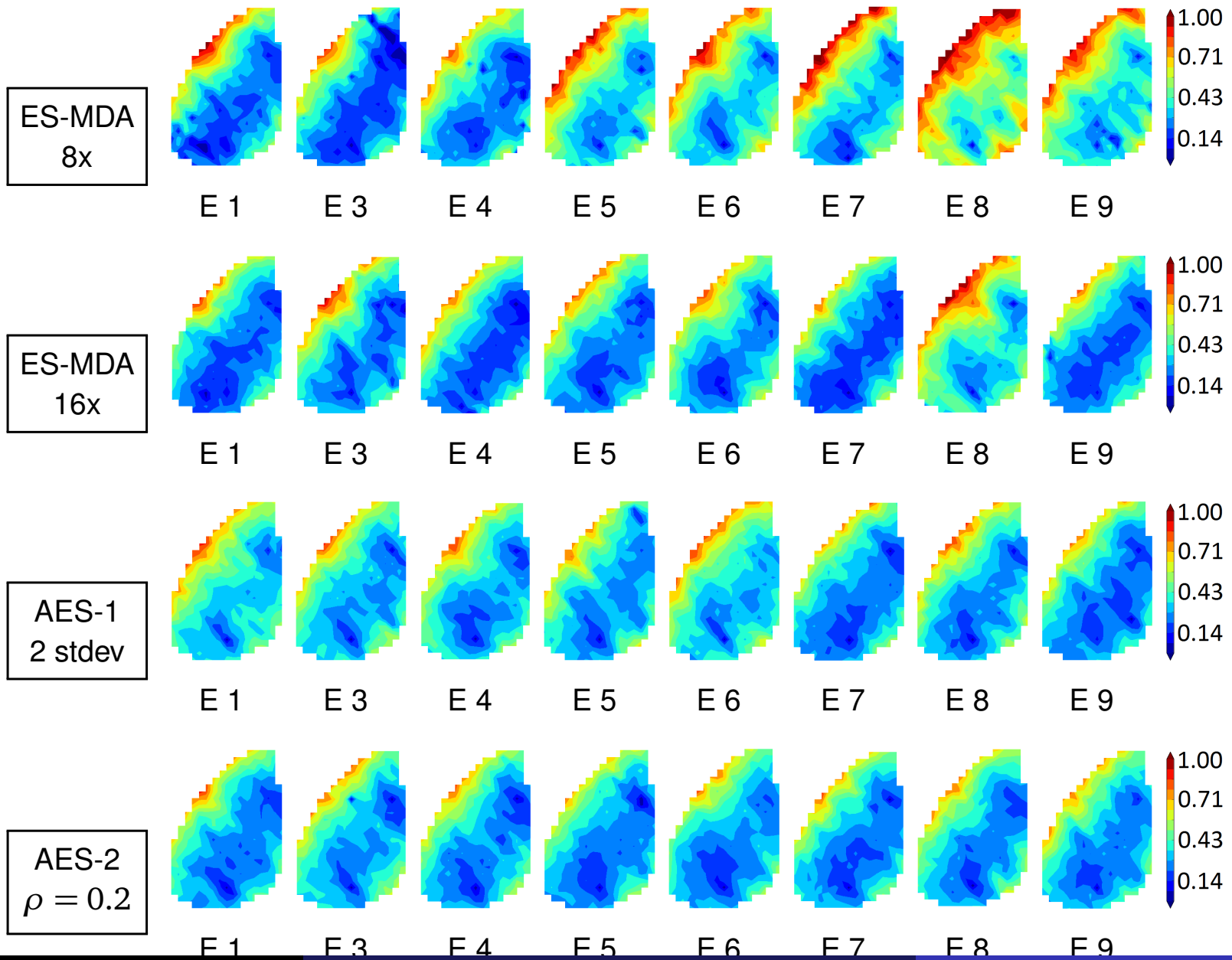


(b) k_{rg} and k_{rog}

Posterior model - 1st realization of each ensemble (L3)

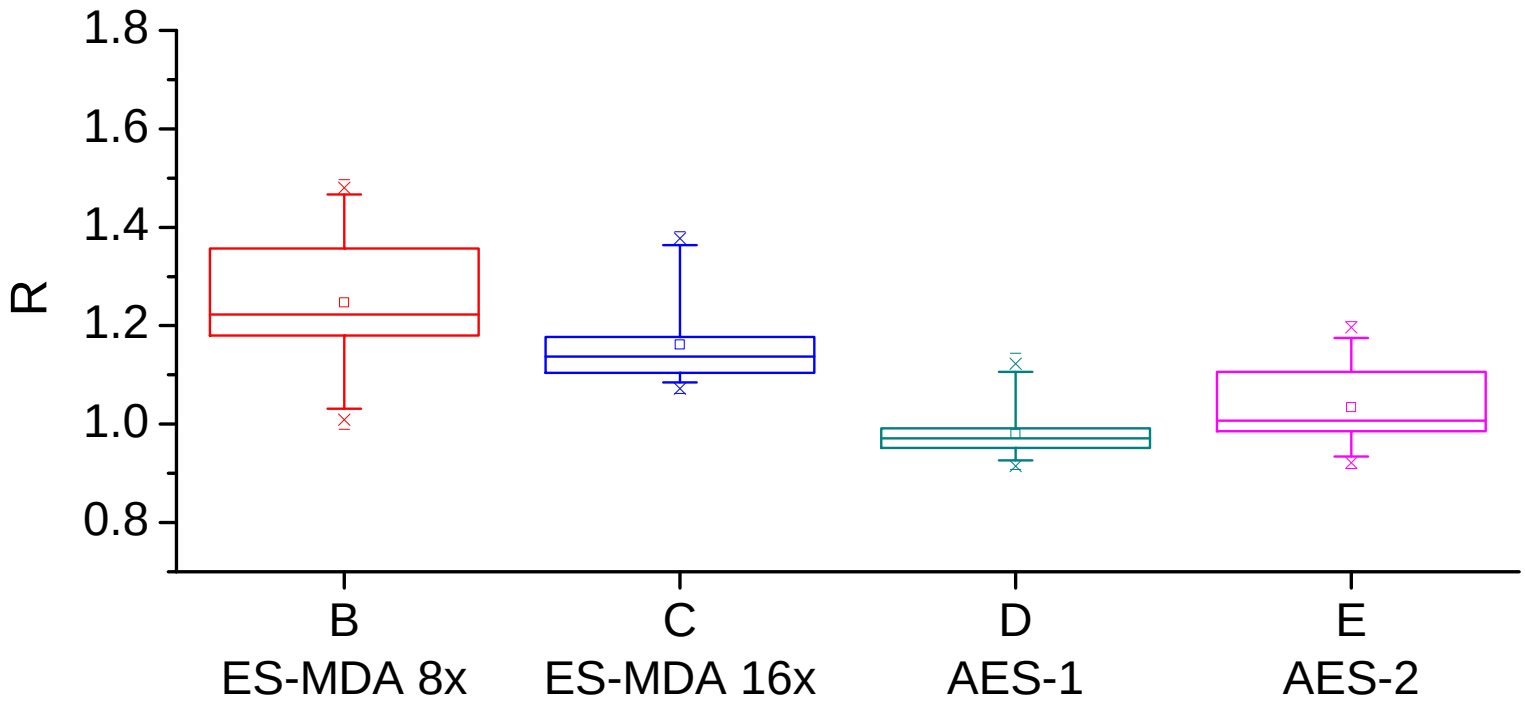


Posterior model - Standard deviation (L3)

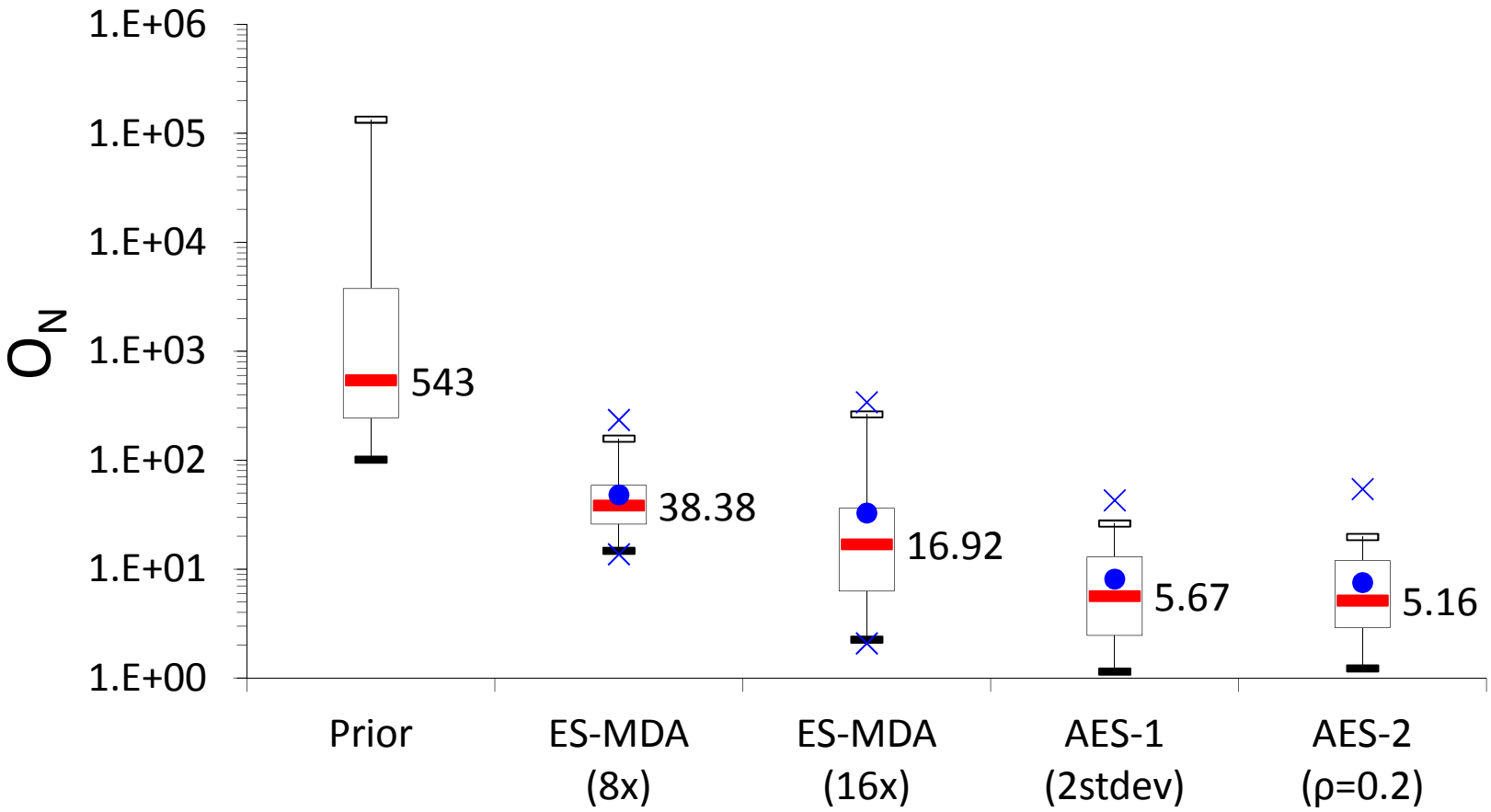


Difference between posterior realizations and true model

$$R_j = \frac{1}{N_m} ||C_M^{-1/2}(m_j - m_{true})||_1$$



Normalized objective function



Two Profound Questions

- Should we really restrict methods for sampling in the nonlinear-non-Gaussian case to those that sample correctly for the linear Gaussian case?
- Why do men's clothes have buttons on the right while women's clothes have buttons on the left?