Ensemble reservoir data assimilation with generic constraints

Xiaodong Luo, Norwegian Research Centre (NORCE), Norway

William C. Cruz, University of Stavanger, Norway

Data assimilation with generic constraints

- Physical constraints ubiquitous in practical data assimilation (DA) problems

- Approaches to handling generic constrained DA problems seemingly under-developed, especially in the context of ensemble DA. Some noticed problems include:

  - often considering equality or inequality constraints, but not both

  - difficulty in dealing with nonlinear constraints

  - applicability to large-scale problems

- The current work presenting a class of constrained ensemble DA algorithms with the potential to narrow the above noticed gaps

- These constrained DA algorithms derived from the generalized iterative ensemble smoother (GIES)

Overview of the generalized iterative ensemble smoother (GIES)

GIES: finding an ensemble of models $\{m_j^a\}$ that solves the following generalized minimum-average-cost (GMAC) problem:

$$min_{\{m_j^a\}} \frac{1}{N_e} \sum_j L(m_j^a), j = 1, 2, \dots, N_e$$

$$L(m) = D[T(d^o) - T(g(m))] + \gamma R[\Gamma(m) - \Gamma(m^b)]$$

Data mismatch term

Regularization term

Overview of a generalized iterative ensemble smoother (GIES)

Umbrella update formula of the GIES:

$$\mathrm{m}_{\mathrm{j}}^{\mathrm{a}} = m_j^{\mathrm{b}} + S_m \left(M_D(\bar{m}^b) + \gamma M_R(m_j^{\mathrm{b}}, \bar{m}^b) \right)^{-1} S_{T \circ g}^T \nabla_D [T(d^o) - T(g(m_j^b))]$$

Details of the GIES available in our previous work*:

*Luo, X. (2021). Novel iterative ensemble smoothers derived from a class of generalized cost functions. *Computational Geosciences*, *25*(3), 1159-1189.

GIES extended to other problems

- Allowing us to tackle certain problems that were previously cumbersome, if not impossible, to handle

- Example: data assimilation with soft constraints (DASC)

GIES for DASC problems

Problem statement

- Available sources of information:
  - Reservoir simulator: $d^{sim} = g(m)$ for a reservoir model m
  - Equality constraints: $f_{eq}(m) = 0$
  - Inequality constraints: $h_{in}(m) \leq 0$

- Constraints not necessarily strictly satisfied (hence the name “soft constraints”)

GIES for DASC problems

DASC as an optimization problem

- In the DASC problem, constraints incorporated in the data mismatch term, e.g.,

$$D\left[T(d^o) - T\left(g(m)\right)\right] = \frac{1}{2}\left(d^o - g(m)\right)^T C_d^{-1}\left(d^o - g(m)\right) + \alpha\, D_{eq}\left(0 - f_{eq}(m)\right) + \beta\, D_{in}(0 - h_{in}(m))$$

  - D_{eq} and D_{in}: distance metrics for equality and inequality constraints, respectively
  - α and β: relative weights

- Regularization term $R[\Gamma(m) - \Gamma(m^b)] = \frac{1}{2}\left(m - m^b\right)^T C_m^{-1}(m - m^b)$

GIES for DASC problems

DASC as an optimization problem

Applying the umbrella update formula of the GIES to the above choices:

$$\mathrm{m}_{\mathrm{j}}^{\mathrm{a}} = m_j^{\mathrm{b}} + K\left(S_g^T C_d^{-1}\left(d^o - g\left(m_j^{\mathrm{b}}\right)\right) + \alpha S_{\mathrm{f_{eq}}}^T \nabla_{D_{eq}}\left(0 - f_{eq}\left(m_j^{\mathrm{b}}\right)\right) + \beta S_{h_{in}}^T \nabla_{D_{in}}\left(0 - h_{in}\left(m_j^{\mathrm{b}}\right)\right)\right)$$

$$K \equiv S_m\left(S_{\mathrm{g}}^T\, C_d^{-1}\, S_{\mathrm{g}} + \alpha S_{\mathrm{f_{eq}}}^T \nabla_{D_{eq}}^2\left[0 - f_{eq}\left(\bar{m}^b\right)\right] S_{\mathrm{f_{eq}}} + \beta S_{h_{in}}^T \nabla_{D_{in}}^2\left[0 - h_{in}\left(\bar{m}^b\right)\right] S_{\mathrm{h_{in}}} + \gamma I\right)^{-1}$$

- Red: impact of equality constraints on model update
- Green: impact of inequality constraints on model update
- $\alpha = \beta = 0 \Rightarrow$ original IES algorithm

Referred to as GIES-DASC algorithm hereafter*

*Luo, X., & Cruz, W. C. (2022). Data assimilation with soft constraints (DASC) through a generalized iterative ensemble smoother. *Computational Geosciences*, *26*(3), 571-594.

GIES for DASC problems

Features of the GIES-DASC algorithm

$$\mathrm{m}_{\mathrm{j}}^{\mathrm{a}} = m_j^{\mathrm{b}} + K\left(S_g^T C_d^{-1}\left(d^o - g\left(m_j^{\mathrm{b}}\right)\right) + \alpha S_{\mathrm{f_{eq}}}^T \nabla_{D_{eq}}\left(0 - f_{eq}\left(m_j^{\mathrm{b}}\right)\right) + \beta S_{h_{in}}^T \nabla_{D_{in}}\left(0 - h_{in}\left(m_j^{\mathrm{b}}\right)\right)\right)$$

$$K \equiv S_m\left(S_{\mathrm{g}}^T\, C_d^{-1}\, S_{\mathrm{g}} + \alpha S_{\mathrm{f_{eq}}}^T \nabla_{D_{eq}}^2\left[0 - f_{eq}\left(\overline{m}^b\right)\right] S_{\mathrm{f_{eq}}} + \beta S_{h_{in}}^T \nabla_{D_{in}}^2\left[0 - h_{in}\left(\overline{m}^b\right)\right] S_{\mathrm{h_{in}}} + \gamma I\right)^{-1}$$

- Closed-form update formula, bearing a similar structure to the original IES algorithm
- Able to simultaneously handle nonlinear equality and inequality constraints in general
- Derivative-free with respect to the constraint-systems (i.e., no gradient of f_{eq} or h_{in} with respect to m)
- User-defined distance metrics D_{eq} and D_{in} => $\nabla_{D_{eq}}$, $\nabla_{D_{in}}$, $\nabla_{D_{eq}}^2$ and $\nabla_{D_{in}}^2$ having known analytical forms

GIES for DASC problems

Localization and handling big model/data size in the GIES-DASC algorithm

$$\mathrm{m}_{\mathrm{j}}^{\mathrm{a}} = m_j^{\mathrm{b}} + K\left(S_g^T C_d^{-1}\left(d^o - g\left(m_j^{\mathrm{b}}\right)\right) + \alpha S_{\mathrm{f_{eq}}}^T \nabla_{D_{eq}}\left(0 - f_{eq}\left(m_j^{\mathrm{b}}\right)\right) + \beta S_{h_{in}}^T \nabla_{D_{in}}\left(0 - h_{in}\left(m_j^{\mathrm{b}}\right)\right)\right)$$

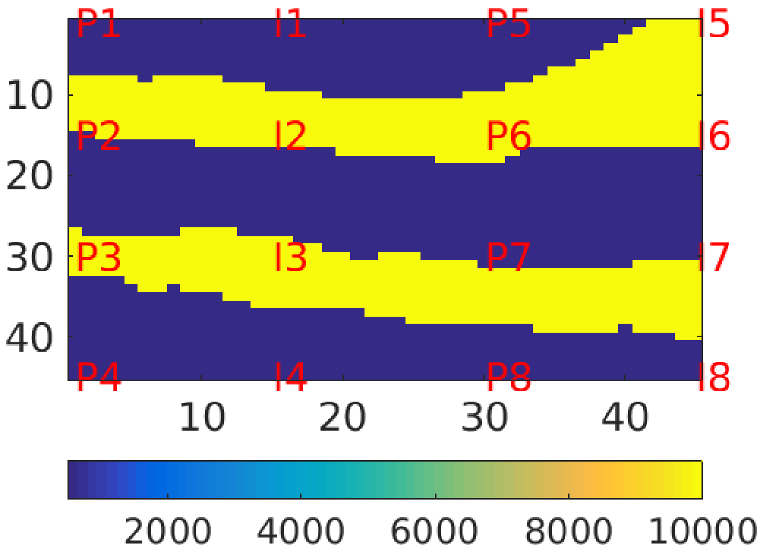
$$K \equiv S_m\left(S_{\mathrm{g}}^T\, C_d^{-1}\, S_{\mathrm{g}} + \alpha S_{\mathrm{f_{eq}}}^T \nabla_{D_{eq}}^2\left[0 - f_{eq}\left(\overline{m}^b\right)\right] S_{\mathrm{f_{eq}}} + \beta S_{h_{in}}^T \nabla_{D_{in}}^2\left[0 - h_{in}\left(\overline{m}^b\right)\right] S_{\mathrm{h_{in}}} + \gamma I\right)^{-1}$$

- Correlation-based localization applied to innovation/gradient projected onto the ensemble sub-space
- Proper choice of D_{eq} and D_{in} => diagonal Hessian matrices $\nabla_{D_{eq}}^2$ and $\nabla_{D_{in}}^2$ (useful for large-scale problems)
- More details available in:

Luo, X., & Cruz, W. C. (2022). Data assimilation with soft constraints (DASC) through a generalized iterative ensemble smoother. Computational Geosciences, 26(3), 571-594.

2D case study

Reference model (truth)



Experimental settings	
Model information	45 x 45 (two phases: oil and water); 8 producers (control mode: fluid rates) + 8 injectors (control mode: fluid rates) Uncertain parameters: PERMX
Reference model	PERMX: 500md (shale) and 10000 md (sand)
Production data used for history matching (history)	Oil and water rates from 8 producers + BHP from 8 injectors; History period: 0 – 1900 days
Production data used for cross-validation (forecast)	Oil and water rates from 8 producers + BHP from 8 injectors; Forecast period: 1900 – 3800 days
HM algorithm	Ensemble size: 100 Ordinary IES vs. GIES for DASC problems Correlation based adaptive localization

EAGE

2D case study

- Inequality-constraints: 100 md <= PERMX <= 15000 md on each active gridblock

  NB: $h_{in}(m) = [100\ - \mathrm{m}; \mathrm{m}\ - 15000] \leq 0$

- Choice of distance metric (barrier function):

  $$D_{in}(\boldsymbol{x}) = -\log(\boldsymbol{x})^T \mathbf{1} \text{ at } \boldsymbol{x} = \mathbf{0}\ - \boldsymbol{h_{in}}(\boldsymbol{m})$$

  NB: if $h_{in}(m)\ \rightarrow 0$, then $-\log\left(0 - h_{in}(m)\right)^T \mathbf{1}\ \rightarrow\ +\infty$

- The gradient $\nabla_{D_{in}}(x)$ and the Hessian $\nabla^2_{D_{in}}(x)$ having analytic forms*

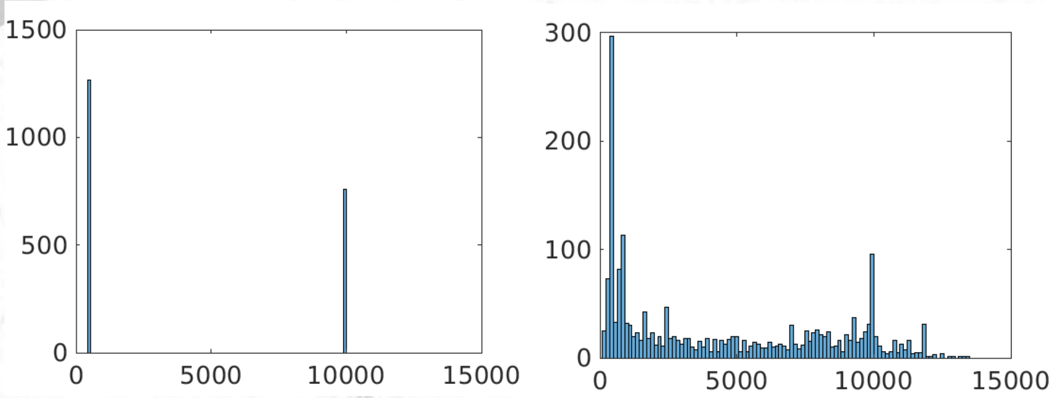
- Diagonal Hessian $\nabla^2_{D_{in}}(x)$, useful for large-scale problems (as in the 3D case later)*

*Luo, X., & Cruz, W. C. (2022). Data assimilation with soft constraints (DASC) through a generalized iterative ensemble smoother. *Computational Geosciences*, *26*(3), 571-594.

EAGE

2D case study

$f_{eq}(m)$ computes the differences between the histogram of the ground truth and that of an estimated reservoir model, bin by bin



- Equality-constraint system: see left-hand side

- Choice of distance metric ("channel" function):

  $D_{eq}(\boldsymbol{x}) = \log(|\boldsymbol{x}|)^T \mathbf{1}$ at $\boldsymbol{x} = \mathbf{0} - \boldsymbol{f}_{\boldsymbol{eq}}(\boldsymbol{m})$

  NB: $\log\left(|0 - f_{eq}(m)|\right)^T \mathbf{1} \rightarrow -\infty$ if $f_{eq}(m) \rightarrow 0$

- No need to evaluate the gradient of a histogram w.r.t PERMX

- The gradient $\nabla_{D_{eq}}(x)$ and the Hessian $\nabla^2_{D_{eq}}(x)$ having analytic forms*

- Diagonal Hessian $\nabla^2_{D_{eq}}(x)$

*Luo, X., & Cruz, W. C. (2022). Data assimilation with soft constraints (DASC) through a generalized iterative ensemble smoother. *Computational Geosciences*, *26*(3), 571-594.

2D case study

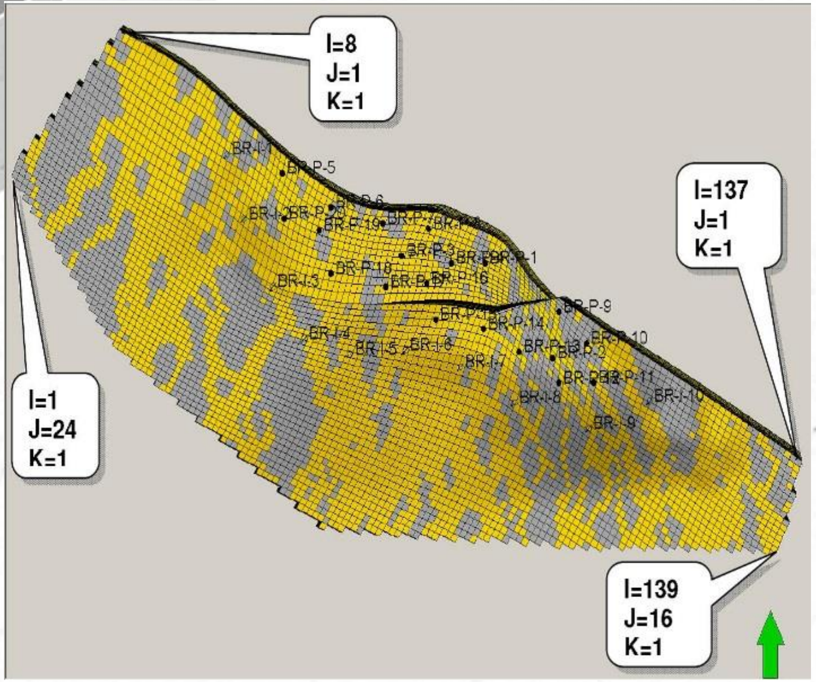
Table 1: Performance of IES algorithms in the 2D case study, in terms of data mismatch (mean ± STD) during history matching and forecast periods, and RMSE (mean ± STD) with respect to the ensembles of reservoir models at the initial or final iteration step(s).

	History-matching data mismatch (mean ± STD) $\times 10^3$	Forecast data mismatch (mean ± STD) $\times 10^3$	RMSE of PERMX (mean ± STD) $\times 10^3$	Values of (w_1, w_2)
Initial ensemble	5.7936 ± 2.2513	7.5940 ± 3.7789	5.2665 ± 0.4404	(N/A, N/A)
O-IES	2.2367 ± 0.8683	5.2670 ± 2.5087	4.5398 ± 0.3245	(0, 0)
C-GIES-EQ	2.3015 ± 0.9678	4.5766 ± 2.2940	4.4773 ± 0.3036	(1, 0)
C-GIES-IN	2.5579 ± 1.0900	5.1431 ± 2.1756	4.4380 ± 0.2916	(0,1)
C-GIES-(IN+EQ)	2.1575 ± 0.6781	4.0826 ± 1.2078	3.9053 ± 0.1863	(0.5, 0.5)

Nomenclature

- **O-IES**: Original IES
- **C-GIES-EQ**: GIES-DASC algorithm with only equality constraint(s)
- **C-GIES-IN**: GIES-DASC algorithm with only inequality constraint(s)
- **C-GIES-(IN+EQ):** GIES-DASC algorithm with both equality and inequality constraints

3D case study



Grid geometry of the Brugge field

Experimental settings	
Model size	139x48x9, with 44550 out of 60048 being active gridcells
Parameters to estimate	PORO, PERMX, PERMY, PERMZ. Total number is 4x44550 = 178,200
Production data (~10 yrs)	BHP, OPR, WCT. Total number is 1400
Constraint system	Upper and lower bounds for each parameter (as in the 2D case). Dimension of the constraint system = 2 x 178,200 = 356,4000
History matching algorithm	Ordinary IES vs. GIES for DASC problems Correlation based adaptive localization

3D case study

Table 2: Performance of the IES algorithms in the Brugge case study, in terms of data mismatch and RMSE (mean ± STD).

	Initial ensemble	O-IES	C-GIES-IN
Data mismatch	$3.6232 \times 10^9 \pm 1.4900 \times 10^{10}$	$(3.9616 \pm 2.9947) \times 10^7$	$(7.0091 \pm 5.5507) \times 10^6$
RMSE (PERMX)	1.6585 ± 0.3827	1.4167 ± 0.2545	1.4119 ± 0.2284
RMSE (PERMY)	1.6612 ± 0.3794	1.4198 ± 0.2515	1.4133 ± 0.2244
RMSE (PERMZ)	2.0077 ± 0.4096	1.8054 ± 0.3101	1.7636 ± 0.2916
RMSE (PORO)	0.0302 ± 0.0033	0.0280 ± 0.0025	0.0285 ± 0.0028
RMSE (all together)	1.5450 ± 0.3362	1.3498 ± 0.2344	1.3327 ± 0.2103

Nomenclature

- **O-IES**: Original IES
- **C-GIES-IN**: GIES-DASC algorithm with only inequality constraint(s)

EAGE

Discussion and conclusion

- A class of ensemble DASC algorithms obtained as a special case of the umbrella GIES update formula

- Features of the GIES-DASC algorithm
  - closed-form and close to IES
  - simultaneously handling nonlinear equality and inequality constraints
  - derivative-free
  - applicable to large-scale problems

- Better data assimilation performance obtained by the GIES-DASC algorithm(s) in both case studies

EAGE

Acknowledgements / Thank You / Questions

XL acknowledges partial financial supports from both the NORCE research project “Assimilating 4D Seismic Data: Big Data Into Big Models” which is funded by industry partners, Equinor Energy AS, Lundin Energy Norway AS, Repsol Norge AS, Shell Global Solutions International B.V., TotalEnergies EP Norge AS, and Wintershall Dea Norge AS, as well as the Research Council of Norway (project number: 295002), and the National Centre for Sustainable Subsurface Utilization of the Norwegian Continental Shelf (NCS2030), which is funded by industry partners and the Research Council of Norway (project number: 331644).

WC acknowledges financial supports from the National IOR centre of Norway (RCN no. 230303), which is funded by the RCN and industry partners ConocoPhillips, Aker BP, Vår Energi, Equinor, Neptune Energy, Lundin, Halliburton, Schlumberger, and Wintershall Dea.

The authors would also like to thank Schlumberger for providing academic licenses to ECLIPSE©