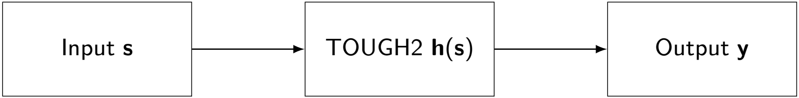
FAST LARGE-SCALE INVERSION FOR DEEP AQUIFER CHARACTERIZATION

Jonghyun Harry Lee
Amalia Kokkinaki
Judith Yue Li
Peter K. Kitanidis

Department of Civil and Environmental Engineering, Stanford University

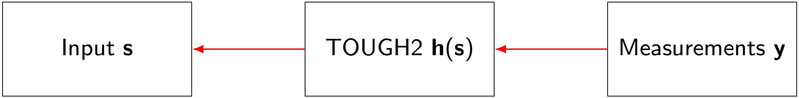
September 30, 2015
TOUGH Symposium 2015

Introduction : Inverse Problem



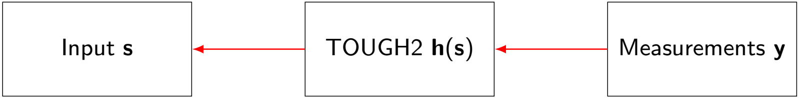
- Given perfect description on **s**, TOUGH2 predicts the state of the system **y**

Introduction : Inverse Problem



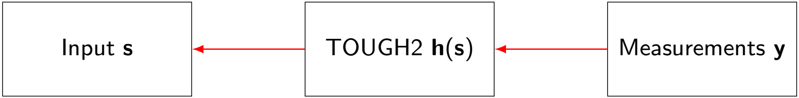
- Given perfect description on **s**, TOUGH2 predicts the state of the system **y**
- In practice, we use measurements of **y** to estimate **s**

Introduction : Inverse Problem



- Given perfect description on **s**, TOUGH2 predicts the state of the system **y**
- In practice, we use measurements of **y** to estimate **s**
- typically $n_{obs} \ll n_{unknowns}$
- Error in **y** and **h** leads to uncertainty in estimation **s**
- Therefore, **s** is characterized in a statistical framework

Introduction : Inverse Problem



- Given perfect description on **s**, TOUGH2 predicts the state of the system **y**
- In practice, we use measurements of **y** to estimate **s**
- typically $n_{obs} \ll n_{unknowns}$
- Error in **y** and **h** leads to uncertainty in estimation **s**
- Therefore, **s** is characterized in a statistical framework
- requires $\mathcal{O}(min(n_{obs}, n_{unknowns}))$ TOUGH2 runs (or more)

Geostatistical Approach [Kitanidis, 1995]

Consider the measurement equation

$$\mathbf{y} = \mathbf{h}(\mathbf{s}) + \mathbf{v} \qquad \mathbf{v} \sim \mathcal{N}(0, \mathbf{\Gamma}_{\text{noise}})$$

where,

$$\begin{aligned}
\mathbf{y} &:= n_{obs} \text{ measurements, e.g., pressure or temperature} \\
\mathbf{h}(\cdot) &:= \text{TOUGH2, e.g., TOUGH2-MP EOS1} \\
\mathbf{s} &:= n_{unknowns} \text{ model parameters, e.g., permeability} \\
\mathbf{v} &:= \text{measurement and model error}
\end{aligned}$$

Geostatistical Approach [Kitanidis, 1995]

Consider the measurement equation

$$\mathbf{y} = \mathbf{h}(\mathbf{s}) + \mathbf{v} \qquad \mathbf{v} \sim \mathcal{N}(0, \mathbf{\Gamma}_{\text{noise}})$$

where,

$$\begin{aligned}
\mathbf{y} &:= n_{obs} \text{ measurements, e.g., pressure or temperature} \\
\mathbf{h}(\cdot) &:= \text{TOUGH2, e.g., TOUGH2-MP EOS1} \\
\mathbf{s} &:= n_{unknowns} \text{ model parameters, e.g., permeability} \\
\mathbf{v} &:= \text{measurement and model error}
\end{aligned}$$

Assuming prior info on $\mathbf{s} \sim \mathcal{N}(\mathbf{s}_{prior}, \mathbf{\Gamma}_{\text{prior}})$, the probabilistic distribution of $\mathbf{s}$

Geostatistical Approach [Kitanidis, 1995]

Consider the measurement equation

$$\mathbf{y} = \mathbf{h}(\mathbf{s}) + \mathbf{v} \qquad \mathbf{v} \sim \mathcal{N}(0, \mathbf{\Gamma}_{\text{noise}})$$

where,

$$\begin{aligned} \mathbf{y} &:= n_{obs} \text{ measurements, e.g., pressure or temperature} \\ \mathbf{h}(\cdot) &:= \text{TOUGH2, e.g., TOUGH2-MP EOS1} \\ \mathbf{s} &:= n_{unknowns} \text{ model parameters, e.g., permeability} \\ \mathbf{v} &:= \text{measurement and model error} \end{aligned}$$

Assuming prior info on $\mathbf{s} \sim \mathcal{N}(\mathbf{s}_{prior}, \mathbf{\Gamma}_{\text{prior}})$, the probabilistic distribution of $\mathbf{s}$

$$p(\mathbf{s}) \sim \exp\left(\underbrace{-\frac{1}{2}(\mathbf{y}-\mathbf{h}(\mathbf{s}))^\top \mathbf{\Gamma}_{\text{prior}}^{-1}(\mathbf{y}-\mathbf{h}(\mathbf{s}))}_{likelihood} \underbrace{-\frac{1}{2}(\mathbf{s}-\mathbf{s}_{prior})^\top \mathbf{\Gamma}_{\text{prior}}^{-1}(\mathbf{s}-\mathbf{s}_{prior})}_{prior} \right)$$

Geostatistical Approach [Kitanidis, 1995]

Consider the measurement equation

$$\mathbf{y} = \mathbf{h}(\mathbf{s}) + \mathbf{v} \qquad \mathbf{v} \sim \mathcal{N}(0, \mathbf{\Gamma}_{\text{noise}})$$

where,

$$\begin{aligned}
\mathbf{y} &:= n_{obs} \text{ measurements, e.g., pressure or temperature} \\
\mathbf{h}(\cdot) &:= \text{TOUGH2, e.g., TOUGH2-MP EOS1} \\
\mathbf{s} &:= n_{unknowns} \text{ model parameters, e.g., permeability} \\
\mathbf{v} &:= \text{measurement and model error}
\end{aligned}$$

Assuming prior info on $\mathbf{s} \sim \mathcal{N}(\mathbf{s}_{prior}, \mathbf{\Gamma}_{\text{prior}})$, the probabilistic distribution of $\mathbf{s}$

$$p(\mathbf{s}) \sim \exp\left(\underbrace{-\frac{1}{2}(\mathbf{y} - \mathbf{h}(\mathbf{s}))^\top \mathbf{\Gamma}_{\text{prior}}^{-1} (\mathbf{y} - \mathbf{h}(\mathbf{s}))}_{likelihood} \underbrace{-\frac{1}{2}(\mathbf{s} - \mathbf{s}_{prior})^\top \mathbf{\Gamma}_{\text{prior}}^{-1} (\mathbf{s} - \mathbf{s}_{prior})}_{prior} \right)$$

and the best estimate is given by

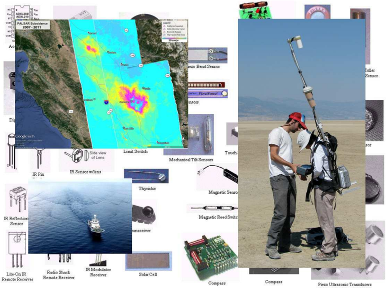
$$\hat{\mathbf{s}} = \mathbf{s}_{prior} + \mathbf{\Gamma}_{\text{prior}} \mathbf{J}^\top \xi$$

where $\mathbf{J}$ is Jacobian (sensitivity) matrix

$$\mathbf{J} = \frac{\partial \mathbf{h}}{\partial \mathbf{s}}_{\mathbf{s}=\hat{\mathbf{s}}}$$

Computational Challenges in Inverse Modeling

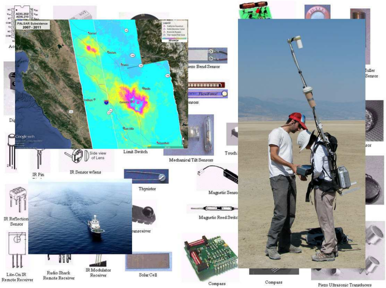
With advances in sensor technology and computational power (e.g.,$\min(n_{obs}, n_{unknowns}) \geq 10^6$), we need a scalable approach!



images adapted from USGS, http://cs.txstate.edu/∼xc10 and CERN

Computational Challenges in Inverse Modeling

With advances in sensor technology and computational power (e.g.,$\min(n_{obs}, n_{unknowns}) \geq 10^6$), we need a scalable approach!



In geostatistical approach (as well as other approaches), computational costs from

- Jacobian-Covariance products (e.g., prior cross-covariance) $\mathbf{J}\boldsymbol{\Gamma}_{\text{prior}}$
- $\mathcal{O}(\min(n_{obs}, n_{unknowns}))$ forward model runs for $\mathbf{J}$
- $\mathcal{O}(n_{unknowns}^2)$ storage for $\boldsymbol{\Gamma}_{\text{prior}}$
- $\mathcal{O}(n_{unknowns}^2 n_{obs})$ multiplication for $\mathbf{J}\boldsymbol{\Gamma}_{\text{prior}}$

images adapted from USGS, http://cs.txstate.edu/∼xc10 and CERN

Computational Challenges in Inverse Modeling

For example, $n_{unknowns} = 6$ mil. and $n_{obs} = 100,000$, how can we compute

$$\mathbf{J}, \quad \mathbf{\Gamma}_{\text{prior}}$$

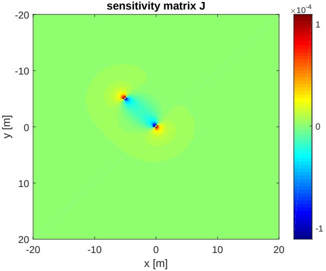
efficiently?

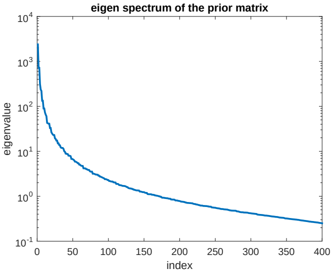
Computational Challenges in Inverse Modeling

For example, $n_{unknowns} = 6$ mil. and $n_{obs} = 100,000$, how can we compute

$$\mathbf{J}, \quad \mathbf{\Gamma}_{\text{prior}}$$

efficiently?





Computational Challenges in Inverse Modeling

For example, $n_{unknowns} = 6$ mil. and $n_{obs} = 100,000$, how can we compute

$$\mathbf{J}\Gamma_{\text{prior}}$$

efficiently?

Principal Component Geostatistical Approach [Lee and Kitanidis, 2014]

$$\mathbf{J\Gamma}_{prior}$$

Principal Component Geostatistical Approach [Lee and Kitanidis, 2014]

$$\mathbf{J}\boldsymbol{\Gamma}_{prior}$$

To avoid explicit construction of $\mathbf{J}$ and $\boldsymbol{\Gamma}$, we use a Jacobian-free algorithm:

Principal Component Geostatistical Approach [Lee and Kitanidis, 2014]

$$\mathbf{J}\boldsymbol{\Gamma}_{prior}$$

To avoid explicit construction of **J** and **Γ**, we use a Jacobian-free algorithm:

$$\boldsymbol{\Gamma}_{prior} \quad = \quad \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T \approx \mathbf{U}_{1:K}\boldsymbol{\Sigma}_{1:K}\mathbf{U}_{1:K}^T = \sum_{i=1}^{K} \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^\top$$

Principal Component Geostatistical Approach [Lee and Kitanidis, 2014]

$$\mathbf{J}\boldsymbol{\Gamma}_{prior}$$

To avoid explicit construction of **J** and **Γ**, we use a Jacobian-free algorithm:

$$\begin{aligned}\boldsymbol{\Gamma}_{prior} &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T \approx \mathbf{U}_{1:K}\boldsymbol{\Sigma}_{1:K}\mathbf{U}_{1:K}^T = \sum_{i=1}^{K} \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^\top \\ \mathbf{J}\boldsymbol{\Gamma}_{prior} &\approx \mathbf{J}\sum_{i=1}^{K} \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^\top = \sum_{i=1}^{K} \left(\mathbf{J}\boldsymbol{\zeta}_i\right) \boldsymbol{\zeta}_i^\top\end{aligned}$$

$$\mathbf{J}\mathbf{\Gamma}_{prior}$$

To avoid explicit construction of **J** and **Γ**, we use a Jacobian-free algorithm:

$$\begin{aligned}
\mathbf{\Gamma}_{prior} &= \mathbf{U}\mathbf{\Sigma}\mathbf{U}^T \approx \mathbf{U}_{1:K}\mathbf{\Sigma}_{1:K}\mathbf{U}_{1:K}^T = \sum_{i=1}^{K} \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^\top \\
\mathbf{J}\mathbf{\Gamma}_{prior} &\approx \mathbf{J}\sum_{i=1}^{K} \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^\top = \sum_{i=1}^{K} \left(\mathbf{J}\boldsymbol{\zeta}_i\right) \boldsymbol{\zeta}_i^\top \\
\mathbf{J}\boldsymbol{\zeta}_i &= \frac{h(\mathbf{s}+\delta\boldsymbol{\zeta}_i) - h(\mathbf{s})}{\delta} + O(\delta)
\end{aligned}$$

Principal Component Geostatistical Approach [Lee and Kitanidis, 2014]

$$\mathbf{J}\boldsymbol{\Gamma}_{prior}$$

To avoid explicit construction of $\mathbf{J}$ and $\boldsymbol{\Gamma}$, we use a Jacobian-free algorithm:

$$\begin{aligned}
\boldsymbol{\Gamma}_{prior} &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T \approx \mathbf{U}_{1:K}\boldsymbol{\Sigma}_{1:K}\mathbf{U}_{1:K}^T = \sum_{i=1}^{K} \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^\top \\
\mathbf{J}\boldsymbol{\Gamma}_{prior} &\approx \mathbf{J}\sum_{i=1}^{K} \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^\top = \sum_{i=1}^{K} (\mathbf{J}\boldsymbol{\zeta}_i)\, \boldsymbol{\zeta}_i^\top \\
\mathbf{J}\boldsymbol{\zeta}_i &= \frac{h(\mathbf{s} + \delta\boldsymbol{\zeta}_i) - h(\mathbf{s})}{\delta} + O(\delta)
\end{aligned}$$

- only $K + 1$ TOUGH2 executions in each iteration; total runs $\sim O(100)$ in most cases

Principal Component Geostatistical Approach [Lee and Kitanidis, 2014]

$$\mathbf{J}\boldsymbol{\Gamma}_{prior}$$

To avoid explicit construction of $\mathbf{J}$ and $\boldsymbol{\Gamma}$, we use a Jacobian-free algorithm:

$$\begin{aligned}
\boldsymbol{\Gamma}_{prior} &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T \approx \mathbf{U}_{1:K}\boldsymbol{\Sigma}_{1:K}\mathbf{U}_{1:K}^T = \sum_{i=1}^{K} \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^\top \\
\mathbf{J}\boldsymbol{\Gamma}_{prior} &\approx \mathbf{J}\sum_{i=1}^{K} \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^\top = \sum_{i=1}^{K} (\mathbf{J}\boldsymbol{\zeta}_i)\, \boldsymbol{\zeta}_i^\top \\
\mathbf{J}\boldsymbol{\zeta}_i &= \frac{h(\mathbf{s}+\delta\boldsymbol{\zeta}_i) - h(\mathbf{s})}{\delta} + O(\delta)
\end{aligned}$$

- only $K+1$ TOUGH2 executions in each iteration; total runs $\sim O(100)$ in most cases
- use TOUGH2 as a black-box; easy to implement without code modification!

Principal Component Geostatistical Approach [Lee and Kitanidis, 2014]

$$\mathbf{J}\boldsymbol{\Gamma}_{prior}$$

To avoid explicit construction of **J** and **Γ**, we use a Jacobian-free algorithm:

$$\begin{aligned}
\boldsymbol{\Gamma}_{prior} &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T \approx \mathbf{U}_{1:K}\boldsymbol{\Sigma}_{1:K}\mathbf{U}_{1:K}^T = \sum_{i=1}^{K} \zeta_i \zeta_i^\top \\
\mathbf{J}\boldsymbol{\Gamma}_{prior} &\approx \mathbf{J}\sum_{i=1}^{K} \zeta_i \zeta_i^\top = \sum_{i=1}^{K} (\mathbf{J}\zeta_i)\, \zeta_i^\top \\
\mathbf{J}\zeta_i &= \frac{h(\mathbf{s}+\delta\zeta_i) - h(\mathbf{s})}{\delta} + O(\delta)
\end{aligned}$$

- only $K+1$ TOUGH2 executions in each iteration; total runs $\sim O(100)$ in most cases
- use TOUGH2 as a black-box; easy to implement without code modification!
- easily parallelizable

Principal Component Geostatistical Approach [Lee and Kitanidis, 2014]

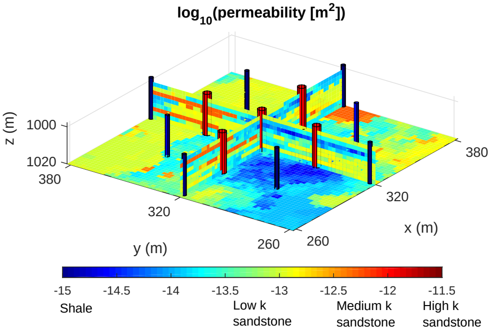
$$\mathbf{J}\boldsymbol{\Gamma}_{prior}$$

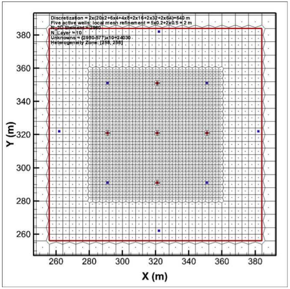
To avoid explicit construction of $\mathbf{J}$ and $\boldsymbol{\Gamma}$, we use a Jacobian-free algorithm:

$$\begin{aligned}
\boldsymbol{\Gamma}_{prior} &= \mathbf{U}\Sigma\mathbf{U}^T \approx \mathbf{U}_{1:K}\Sigma_{1:K}\mathbf{U}_{1:K}^T = \sum_{i=1}^{K} \boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^\top \\
\mathbf{J}\boldsymbol{\Gamma}_{prior} &\approx \mathbf{J}\sum_{i=1}^{K} \boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^\top = \sum_{i=1}^{K} \left(\mathbf{J}\boldsymbol{\zeta}_i\right)\boldsymbol{\zeta}_i^\top \\
\mathbf{J}\boldsymbol{\zeta}_i &= \frac{h(\mathbf{s}+\delta\boldsymbol{\zeta}_i) - h(\mathbf{s})}{\delta} + O(\delta)
\end{aligned}$$

- only $K+1$ TOUGH2 executions in each iteration; total runs $\sim O(100)$ in most cases
- use TOUGH2 as a black-box; easy to implement without code modification!
- easily parallelizable
- previously applied to 3-D hydraulic tomography ($n_{unknown} \sim 3$ mil.), MRI-imaged tracer data inversion ($n_{obs} \sim 6$ mil.), arsenic-bearing mineral characterization (flow - transport - multi-species reaction) and so on

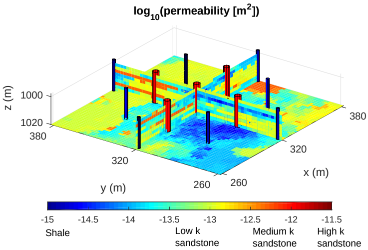
Application : aquifer characterization using CO_2 leakage monitoring network

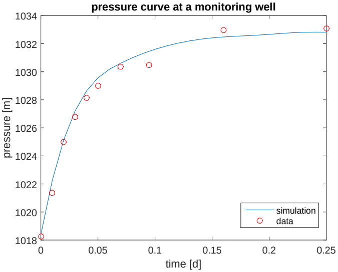




- 120 m x 120 m x 20 m permeability field estimation
- 640 m x 640 m x 20 m model domain
- 5 pumping & 8 monitoring wells
- "True" field from TPROGS [Carle and Fogg, 1996], $n_{unknowns} = 24{,}040$
- pressure and heat tracer data generated from TOUGH2-MP EOS1

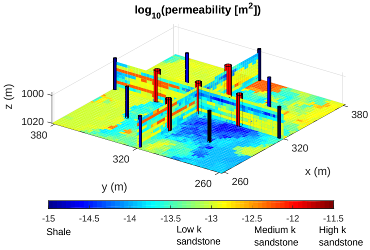
Pressure Data Inversion





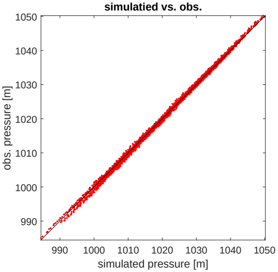
- 4 dipole pumping tests (extraction at the center well)
- $n_{pres.obs.} = 4{,}400$ (4 dipole tests x 11 wells x 10 ports x 10 transient pts)

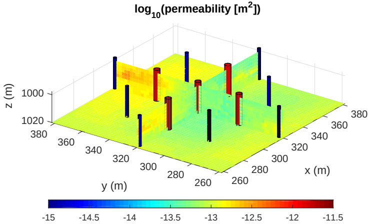
Pressure Data Inversion



- 4 dipole pumping tests (extraction at the center well)
- $n_{pres.obs.} = 4{,}400$ (4 dipole tests x 11 wells x 10 ports x 10 transient pts)
- $K = 100$ TOUGH2 runs, converged in 3 iterations
- total inversion $\sim$ 300 TOUGH2 runs for each test
- On a 16-core workstation, it took less than 2 hours.

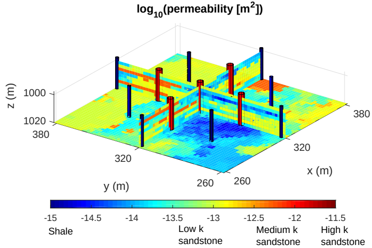
Pressure Data Inversion

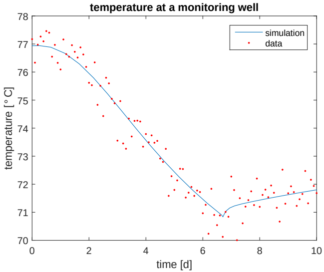




- 4 dipole pumping tests (extraction at the center well)
- $n_{pres.obs.} = 4{,}400$ (4 dipole tests x 11 wells x 10 ports x 10 transient pts)
- K = 100 TOUGH2 runs, converged in 3 iterations
- total inversion $\sim$ 300 TOUGH2 runs for each test
- On a 16-core workstation, it took less than 2 hours.

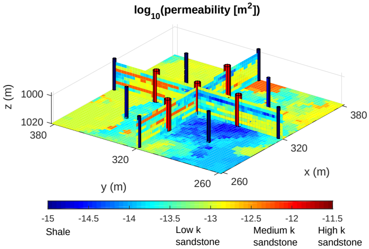
Pressure & Temperature Data Inversion

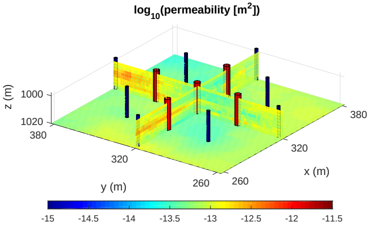




- inject cold water (by 20 °C) and extract in 4 wells for a week
- Here, we use "total heat" (zero-th moment of temperature) as measurements, $n_{temp.obs.} = 40$

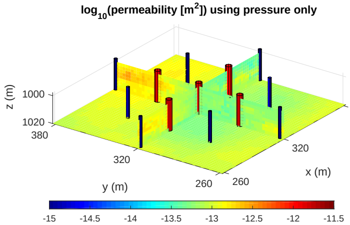
Pressure & Temperature Data Inversion

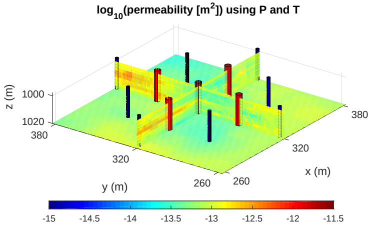




- inject cold water (by 20 °C) and extract in 4 wells for a week
- Here, we use "total heat" (zero-th moment of temperature) as measurements, $n_{temp.obs.} = 40$
- K = 100 principal components used, 4 iterations
- total inversion $\sim$ 400 TOUGH2 runs for each test
- On a 16-core workstation, it took 8 hours.

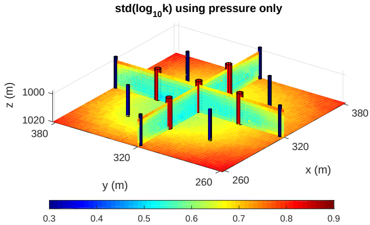
Pressure & Temperature Data Inversion





- inject cold water (by 20 °C) and extract in 4 wells for a week
- Here, we use "total heat" (zero-th moment of temperature) as measurements, $n_{temp.obs.} = 40$
- K = 100 principal components used, 4 iterations
- total inversion $\sim$ 400 TOUGH2 runs for each test
- On a 16-core workstation, it took 8 hours.

Pressure & Temperature Data Inversion



- inject cold water (by 20 °C) and extract in 4 wells for a week
- Here, we use "total heat" (zero-th moment of temperature) as measurements, $n_{temp.obs.} = 40$
- K = 100 principal components used, 4 iterations
- total inversion ∼ 400 TOUGH2 runs for each test
- On a 16-core workstation, it took 8 hours.
- Estimation uncertainty is reduced with additional temperature data

Conclusion

Contributions

- use of TOUGH2-MP as a black box to perform joint inversion
- total $\sim \mathcal{O}(100)$ TOUGH2-MP executions for complex inverse problems
- Potential benefits using temperature tracer data to improve the inversion in CO2 leakage monitoring network

References

- Lee and Kitanidis, Large-Scale Hydraulic Tomography and Joint Inversion of Head and Tracer Data using the Principal Component Geostatistical Approach (PCGA), *WRR*, 2014
- Kitanidis and Lee, Principal Component Geostatistical Approach for Large-Dimensional Inverse Problem, *WRR*, 2014
- Kitanidis, Quasi-linear Geostatistical Theory for Inversing, *WRR*, 1995
- Carles and Fogg, Transition Probability-based Indicator Geostatistics, *Math. Geology*, 1996

Acknowledgements

- This material is based upon work supported by US Department of Energy, National Energy Technology Laboratory (DOE, NETL) under the award DE-FE0009260: "An Advanced Joint Inversion System for CO2 Storage Modeling with Large Data Sets for Characterization and Real-Time Monitoring".
- We would like to thank Dr. Quanlin Zhou of LBNL for support on TOUGH2.