SAND2016-XXXX

SAND2016-4866C

# A Parallel MCMC Method


Laura Swiler and Jaideep Ray

Sandia National Laboratories, MS 1318
Albuquerque, NM 87185
lpswile@sandia.gov, jairay@sandia.gov


Center for Computing Research

*Exceptional service in the national interest*


Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

# Outline

- Parallel Chains:  Motivation and Issues
- Convergence metrics
  - Classic tests:  Raftery-Lewis, Gelman, Geweke
  - Proposed Bootstrap Metric
- Results
  - Correlated Gaussian
  - Rosenbrock
  - Hydrogeological inverse problem – Ground Penetrating Radar

# Bayesian Calibration

- Generate posterior distributions on model parameters, given
  - Experimental data
  - A prior distribution on model parameters
  - A presumed probabilistic relationship between experimental data and model output that can be defined by a likelihood function

$$\pi(\theta \mid d) \propto \pi(\theta) L(d \mid \theta)$$

Model parameters — $\theta$

Observed Data — $d$

Prior parameter distribution — $\pi(\theta)$

Likelihood function which incorporates the model discrepancy: we assume Gaussian — $L(d \mid \theta)$

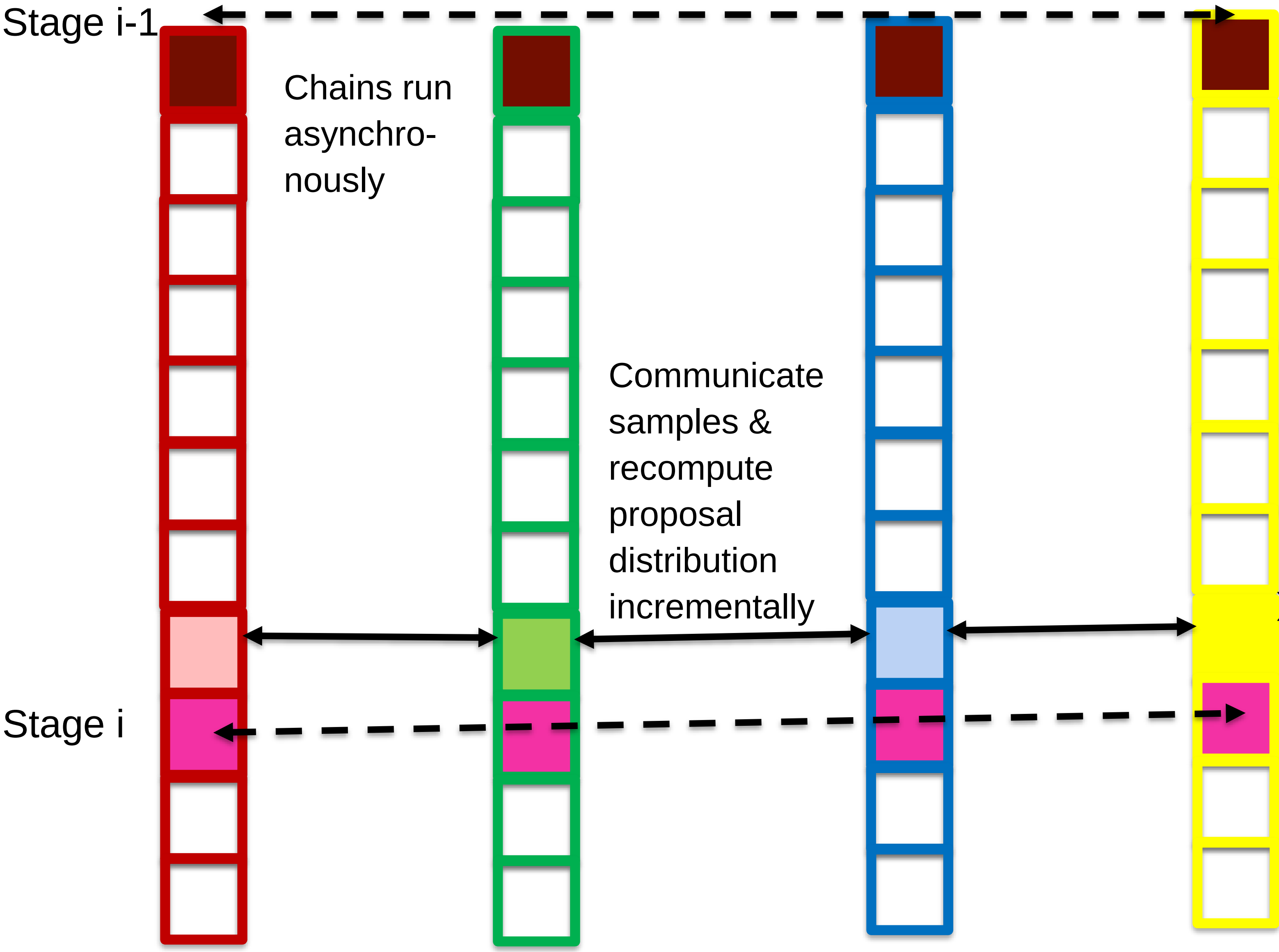
# Markov Chain Monte Carlo

- How do we obtain the posterior?
  - It is usually too difficult to calculate analytically
  - We use a technique called Markov Chain Monte Carlo (MCMC)
  - In MCMC, the idea is to *generate a sampling density that is approximately equal to the posterior distribution.*
  - Metropolis-Hastings is a commonly used algorithm
- MCMC depends on asymptotic behavior of the chain. Ideally, you want to run for 100,000+ samples.

  COMPUTATIONALLY VERY EXPENSIVE!
  - Typically, a limited number of model runs are used to generate a surrogate model and the MCMC sampling is performed on the surrogate
  - We want to avoid surrogates
- Limitation of MCMC: it is inherently sequential.
- We want to exploit some parallelism by using multiple chains

**SOLUTION: PARALLEL DRAM on the actual model**

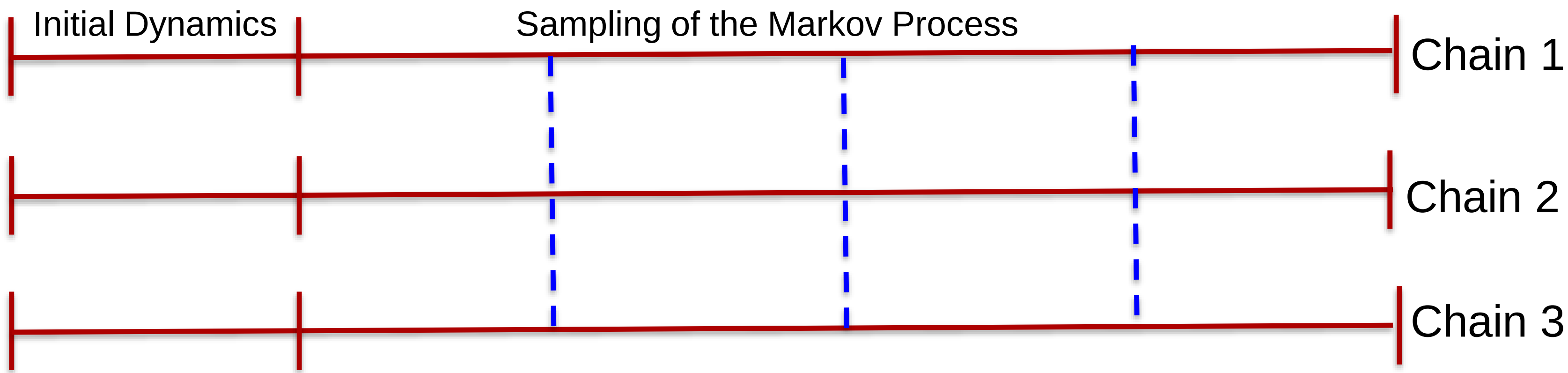
# Parallel DRAM




Full formula given in "Solonen et al. "Efficient MCMC for Climate Model Parameter Estimation: Parallel Adaptive Chains and Early Rejection. Bayesian Analysis (2012) 7(2), pp. 1-22.

$$\Sigma_i = \frac{(i-1)}{i}\,\Sigma_{i-1} + \frac{1}{n}\left[\theta_n - \bar{\theta}\right]^2$$

# Parallel MCMC




- Premise: Unless the posterior is pathological, most of the chain is spent in the sampling part.
  - If we need a number of samples Q to obtain statistical properties of the posterior, *m* chains should allow us to be approximately *m* times more efficient (e.g. split up the Q samples into Q/*m*).
- Issues of how to aggregate chains
  - Pooling acceptable for examination of statistics
  - Concern about Markovian properties

# Convergence Metrics

- There are a variety of metrics to test for chain convergence
- They measure different things:  be careful!
- Convergence metrics
  - Raftery-Lewis:  provide a way of bounding the variances of estimates of the quantiles.  R-L is posed as "what number of samples are needed to obtain a quantile q to within an accuracy of +/- r with probability p."
  - McGibbsit, by Gregory Warnes, is a multi-chain extension of Raftery-Lewis.
  - Gelman-Rubin: Start with multiple chains, starting at "overdispersed" starting points.  The "shrink" factor approaches 1 when the pooled within-chain variances dominates the between-chain variance, so that all chains have escaped the influence of their starting points.
  - Geweke and others.
- PROBLEM:  These all assume one long chain or multiple independent chains.  We have neither.

# Bootstrap Metric

- Bootstrap:
  - Resampling method with replacement approach that allows one to assign measures of accuracy (e.g. variance, confidence intervals) for complex estimators of a distribution, such as percentiles, proportions, etc. [1]
  - A common use case: we generate an MCMC chain with 10K samples. What is the uncertainty in the 95$^{th}$ percentile?
    - We draw 500 samples, where each of the 500 samples is a resampling of the 10K with replacement. For each of the 500 samples, we generate the 95$^{th}$ percentile, then examine the statistics over the 500.

Proposed metric = std(bootstrap QoI)/abs_value(QoI).

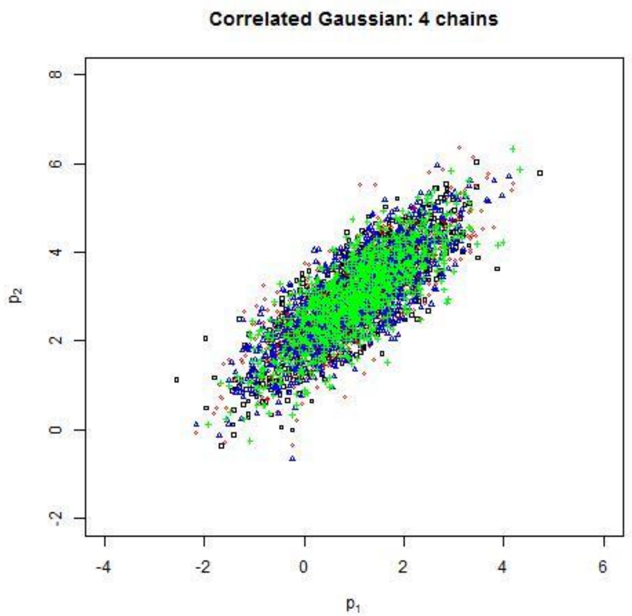
  - This metric gives an indication of the size of the relative error in the QoI. .
  - As we increase the number of generations in the chain, this ratio can be used as a stopping criterion for m-chain MCMC, *regardless of what m is.*
  - The bootstrap metric is valid whether we have 1-chain or a pooled m-chain.

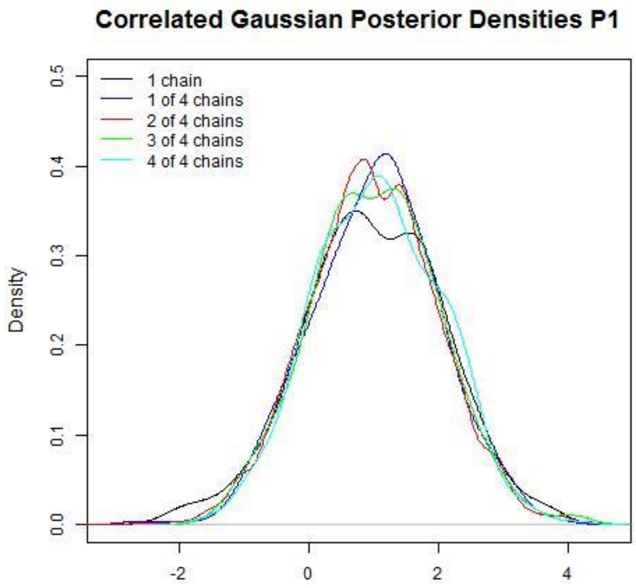


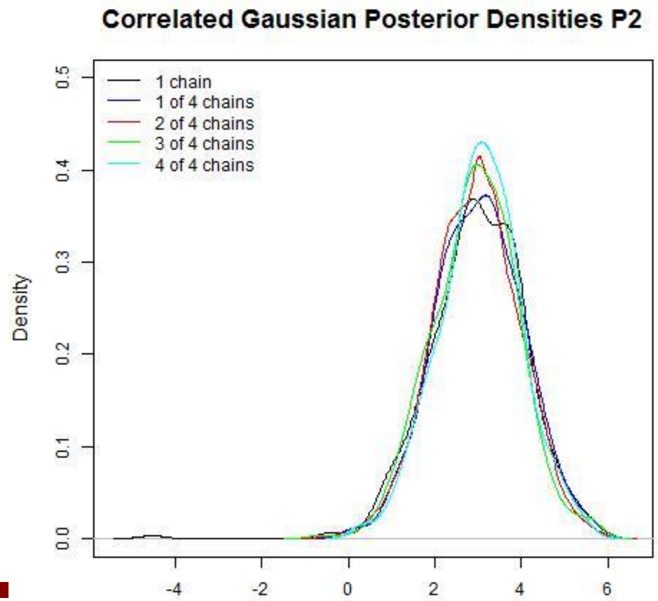
[1]**Efron, B and R. Tibshirani. An Introduction to the Bootstrap. Boca Raton, FL: Chapman & Hall/CRC. (1993)**

# Results

# Correlated Gaussian




- Prior are uniforms from [-5,8]
- Likelihood is correlated bivariate Gaussian, with mean(1,3) and correlation of 0.8
- Marginal posteriors are reasonable: means = (0.99, 2.99, etc.)

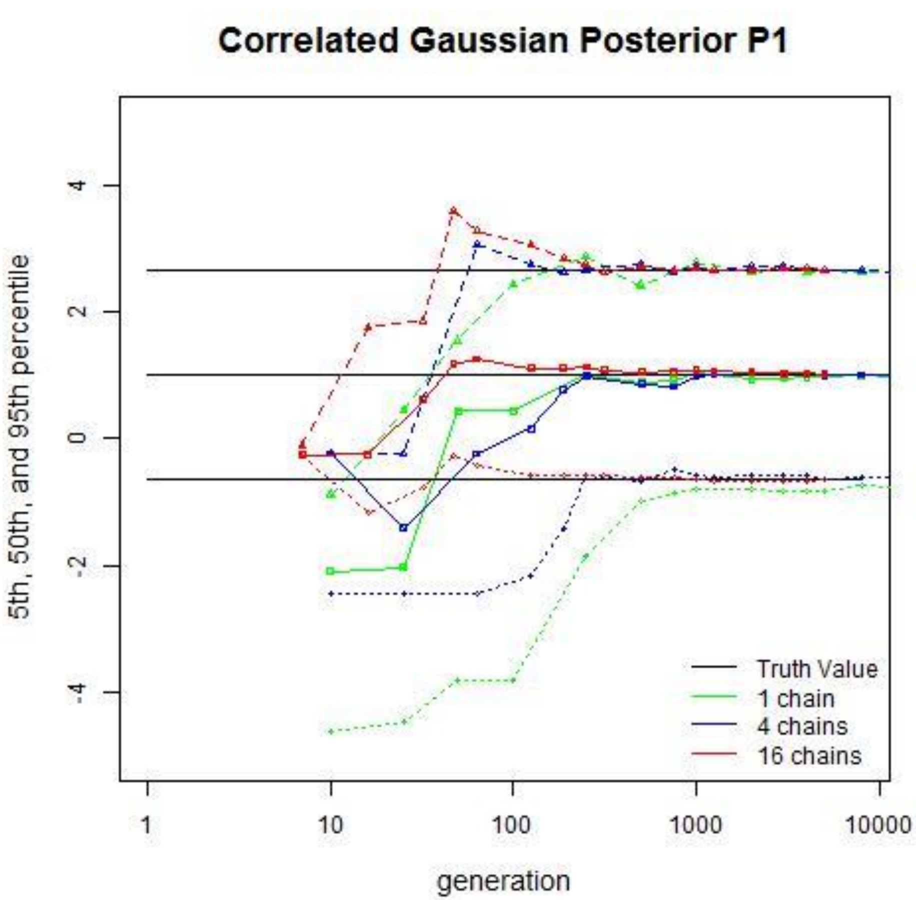
# Summary of correlated Gaussian convergence

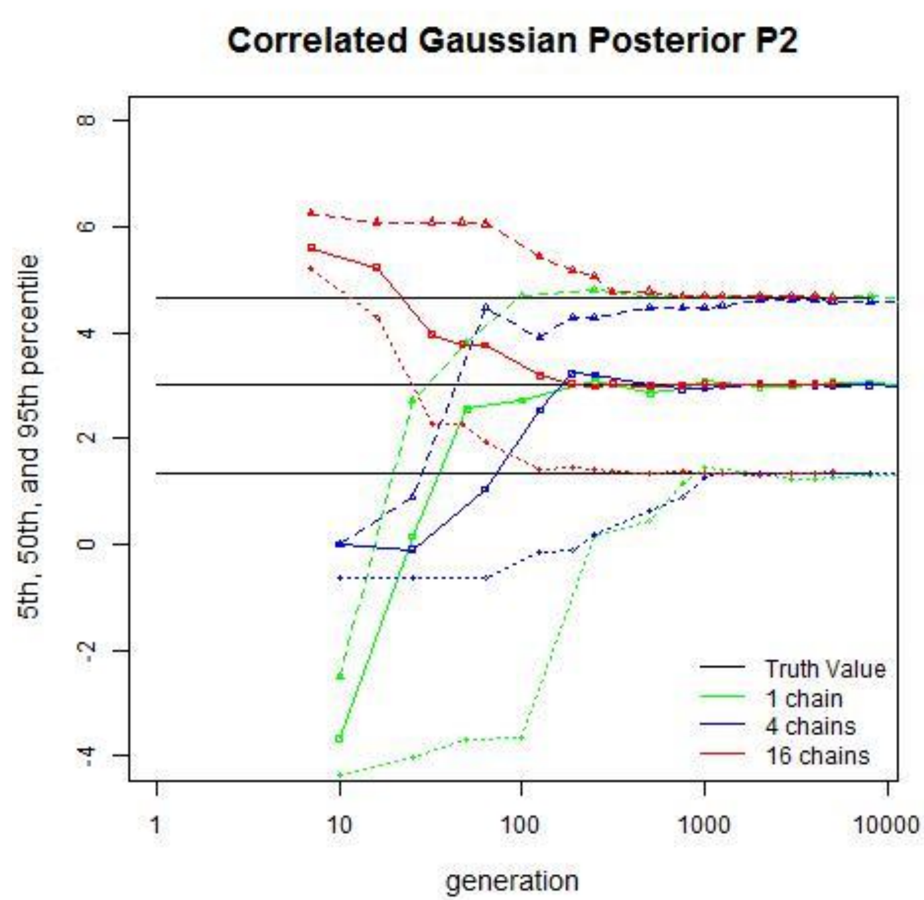
| Chains | Iters run | R-L [p1] [p2]<br>NOTE: for q=0.05,r=0.01 | Gelman |
|---|---|---|---|
| 1 | 100K | [10K], [10K] | |
| 4 | 20K/chain | [10K],[10K] | [6K,6K] |
| 16 | 5K/chain | Average : [9K], [9K] | [3K,3K] |

- Note: both Mcgibbsit and Raftery-Lewis were assessed for quantile q=0.05 at accuracy +/- 0.01 in the table above. In the table below, the accuracy requirements are increased to+/-0.005

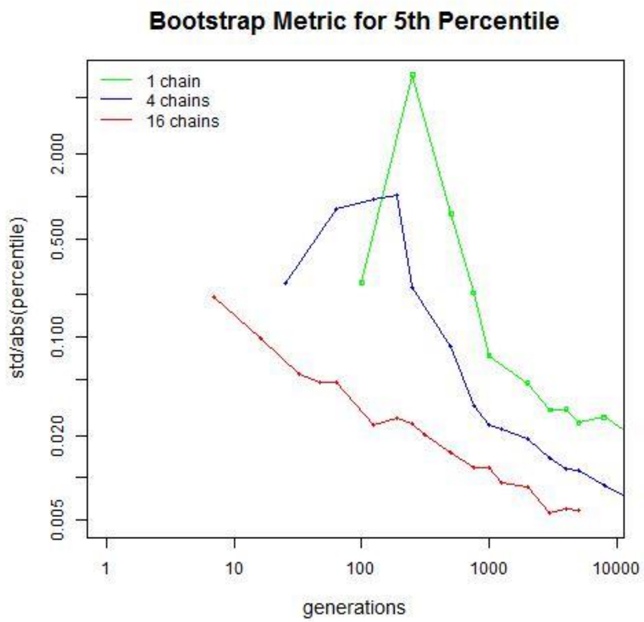
| Chains | Iters run | R-L [p1] [p2]<br>NOTE: for q=0.05, r=0.005 |
|---|---|---|
| 1 | 100K | [41K], [41K] |
| 4 | 50K/chain | [39K],[40K] |
| 16 | 50K/chain | |

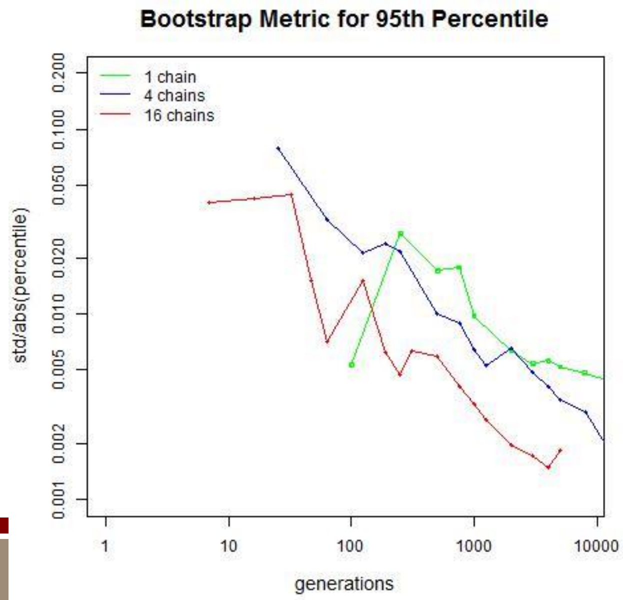
# Summary of correlated Gaussian convergence

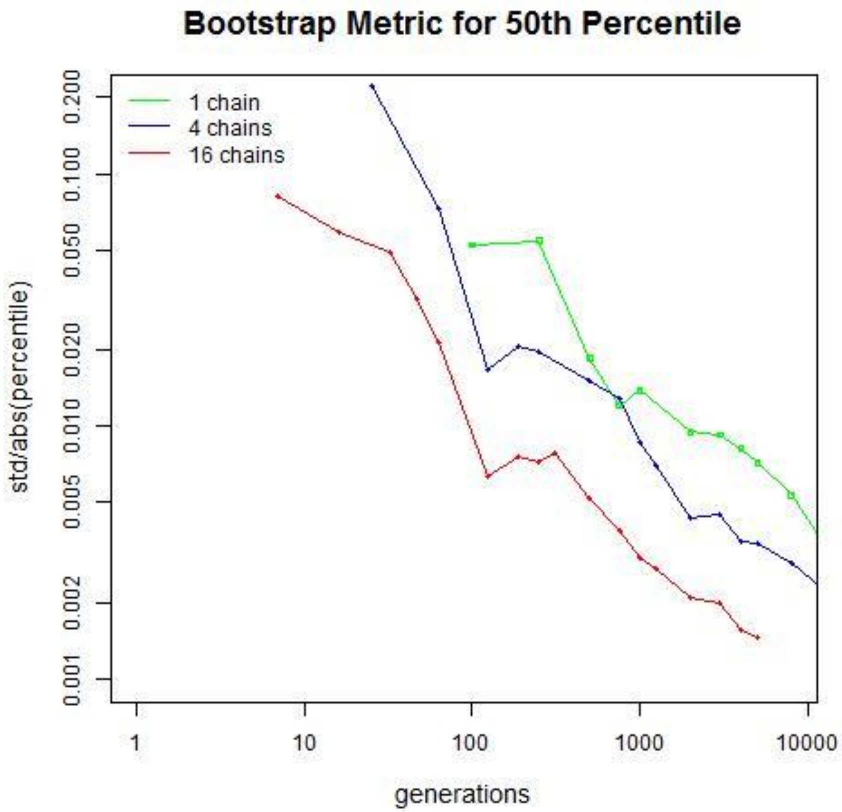


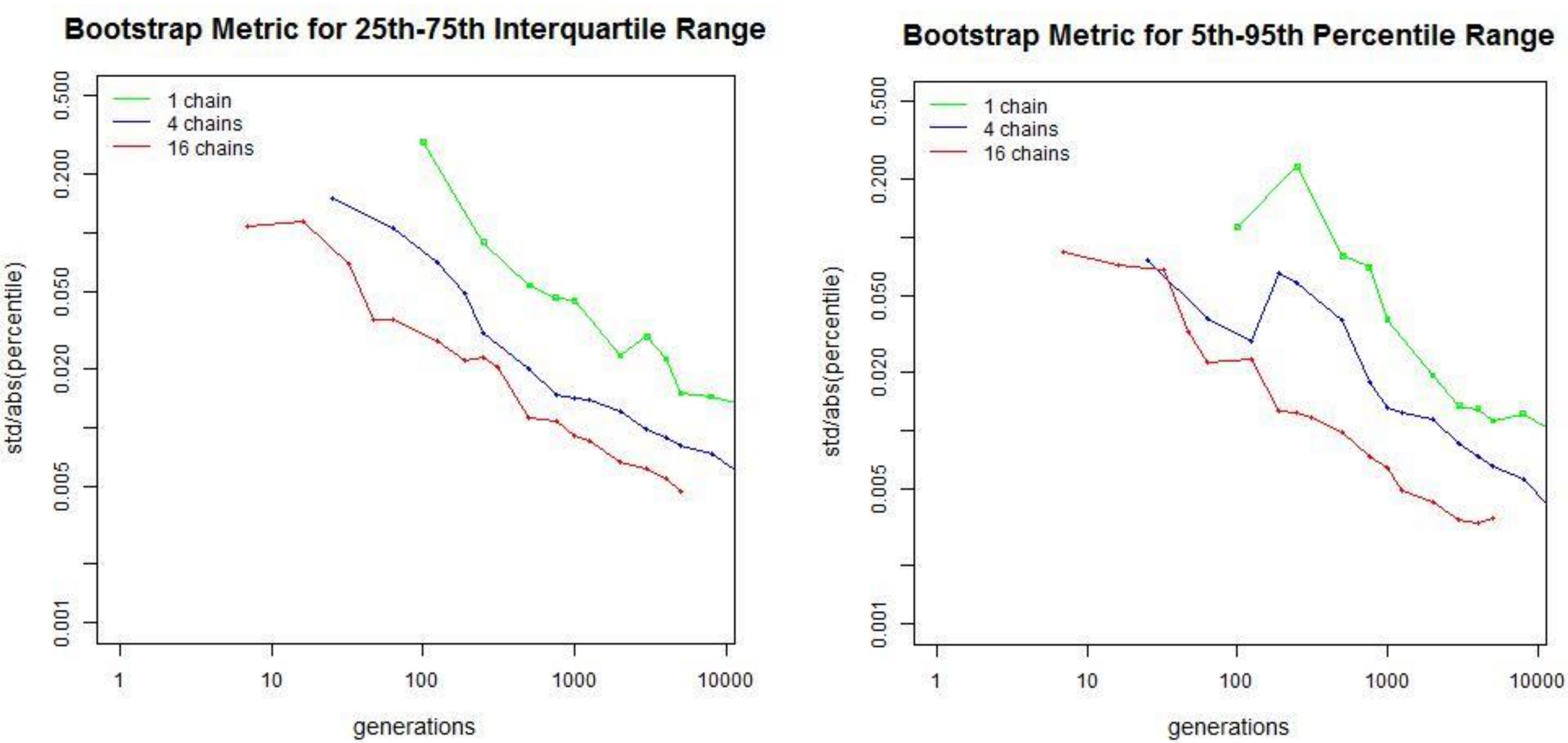
# Summary of correlated Gaussian convergence



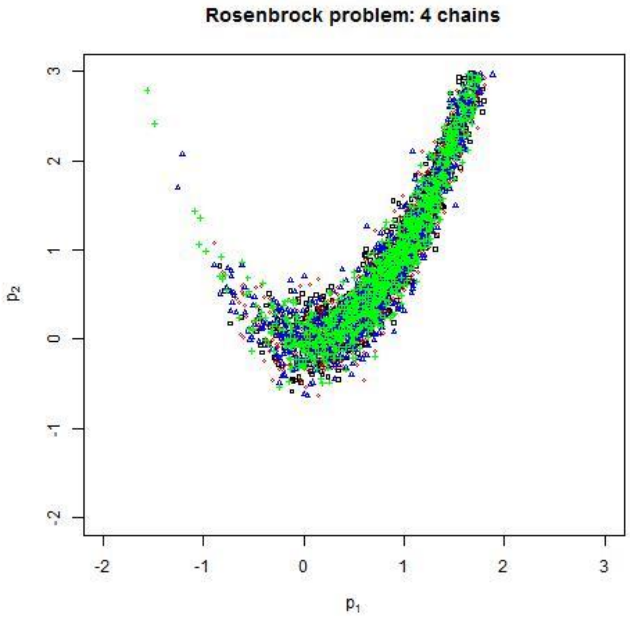



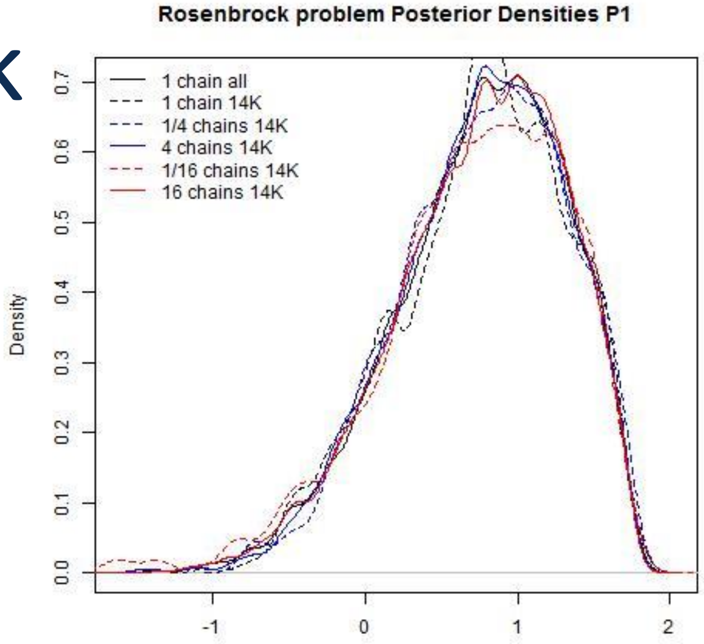



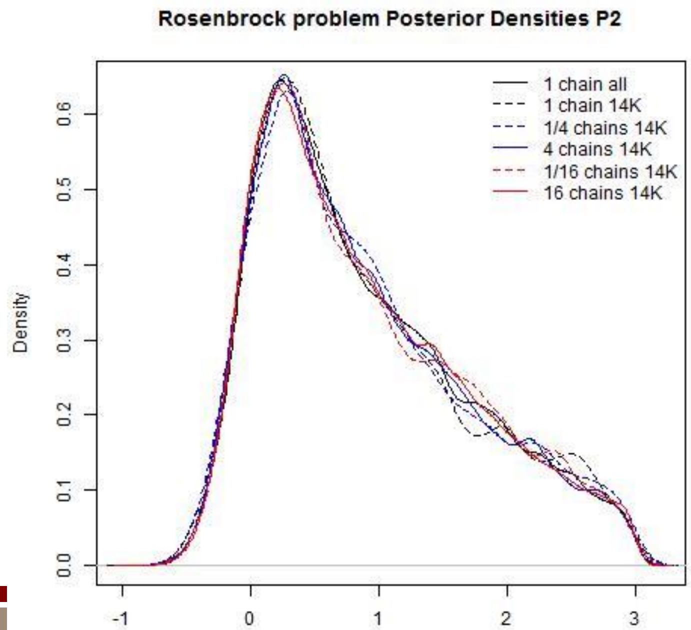
# Summary of correlated Gaussian convergence



# Rosenbrock




- Prior are uniforms from [-2,3]
- Likelihood is the Rosenbrock function

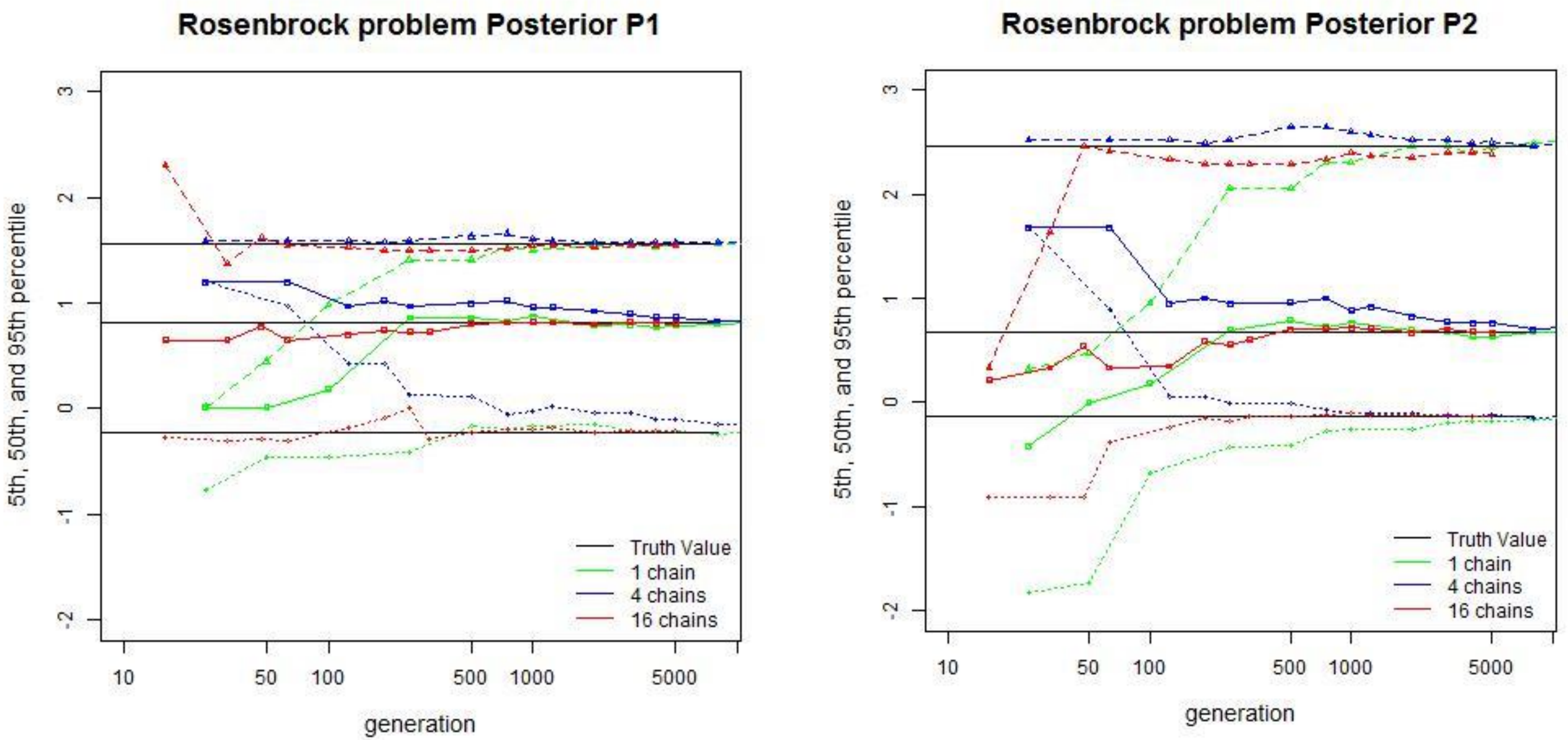
# Summary of Rosenbrock convergence runs

| Chains | Iters run | Mcgibbsit [p1][p2] | R-L [p1] [p2] NOTE: for q=0.05, r=0.01 | Gelman |
|---|---|---|---|---|
| 1 | 100K | [14K][4K] | [14K], [4K] | |
| 4 | 20K/chain | [15K], [4K] | [8K, 10K, 12K, 11K], [4K, 4K, 4K, 4K] | [5K,3K] samples |
| 16 | 5K/chain | [24K], [4K] | Average : [13K], 4K] | [3K,3K] samples |

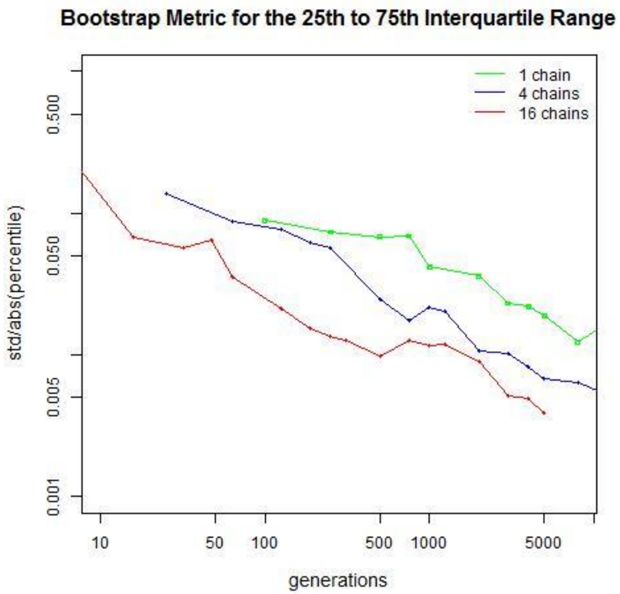
- Note: both Mcgibbsit and Raftery-Lewis were assessed for quantile q=0.025 at accuracy +/- 0.0125 in the table above. In the table below, the accuracy requirements are increased to+/-0.005

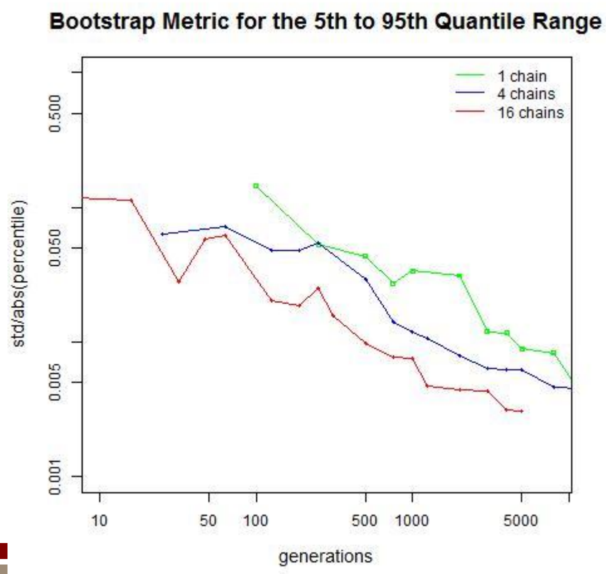
| Chains | Iters run | Mcgibbsit [p1][p2] | R-L [p1] [p2] NOTE: for q=0.025, r=0.005 | Gelman |
|---|---|---|---|---|
| 1 | 100K | [82K], [22K] | [86K], [23K] | |
| 4 | 50K/chain | [93K], [22K] | [47K, 60K, 74K, 68K], [23K, 22K, 22K, 22K] | [5K,3K] samples |
| 16 | 50K/chain | [140K], [23K] | Average : [79K], [23K] | [3K,3K] samples |

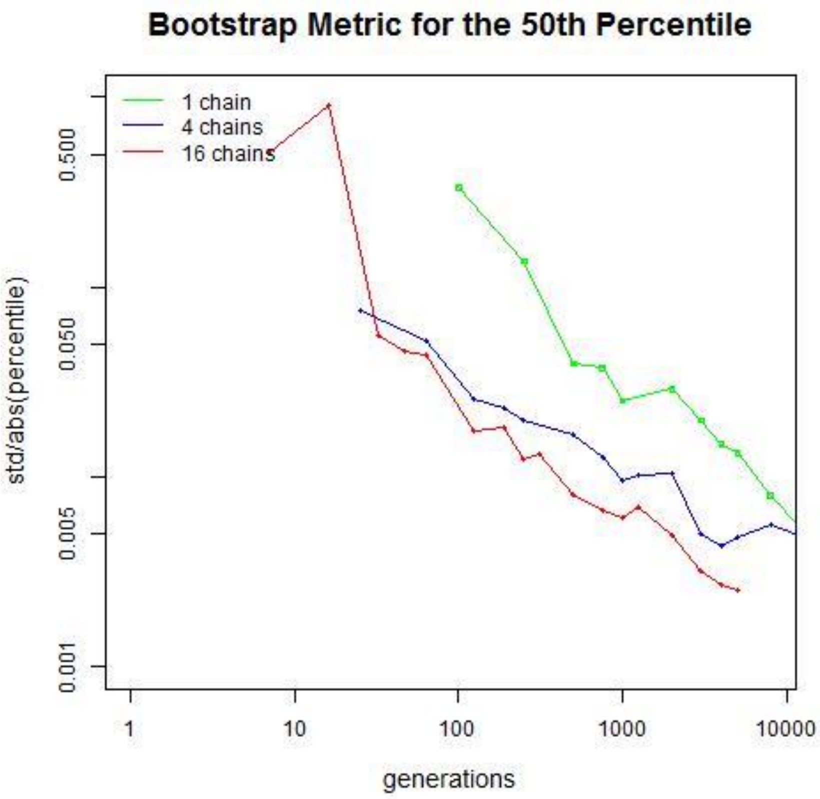
# Rosenbrock Convergence of Quantiles



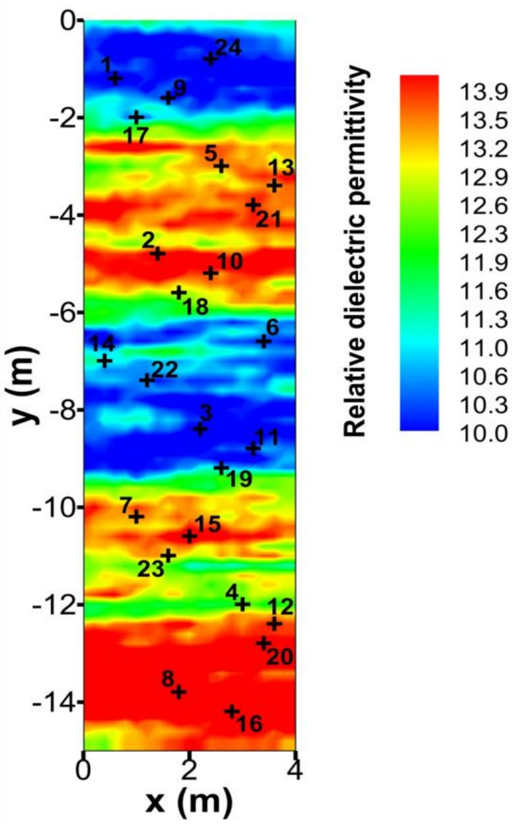
# Rosenbrock Convergence of Quantiles



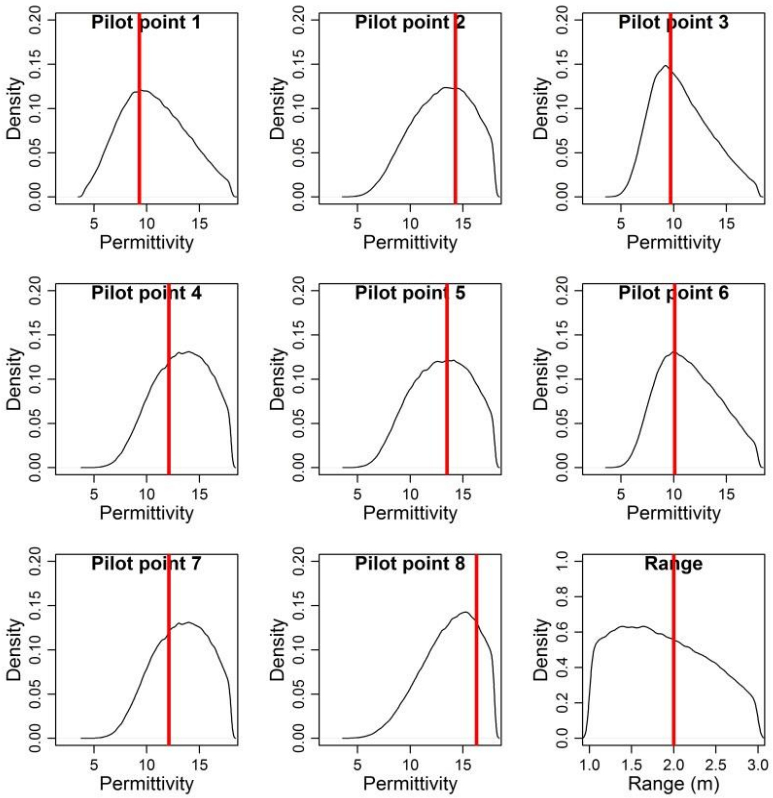



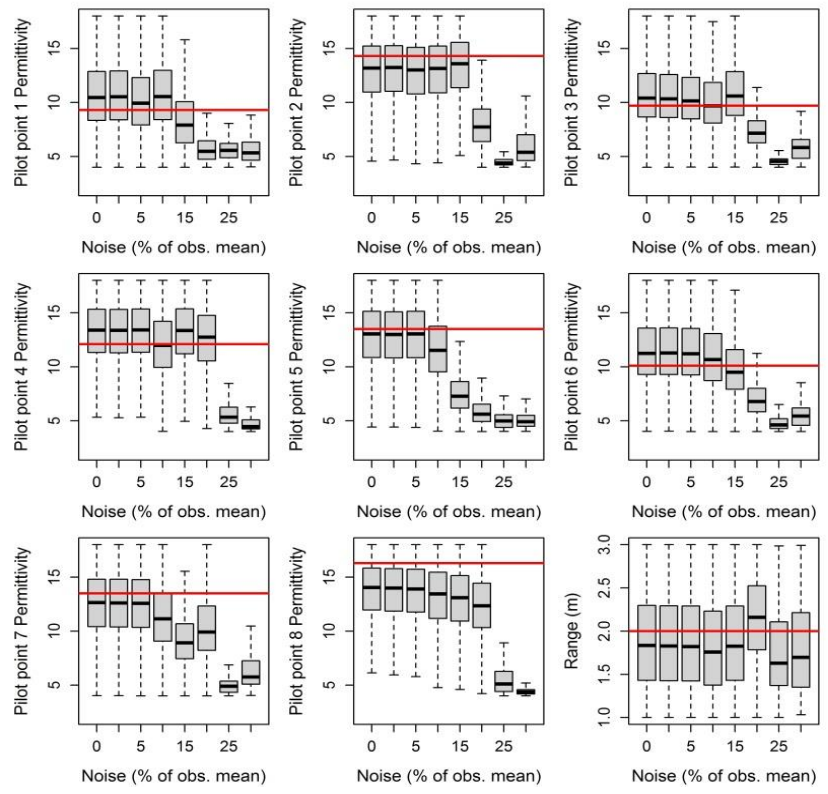
# GPR problem




- Monitoring soil moisture variations using tomographic ground penetrating radar (GPR) travel time data
- Tomographic GPR is a borehole-based geophysical imaging technique.
- It involves transmitting an electromagnetic (EM) pulse from a source in one borehole and recording the arrival of EM energy at a receiver position in a separate borehole.
- Inversion of the first arrival times of the EM energy is used to estimate the velocity and the dielectric permittivity ($\epsilon$) distribution between the boreholes.
- Use of pilot points to model the dielectric permittivity field
- Challenges exist in the inversion of GPR tomographic data for handling non-uniqueness and high-dimensionality of unknowns.

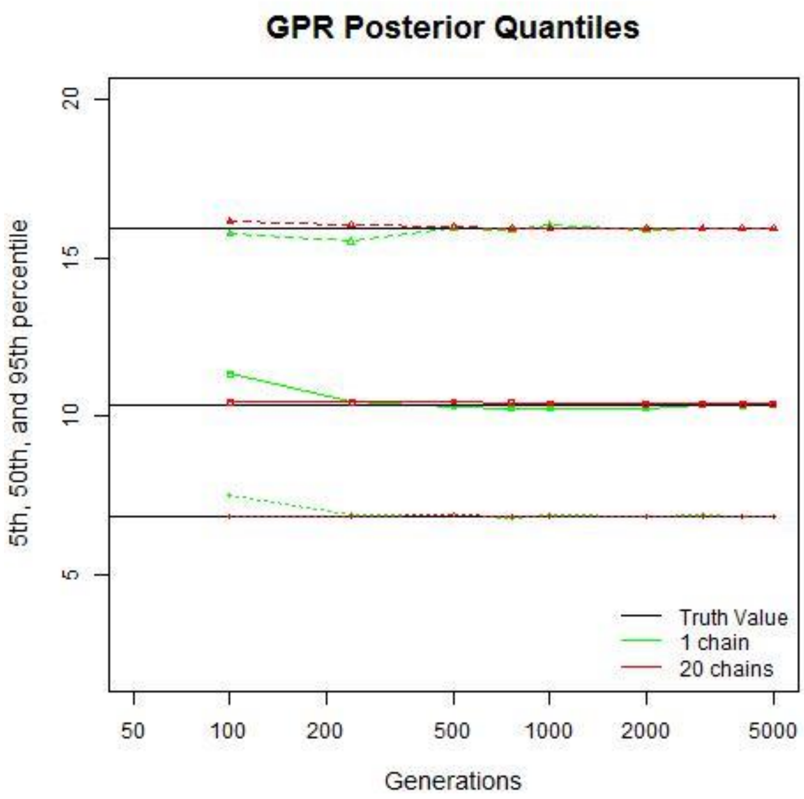
# Introduce GPR problem



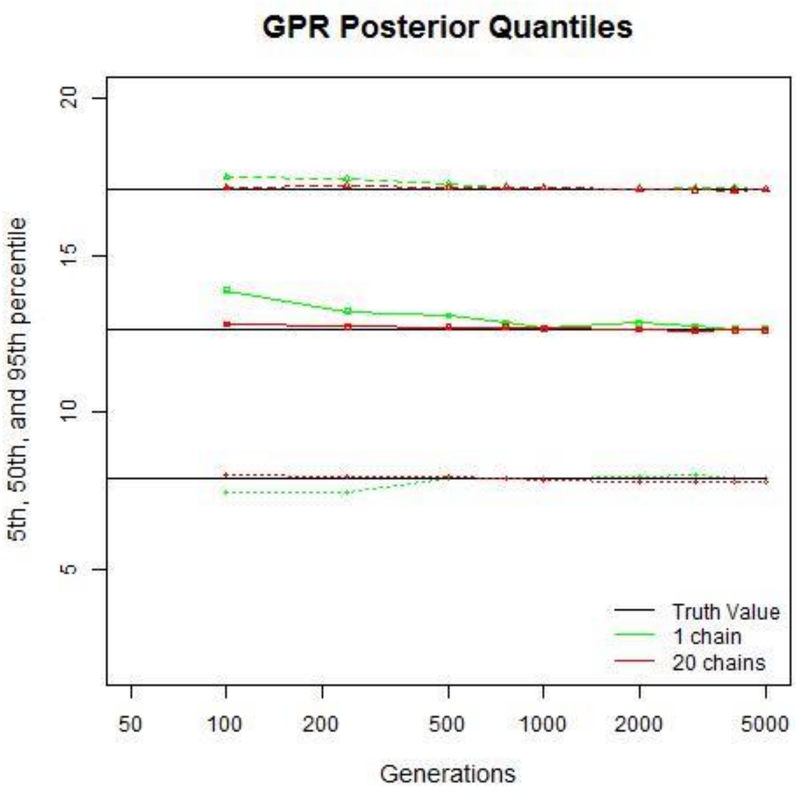



# 20 Chains on the GPR problem

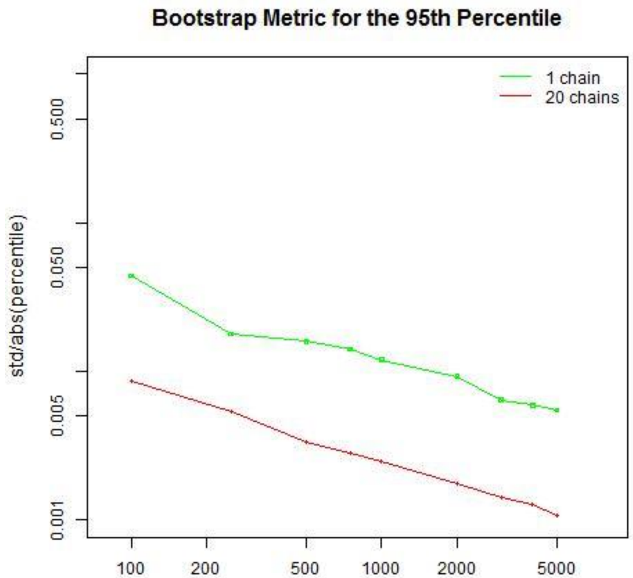
Parameter 3

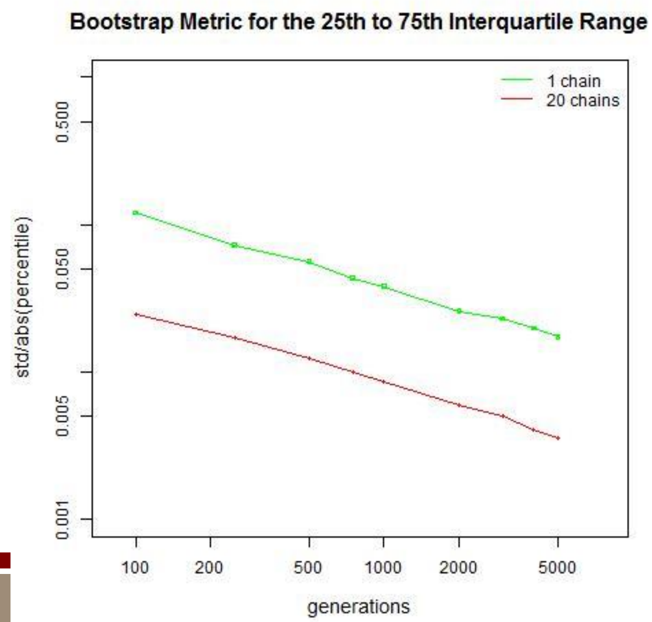


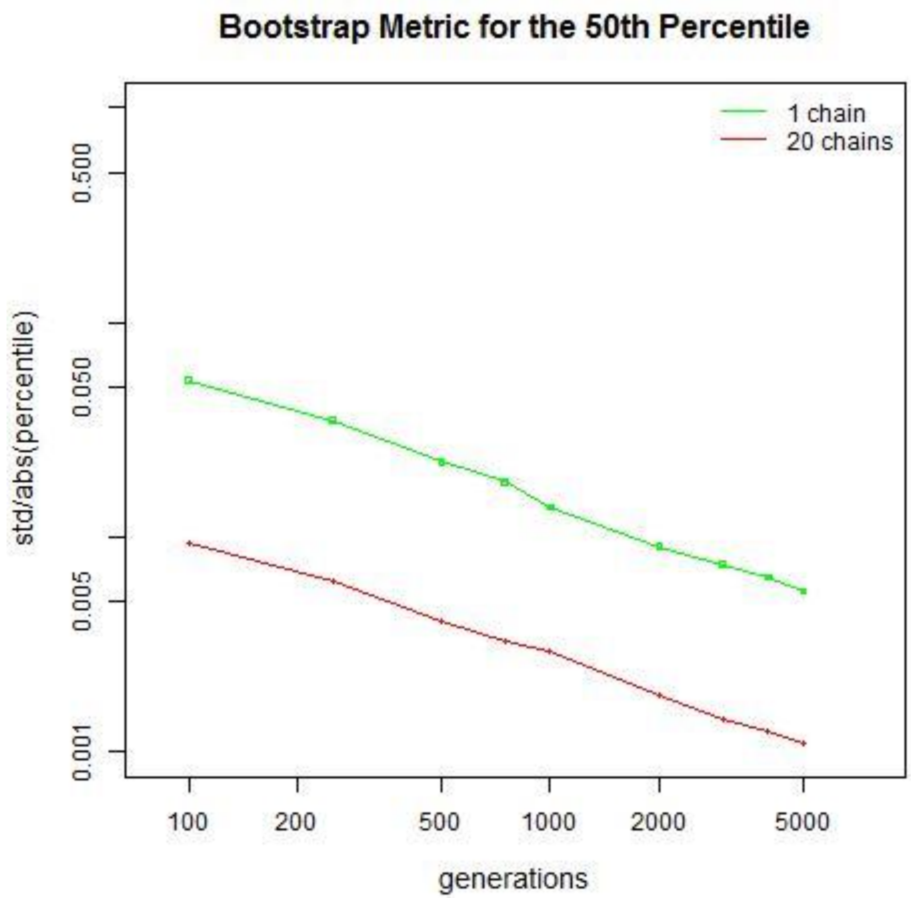

Parameter 7



# 20 Chains on the GPR problem



# Summary

- We have demonstrated parallel MCMC on two analytic problems (a correlated Gaussian and Rosenbrock) and one real hydrogeological problem
- The percentile estimates from pooled parallel MCMC chains are reasonable
- We get good convergence from these pooled chains with respect to the "truth" values in our analytic cases
- We developed a metric based on the bootstrap, the ratio of the std(bootstraps of the quantile)/abs(quantile).
- This metric can be used as a stopping criterion for m-chain MCMC.