Bayeslands: Framework featuring Bayesian inference for uncertainaity quanitification in basin and landscape evolution models (Badlands)

Bayeslands Workshop

Rohitash Chandra

EarthByte Group and Centre for Translational Data Science
The University of Sydney, Sydney, Australia

July 4, 2018

Introduction

- Models require methods that uncover the free parameters that best describes the data.
- Data in many cases is sparse, limited, or incomplete.
- The search for the free parameters in models: optimisation methods, eg. gradient based methods, simplex search, genetic or evolutionary algorithms, meta-heuristics [in case when no gradient information from model is available].

Introduction

- Challenges in optimisation methods given large scale of parameters - when gradient information is not available. The limitations are in terms of uncertainty quantification. The need to run multiple experiments with different initial set of parameters to check robustness for convergence. Limitations of p values for statistical tests.
- Bayesian inference methods - probability distributions instead of single point estimates. There is no need to run multiple given that the inference algorithm has converged to a distribution.

Bayesian inference

- Bayesian inference provides a principled approach towards uncertainty quantification of free parameters in geophysical forward models.
- The use of MCMC methods in the geosciences have been well established, with applications spanning from modelling geochronological ages, inferring sea-level and sediment supply from the stratigraphic record, and inferring groundwater contamination sources.

Bayesian inference

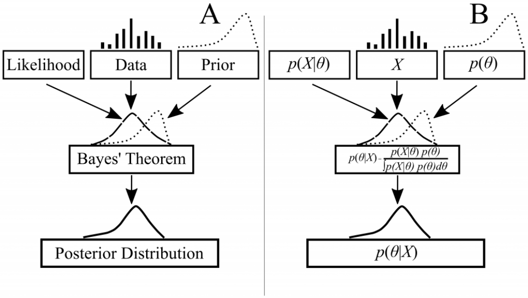


Figure 1: Bayesian inference overview

Markov Chain Monte Carlo sampling methods (MCMC) implement Bayesian inference that sample from a probability distribution. This is based on constructing a Markov chain after a number of steps that has the desired distribution as its equilibrium distribution.

Bayesian inference

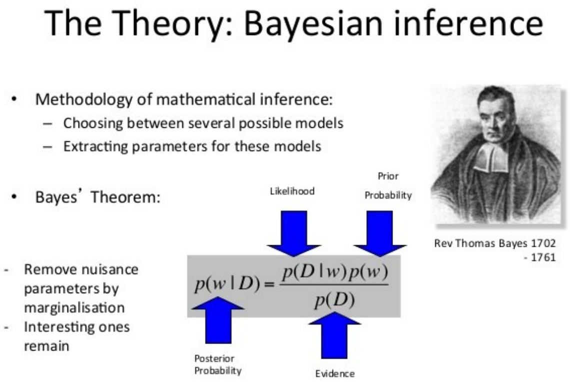


Figure 2: Bayesian inference overview

Bayesian inference

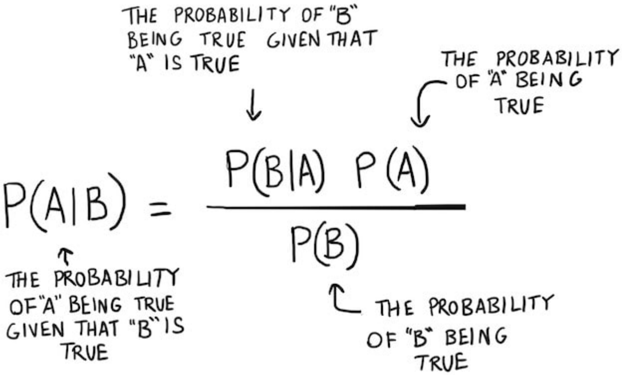


Figure 3: Bayesian inference overview

Bayesian inference

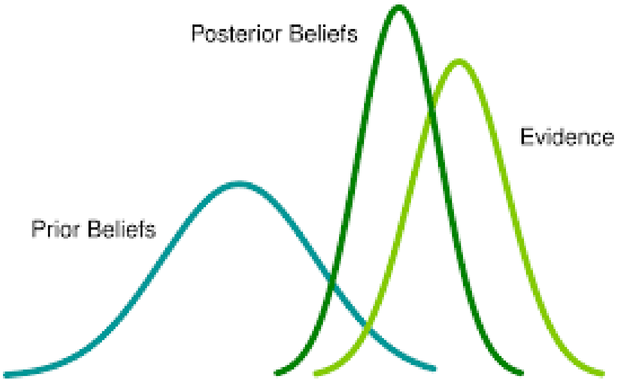


Figure 4: Bayesian inference overview

Bayesian inference

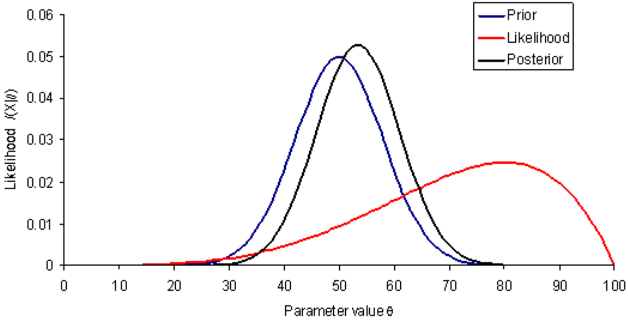


Figure 5: Bayesian inference overview

Surface

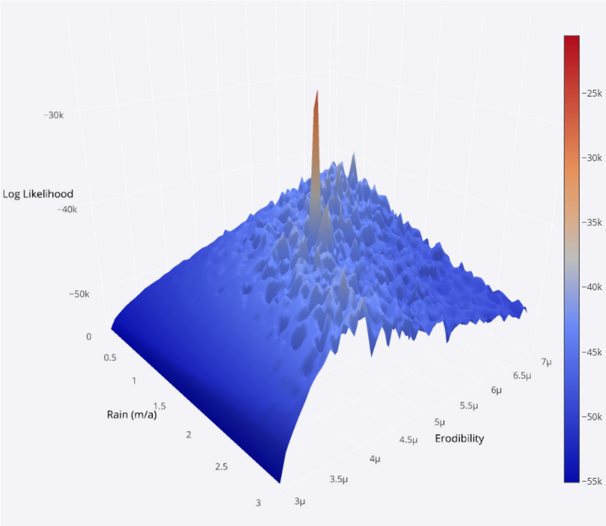


Figure 6: Likelihood surfaces of the Continental Margin topography for the rainfall and erodibility parameters only

MCMC framework

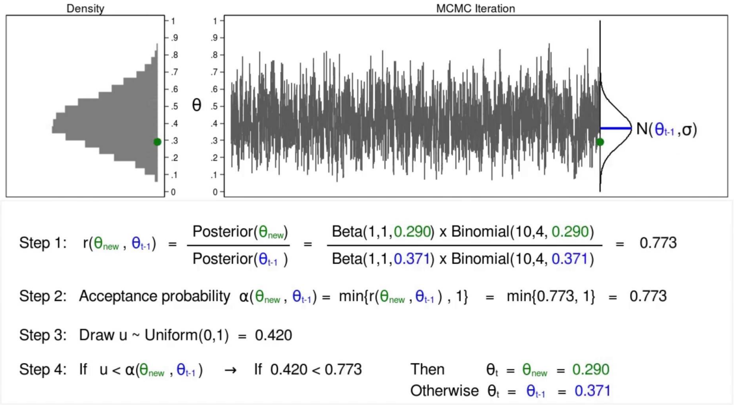
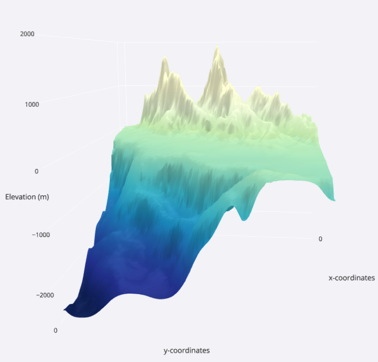
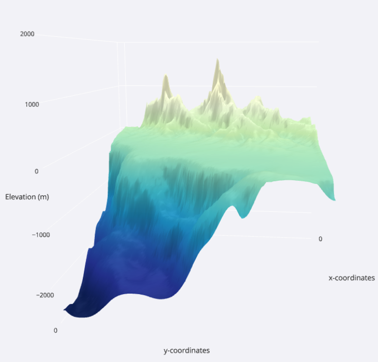


Figure 7: MCMC sampling

Badlands



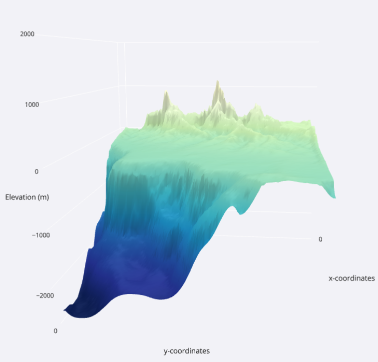
(a) CM-extended predicted topography after 25 % evolution time



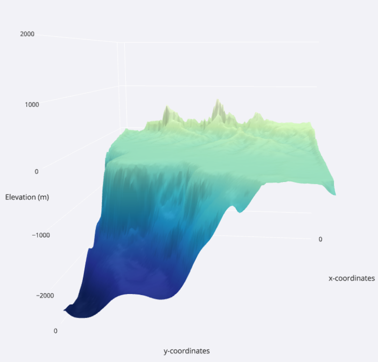
(b) CM-extended predicted topography after 50 % evolution time

Figure 8: CM-extended evolution over 1 000 000 years.

Badlands



(a) CM-extended predicted topography after 75 % evolution time



(b) CM-extended predicted topography after 100 % evolution time

Figure 9: CM-extended

Basin and landscape dynamics via Bayeslands

- Bayeslands is a framework for inference and uncertainty quantification in the Badlands model for basin and landscape evolution.
- Bayeslands extends *Badlands* by placing probability distributions over the free parameters such as rainfal and erobility - thereby turning a deterministic model into a probabilistic one.

Bayeslands framework

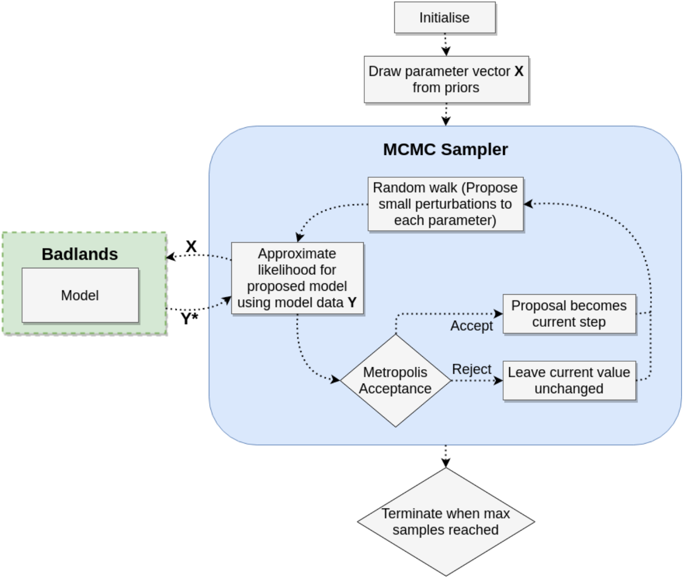


Figure 10: *Bayeslands* framework for Badlands.

Problem formulation

The observed data can be varied in type and scale and is denoted by $\mathcal{D}$. Mathematically, this relationship between prior and posterior beliefs is expressed by Bayes theorem:

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})}$$

where $p(\theta|\mathcal{D})$ is the posterior probability of θ, $p(\mathcal{D}|\theta)$ is the likelihood function and is interpretable as how likely the observed data are, if the true state of the world is given by θ, and $p(\theta)$ is the prior belief regarding θ.
Note that
$p(\mathcal{D}|\theta) = p(\mathcal{D}|g(\theta))$, because $g(\theta)$ is a deterministic function of θ. This equality allows us to make inference regarding geoscientific unknowns (θ), by connecting the true simulated values η to measurable quantities contained in $\mathcal{D}$.

Bayes Theorem

Conditional Probability

- The probability of A given B:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ if } P(B) > 0$$

- Multiplication rule:

$$P(A \cap B) = P(A)P(B|A)$$

- Bayes law:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

Acceptance probability

Algorithm acceptance probability derivation:

$$\alpha = \frac{P(\theta^p|\mathcal{D})Q(\theta^c|\theta^p)}{P(\theta^c|\mathcal{D})Q(\theta^c|\theta^p)}$$

$$\frac{P(\theta^p|\mathcal{D})}{P(\theta^c|\mathcal{D})} = \frac{P(\mathcal{D}|\theta^p)P(\theta^p)/\cancel{P(\mathcal{D})}}{P(\mathcal{D}|\theta^c)P(\theta^c)/\cancel{P(\mathcal{D})}}$$

$$= \frac{P(\mathcal{D}|\theta^p)P(\theta^p)}{P(\mathcal{D}|\theta^c)P(\theta^c)}$$

$$= \frac{P(\mathcal{D}|g(\theta^p))P(\theta^p)}{P(\mathcal{D}|g(\theta^c))P(\theta^c)}$$

where $g(\theta)$ refers to Badlands model.

Parallel tempering

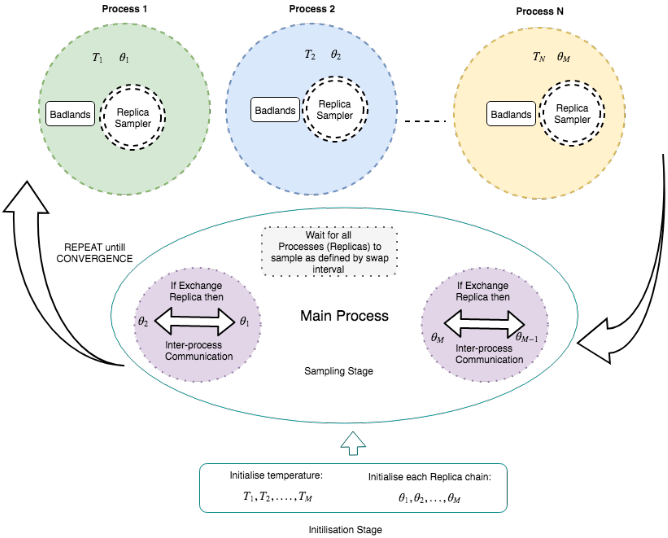


Figure 11: An overview of the different replicas that are executed on a multi-processing architecture. Note that the main process controls the given replicas and enables them to exchange the neighboring replicas given the swap time and probability of exchange is satisfied.

Parallel tempering

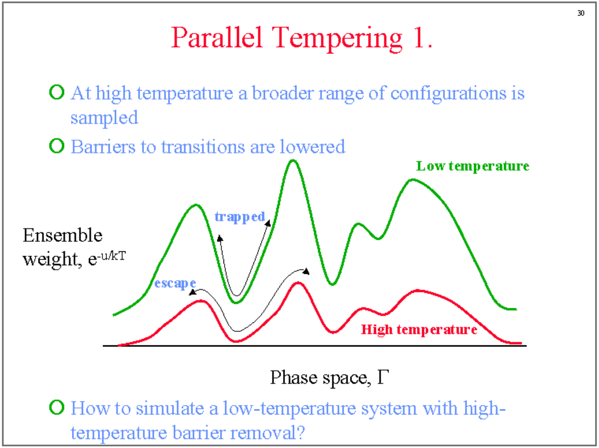
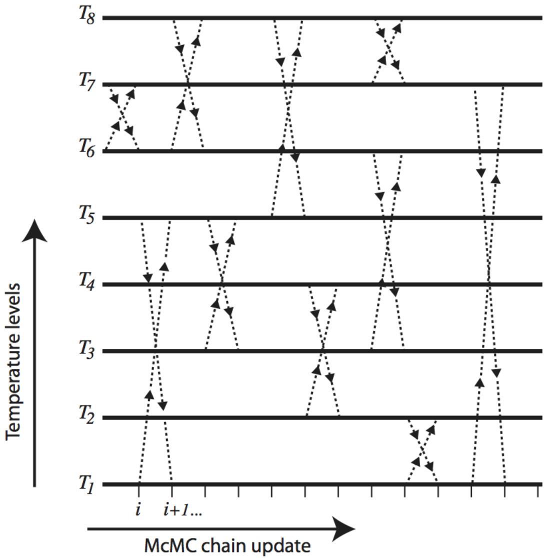


Figure 12: Temperature effects

Parallel tempering



Parallel tempering

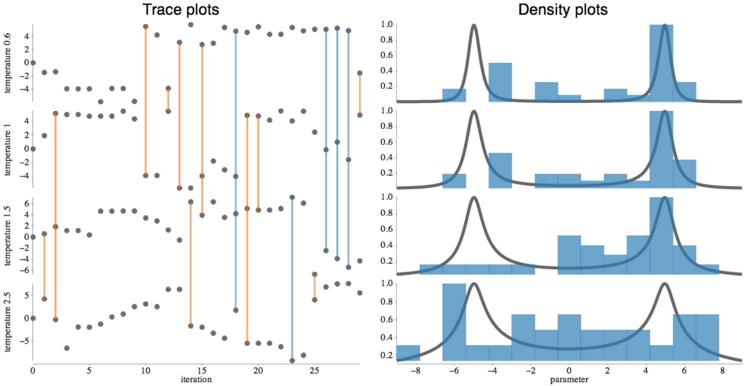


Figure 14: Parallel tempering - sampling

Parallel tempering

Alg. 1 Metropolis-Hastings (MH)

Result: Drawing from $p(\theta|\mathbf{D})$

Set maximum number of iterations, K, and initialize $\theta = \theta^{[0]}$

Set the current value of θ, $\theta^c = \theta^{[0]}$

for $k = 1, \ldots, K$ **do**

1. Propose a new value of θ, $\theta^p \sim q(\theta|\theta^c)$

2. Compute acceptance probability

$$\alpha = \min\left(1, \frac{p(\theta^p|\mathbf{D})q(\theta^c|\theta^p)}{p(\theta^c|\mathbf{D})q(\theta^p|\theta^c)}\right)$$

3. Draw $u \sim U[0,1]$

if $u < \alpha$ **then**

$\theta^{[k]} = \theta^p$

else

$\theta^{[k]} = \theta^c$

end

end

Figure 15: Algorithm: MCMC sampler

Parallel tempering

Alg. 2 Metropolis-Hastings Parallel Tempering (MHPT)

Result: Drawing from $p(\theta|\mathbf{D})$
Set maximum number of iterations, K, and initialize $\theta = \theta^{[0]}$, and $m = m^{[0]}$
Set the current value of θ, to $\theta^c = \theta^{[0]}$ and m to $m^c = m^{[0]}$
for $k = 1, \ldots, M$ **do**

1. Update $\theta^{[k]}$, from the chain with $p(\theta|\mathbf{D})^{\beta_{m^c}}$ as its invariant distribution according to Alg 1.
2. Propose a new value of m^p, from $q(m^p|m^c)$.
3. Compute acceptance probability

$$\alpha = \min\left(1, \frac{p(\theta^{[k]}|\mathbf{D})^{\beta_{m^p}} p(m^p) q(m^c|m^p)}{p(\theta^{[k]}|\mathbf{D})^{\beta_{m^c}} p(m^c) q(m^p|m^c)}\right)$$

4. Draw $u \sim U[0, 1]$

  if $u < \alpha$ **then**
    $m^{[k]} = m^p$
  else
    $m^{[k]} = m^c$
  end
end

Figure 16: Algorithm: Parallel tempering

Synthetic topography - ground truth

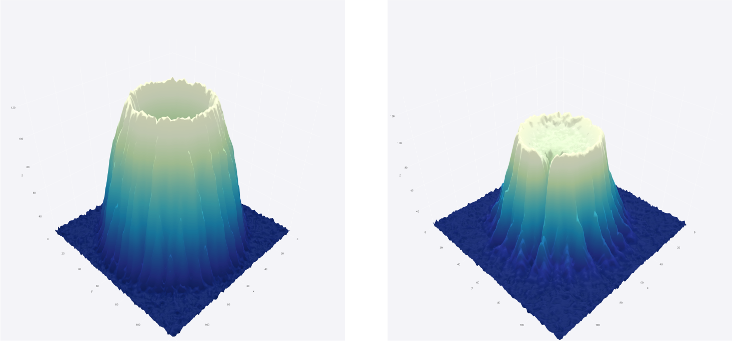
Topography	Rainfall (m/a)	Erod.	n-value	m-value	c-marine	c-surface	Uplift (mm/a)
Crater	1.5	5.0-e05	1.0	0.5	-	-	-
Crater-extended	1.5	5.0-e05	1.0	0.5	-	-	-
CM	1.5	5.0-e06	1.0	0.5	0.5	0.8	-
CM-extended	1.5	5.0-e06	1.0	0.5	0.5	0.8	-
Mountain	1.5	5.0-e06	1.0	0.5	-	-	1.0

Table 1: True values of parameters

Topography	Rainfall (m/a)	Erod.	n-value	m-value	c-marine	c-surface	uplift
Crater	[0,3.0]	[3.0-e05, 7.0-e05]	-	-	-	-	
Crater-ext	[0,3.0]	[3.0-e05, 7.0-e05]	[0, 2.0]	[0, 2.0]	-	-	
CM	[0,3.0]	[3.0-e06, 7.0-e06]	-	-	-	-	-
CM-ext.	[0,3.0]	[3.0-e06, 7.0-e06]	[0, 2.0]	[0, 2.0]	[0.3, 0.7]	[0.6, 1.0]	-
Mountain	[0,3.0]	[3.0-e06, 7.0-e06]	[0, 2.0]	[0, 2.0]	-	-	[0.1, 1.7]

Table 2: Prior distribution range of model parameters

Example 1: Crater

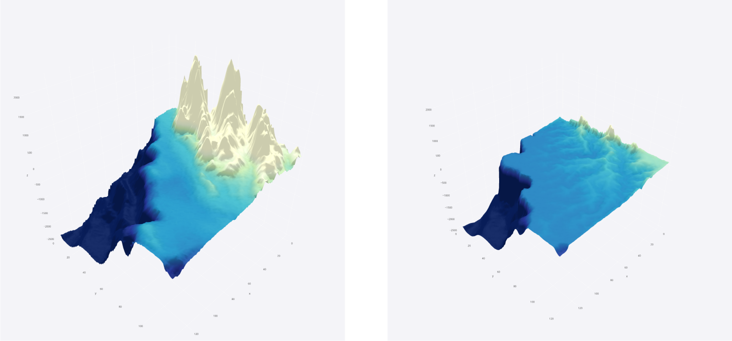


(a) Crater-extended initial topography

(b) Crater-extended synthetic ground-truth topography

Figure 17: Crater-extended: Initial and eroded ground-truth topography and sediment deposition after 50 000 years.

Example 1: Continental Margin

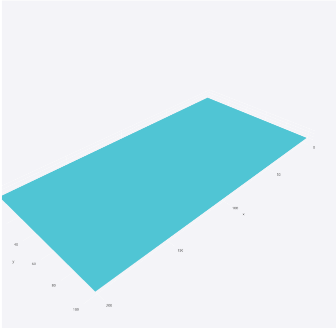


(a) CM-extended initial topography

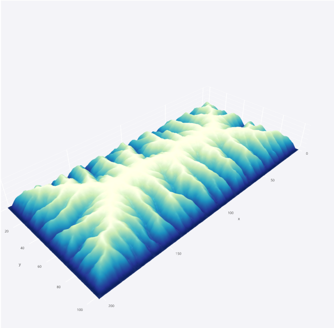
(b) CM-extended synthetic ground-truth topography

Figure 18: Continental Margin (CM)-extended: Initial and eroded ground-truth topography and sediment after 1 000 000 years.

Example 1: Mountain



(a) Mountain initial topography



(b) Mountain synthetic ground truth topography

Figure 19: Mountain: Initial and eroded ground-truth topography after 1 000 000 years evolution.

Model: Likelihood

Let the initial topography be denoted by $\boldsymbol{D}_0$, with $\boldsymbol{D}_0 = (D_{0,s_1} \ldots, D_{0,s_n})$, where s_i corresponds to site s_i, with co-ordinates latitude, u_i, and longitude, v_i. Suppose that we are interested in the topography t years into the future, we will denote this by $\boldsymbol{D}_t$, with $\boldsymbol{D}_T$ defined to be the current topography. Our model for the process which generates the topography is

$$D_{t,s_i} = f_{t,s_i}(\boldsymbol{\theta}) + \epsilon_{t,s_i} \text{ with } \epsilon_{t,s_i} \sim (0, \tau^2) \tag{1}$$

Model: Likelihood

for $t = 0, 1, \ldots, T$ and $i = 1, \ldots, n$, where $\boldsymbol{\theta}$ are the parameters of the Badlands model and $f_{t,s_i}(\boldsymbol{\theta})$ is the output of the Badlands forward model. This model states that the topography is function of the Badlands forward model given parameters $\boldsymbol{\theta}$, plus some Gaussian noise with zero mean and variance τ^2. The likelihood function $L_l(\boldsymbol{\theta})$,is given by

$$L_l(\boldsymbol{\theta}) = \frac{1}{(2\pi\tau^2)^{n/2}} \exp\left\{ -\frac{1}{2} \frac{\sum_{t=1}^{T} \sum_{i=1}^{n} (D_{t,s_i} - f_{t,s_i}(\boldsymbol{\theta}))^2}{\tau^2} \right\}$$

where the subscript l, in $L_l(\theta)$, denotes that it is the landscape likelihood.

Model: Likelihood

The sediment likelihood, $L_s(\theta)$ is

$$L_s(\boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{mT/2}} \exp\left\{ -\frac{1}{2} \sum_{t=1}^{T} \sum_{j=1}^{m} \frac{(Z_{t,s_j} - g_{t,s_j}(\boldsymbol{\theta}))^2}{\sigma^2} \right\}$$

giving a total likelihood as a function of theta $L(\boldsymbol{\theta})$ to be

$$L(\boldsymbol{\theta}) = L_s(\boldsymbol{\theta}) \times L_l(\boldsymbol{\theta}).$$

Design of Experiments

- Step 1: Investigate the effects on computational time and accuracy when increasing the number of replicas;
- Step 2: Evaluate the number of samples required for convergence defined in terms of prediction accuracy;
- Step 3: Using knowledge from above experiments, apply PT-Bayeslands to all the given problems and report the posterior distributions, computational time, uncertainty and accuracy in topography and sediment predictions.

Results

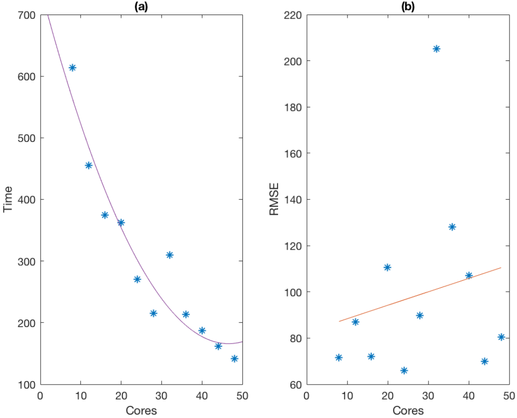


Figure 20: The effect of the number of cores on time, panel (a) and on the RMSE, panel (b). Note that a 95% confidence interval for the slope of the regression line in panel (b), is [-14.2,2.6], showing that the slope is not significantly different from zero.

Results

Replica (cores)	Time (minutes)	Pred. RMSE	Accepted %
2	3540	78.8	0.1
4	1384	62.7	0.1
8	614	71.5	0.4
12	455	87.0	0.5
16	375	72.0	0.6
20	362	110.6	0.3
24	270	65.9	0.4
28	215	89.7	0.5
32	310	205.1	0.4
36	213	128.0	0.6
40	187	107.0	0.6
44	162	70.0	0.6
48	142	80.5	0.8

Table 3: Effect of number replicas/cores for the CM-extended topography with 100 000 samples.

Results

Num. Samples	Time (mins.)	Pred. RMSE	Accepted %
1000	3	120.6	14.0
2000	6	222.5	15.3
4000	10	200.8	9.4
6000	18	98.3	3.7
8000	18	213.1	8.3
10000	27	77.3	2.8
50000	131	77.7	1.3
100000	258	95.1	1.0
150000	436	69.3	0.2

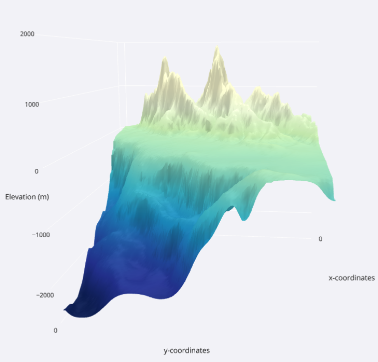
Table 4: Effect of the number of samples for the CM-extended topography running with 24 replicas on 24 cpus.

Results

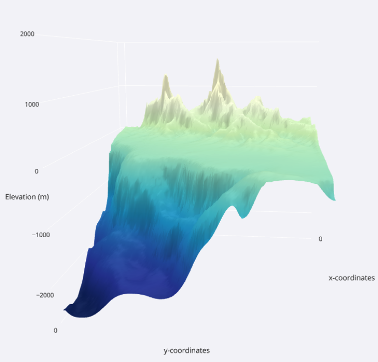
Topography	Time (min)	Sed. RMSE	Elev. RMSE	Pred. RMSE	Acc. %
Crater	80	4.2	1.1	5.3	12.2
Crater-ext	229	0.2	1.0	1.2	1.9
CM	50	2.6	19.9	22.5	1.0
CM-ext	257	50.2	49.0	99.2	0.6
Mountain	375	-	617.0	617.0	0.63

Table 5: Typical results for respective problems (24 replicas/cpus and 100 000 samples). Note that '-' in case of Mountain Sed. RMSE indicates that it was not part of the likelihood function

Results



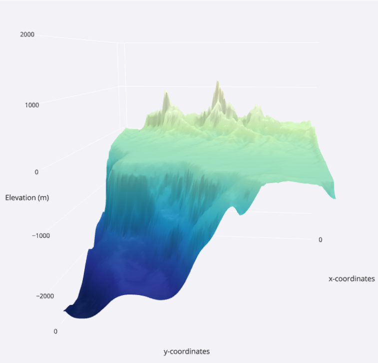
(a) CM-extended predicted topography after 25 % evolution time



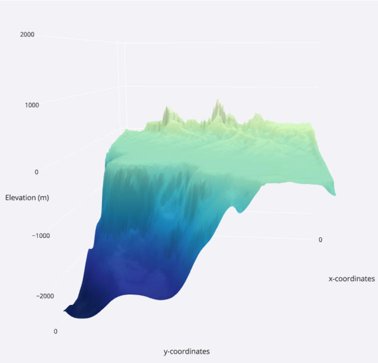
(b) CM-extended predicted topography after 50 % evolution time

Figure 21: CM-extended evolution over 1 000 000 years.

Results



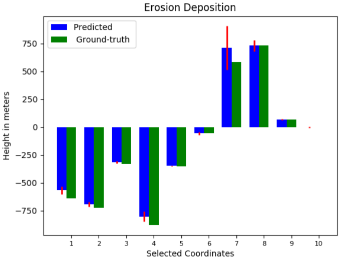
(a) CM-extended predicted topography after 75 % evolution time



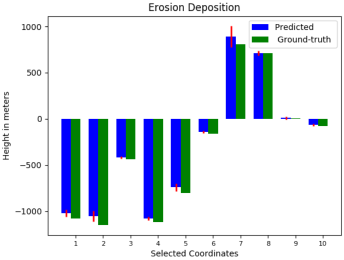
(b) CM-extended predicted topography after 100 % evolution time

Figure 22: CM-extended: Topography and erosion-deposition development for selected time frames. Note that erosion (positive) and deposition (negative) values given by the height in meters

Results



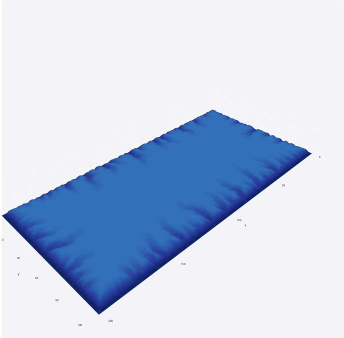
(a) CM-extended erosion-deposition after 50 % evolution time



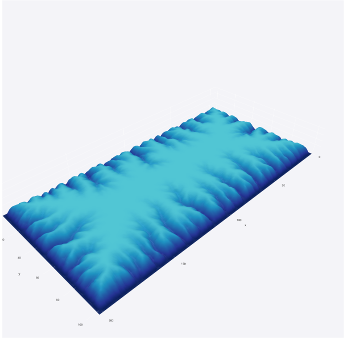
(b) CM-extended erosion-deposition after 100 % evolution time

Figure 23: CM-extended: Topography and erosion-deposition development for selected time frames. Note that erosion (positive) and deposition (negative) values given by the height in meters

Results



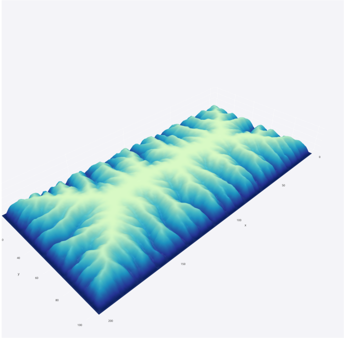
(a) Mountain predicted topography after 25 % evolution time



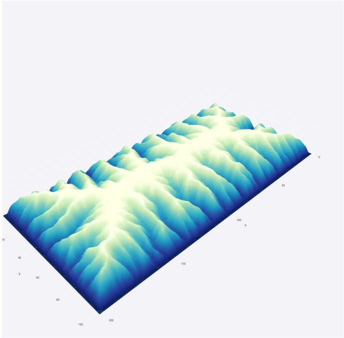
(b) Mountain predicted topography after 50 % evolution time

Figure 24: Mountain evolution over 1 000 000 years.

Results



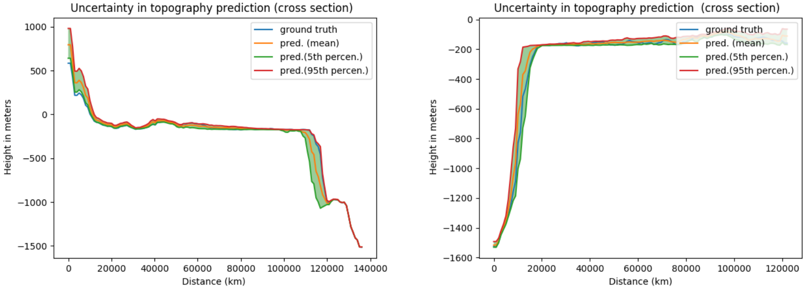
(a) Mountain predicted topography after 75 % evolution time



(b) Mountain predicted topography after 100 % evolution time

Figure 25: Mountain evolution over 1 000 000 years.

Results

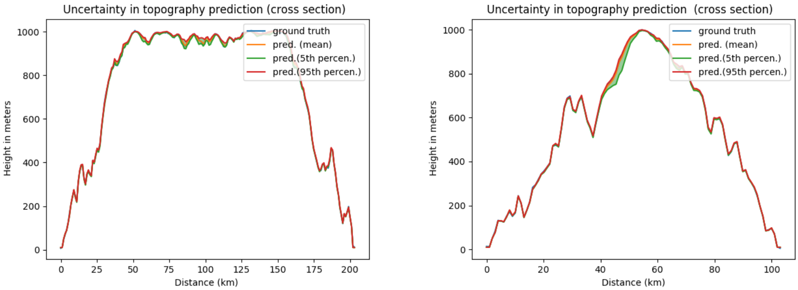


(a) Elevation cross-section along x-axis

(b) Elevation cross-section along y-axis

Figure 26: CM-extended: Cross-section of the elevation along the x-axis comparing the ground-truth of the Badlands-model with the PT-Bayelands predictions.

Results

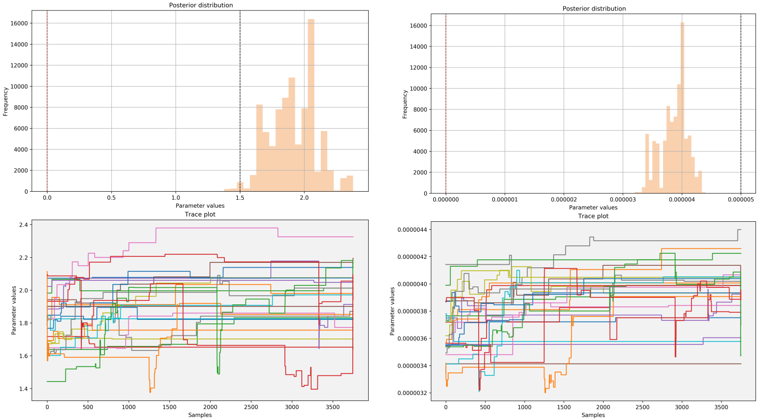


(a) Cross-section along the x-axis.

(b) Cross-section along the y-axis

Figure 27: Typical topography for the evolved Mountain problem, comparing PT-Bayeslands predictions and the Badlands ground-truth. Result is for 24 replicas, 100,000 samples.

Results



(a) CM-extended rainfall distribution

(b) CM-extended erodibility distribution

Figure 28: Typical posterior distribution (top panel) and replica traces (bottom panel) exploring the marine diffusion parameter-space for the CM-extended topography. In this model run 24 replicas are represented by the coloured lines. Parameter, c-marine, units are $m^2 a^{-1}$.

Results

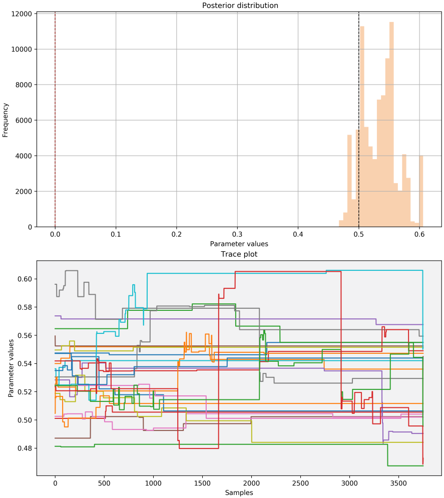


Figure 29: Typical posterior distribution (top panel) and replica traces (bottom panel) exploring the marine diffusion parameter-space for the CM-extended topography. In this model run 24 replicas are represented by the coloured lines. Parameter, c-marine, units are m^2a^{-1}.

Results - Comparison MCMC vs PT

Topography	Samples	Time(mins)	Sed RMSE	Elev RMSE	Pred RMSE	Accepted%
MCMC Bayeslands results						
Crater	10,000	136	8.24	1.06	9.30	2.35
Crater	100,000	1023	8.24	1.06	9.30	2.57
CM	10,000	101	459.85	67.37	527.22	0.47
CM	100,000	729	387.92	10.60	398.53	0.02
PT-Bayeslands using 24 replicas						
Crater	10,000	8	4.2	1.0	5.2	15.0
Crater	100,000	78	4.2	1.1	5.3	12.2
CM	10,000	7	7.4	25.3	32.7	4.9
CM	100,000	50	2.6	19.9	22.5	1.0

Table 6: Comparison of results of PT-Bayeslands with single-threaded MCMC Bayeslands shows significant difference in performance given computational time and accuracy in prediction.

Surface

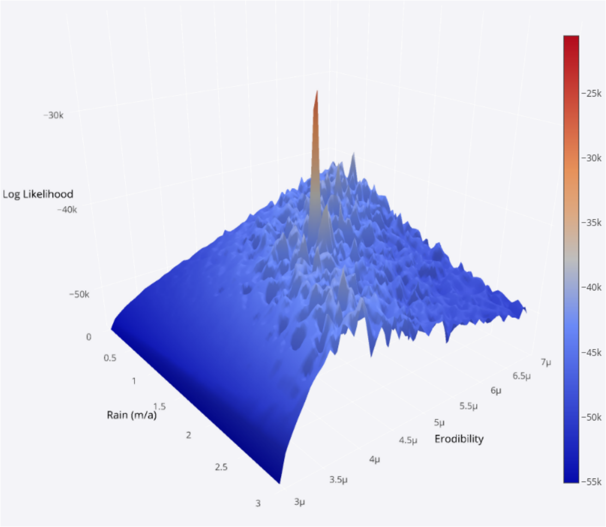


Figure 30: Likelihood surfaces of the Continental Margin topography for the rainfall and erodibility parameters only

Future work

- Surrogate assisted likelihood functions
- Robust sampling techniques such as parallel tempering that explore multi-modality.

Future work

- Way ahead - real world applications is to employ multi-core parallel tempering with Surrogates for Likelihood evaluations.
- Also exploring GPU based acceleration for Badlands. Challenge will be to harness the power of GPU with multi-core parallel tempering implementations.

Implications

- In future, it would be helpful to model region-based or time-varying rainfall distributions in Badlands.
- The distributions could be used to generate more information about the effects of climate change in geological timescales.

Fusion with optimisation methods

- Scope for optimisation methods such as evolutionary algorithms for models with hundreds of parameters
- Scope for fusion of evolutionary algorithms with MCMC methods such as Parallel Tempering for better proposals.
- Can be extended to other geophysical inversion problems.
- Gradient information from Forward Models can be useful for Bayesian Inference and Optimisation
- Library - software package.

Thanks

Project members: Danial Azam, Dietmar Muller, Sally Cripps, Tristan Salles, Ratneel Deo, and Nathaniel Butterworth
Special thanks: Sabin Zahirovic, Arpit Kapoor and Konark Jain

- Bayeslands [1]
- PT-Bayeslands [2]
- R. Chandra, D. Azam, R. D. Müller, T. Salles, S. Cripps, BayesLands: A Bayesian inference approach for parameter uncertainty quantification in Badlands, arxiv 2018 [3]
- R. Chandra, R. D. Müller, R. Deo, N. Butterworth, T. Salles, S. Cripps, Multi-core parallel tempering Bayeslands for basin and landscape evolution, arxiv 2018 [4]

---

[1] https://github.com/badlands-model/BayesLands
[2] https://github.com/badlands-model/paralleltemp_Bayeslands
[3] https://arxiv.org/abs/1805.03696
[4] https://arxiv.org/abs/1806.10939