Exploring Various Monte Carlo Simulations for Geoscience Applications

J. A. Rod Blais
Dept. of Geomatics Engineering
Pacific Institute for the Mathematical Sciences
University of Calgary, Calgary, Alberta T2N 1N4
blais@ucalgary.ca   www.ucalgary.ca/~blais

Introduction

- **Stochastic simulations are often critical in geoscience!**
- **Monte Carlo estimates are needed for direct and inverse problems**
- **Pseudorandom sequences imply simple quadrature computations**
- **Quasirandom (i.e. equidistributed) sequences offer alternatives**
- **Chaotic random sequences have been claimed to be superior**
- **Numerical experimentation is first required for analysis**
- **Variance reduction techniques can often improve results**
- **Geodetic and climate applications abound among others**
- **Investigations are continuing especially for predictions**

Randomness

- **In mathematics, only processes can be random!**
- **In physics, random usually means noncomputable or unpredictable**
- **In practice, there are various ways to simulate random sequences**
- **Pseudorandom sequences are commonly generated using some linear congruential model applied recursively, such as**

  $x_n \equiv c \odot x_{n-1}$ **modulo p** **(for large prime p and constant c)**

  or lagged Fibonacci congruential sequence, such as

  $x_n \equiv x_{n-u} \odot x_{n-v}$ **modulo p** **(for large primes p and u, v)**

  in which $\odot$ **usually stands for ordinary multiplication**
- **Quasirandom sequences are equidistributed (′random′) sequences**

  e.g. using digits from $\pi = 3.14159\ldots \Rightarrow \{0.1, 0.4, 0.1, 0.5, 0.9, \ldots\}$

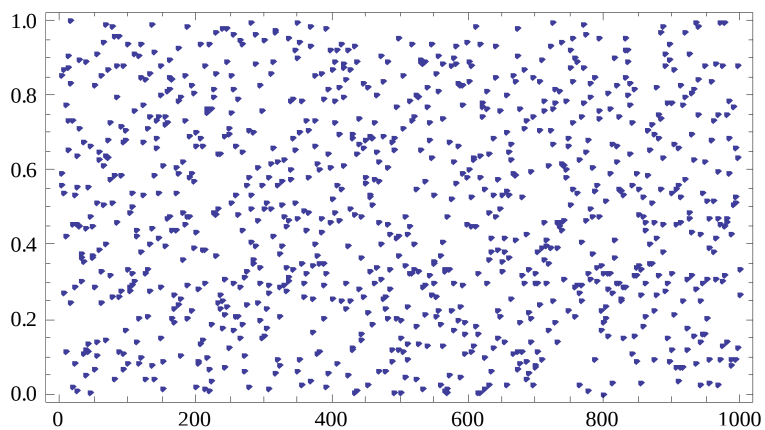
Chaos & Chaotic Randomness

- **Chaos refers to unstable dynamical nonlinear systems which are especially sensitive to their initial conditions**
- **Chaotic maps can be erratic, mixing / ergodic and thus ′random′**
- **The logistic map generated by $x_n = 4\, x_{n-1}\,(1-x_{n-1})$, $n = 1, 2, \ldots,$ for some seed x_0, over the interval (0, 1), exhibits randomness with an approximate density**

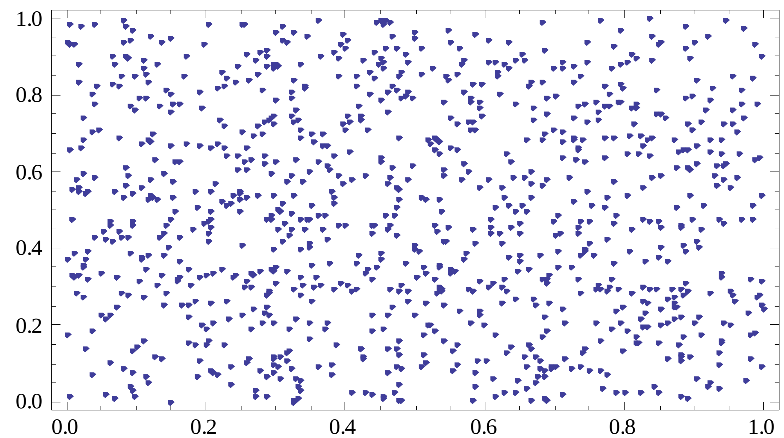
  $$\rho(x) = 1 / \pi\, [x\,(1 - x)]^{1/2}$$

  which needs to be taken into account in Monte Carlo applications
- **However, $x_n = \sin^2(2^n \theta)$ satisfies the Logistic Equation for any θ, $n = 0, 1, 2\ 3, \ldots,$ and according to Makila [2004], it is possible that $2^n\theta \to$ some integer thus making $x_n \to 0$ as $n \to \infty$ [Blais, 2010]**

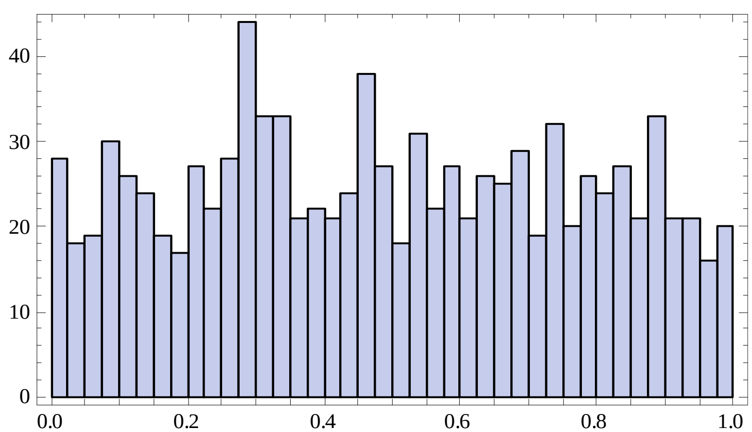
Pseudorandom Sequences



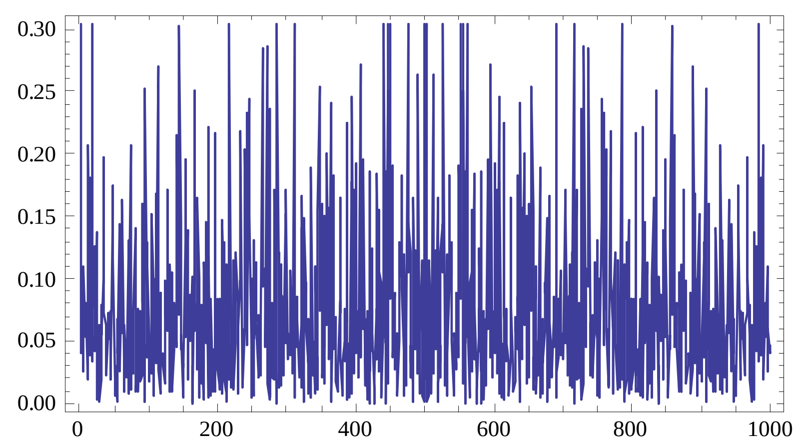
Spatial Plot of Pseudorandom Sequence



Phase Plot of Pseudorandom Sequence

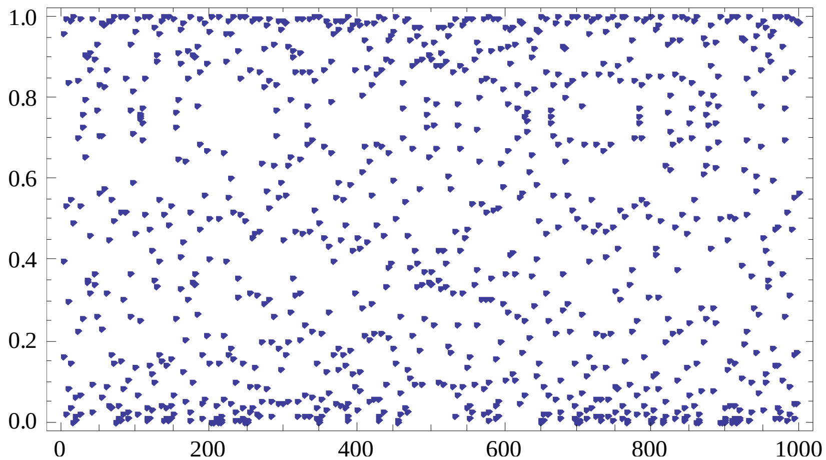


Histogram of Pseudorandom Sequence

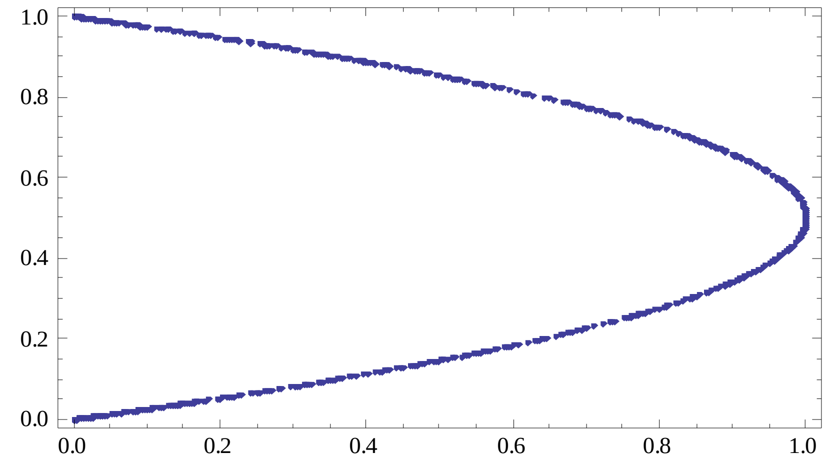


Periodogram of Pseudorandom Sequence

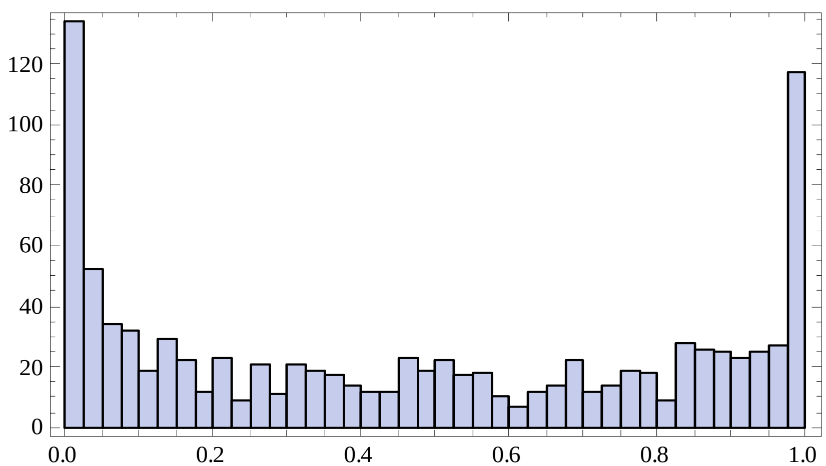
Chaotic Random Sequences



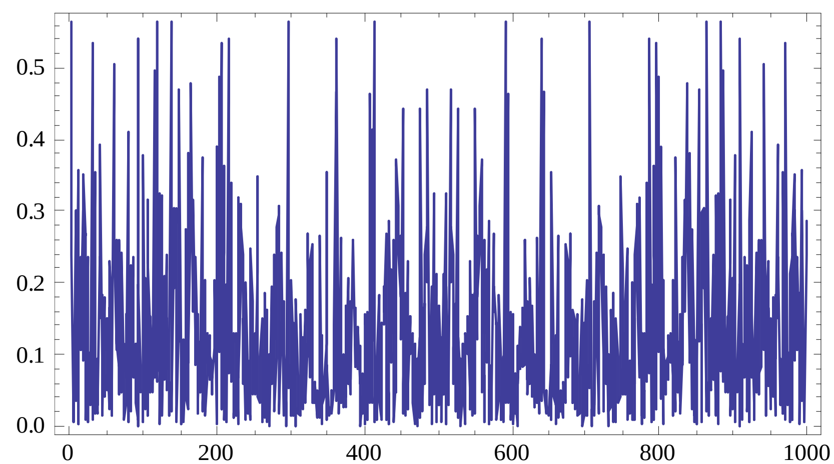
Spatial Plot of Chaotic Random Sequence



Phase Plot of Chaotic Random Sequence

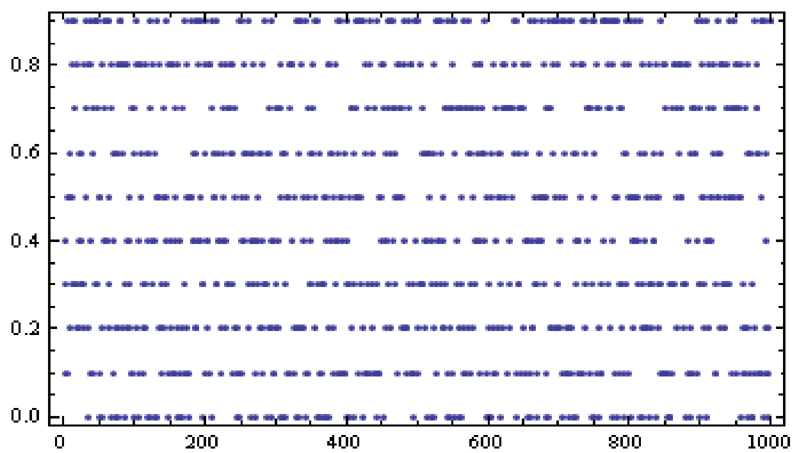


Histogram of Chaotic Random Sequence

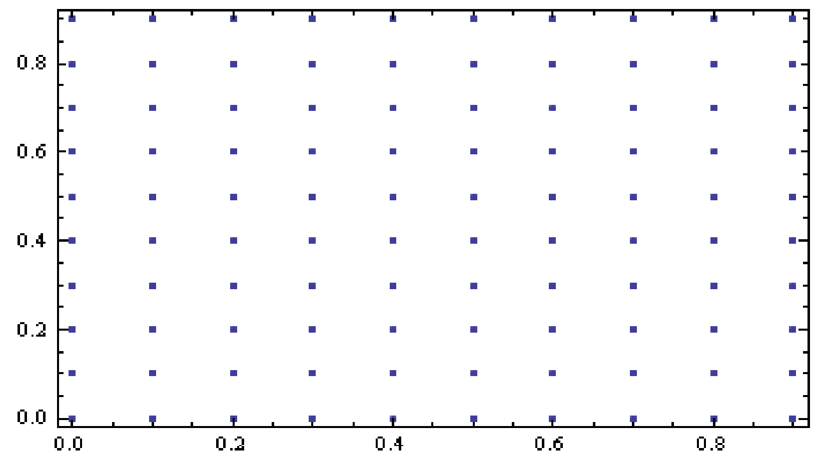


Periodogram of Chaotic Random Sequence

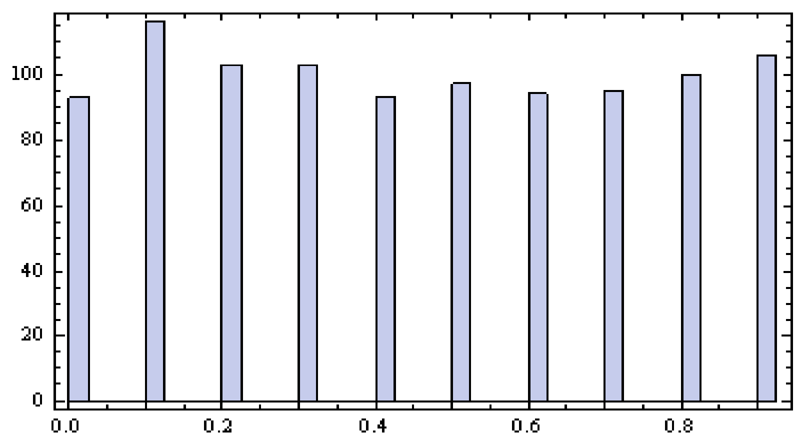
Quasirandom Sequences



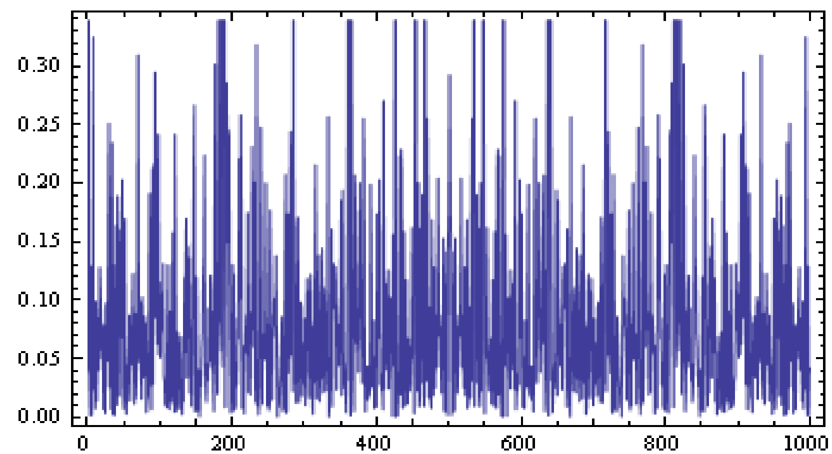
Spatial Plot of Digits of π



Phase Plot of Digits of π



Histogram of Digits of π



Periodogram of Digits of π

Monte Carlo Simulations

Numerical Recipes state:

$$\int_V f\, dV \approx V\langle f\rangle \pm \sqrt{\left(\langle f^2\rangle - \langle f\rangle^2\right)/N}$$ **implying a variance $O(N^{-1})$**

More specifically,

Random Number Generators	**Variance of Error**
Standard Pseudorandom Numbers	$\mathbf{O(N^{-1})}$
Quasirandom Numbers (General, spatial dim. s)	$\mathbf{O((\ln N)^{2s} N^{-2})}$
Chaotic Monte Carlo (General)	$\mathbf{O(N^{-1})}$
′Superefficient′ Chaotic Monte Carlo*	$\mathbf{O(N^{-2})}$

* Under the ′superefficiency condition′ implied by the dynamical correlation for large N, see e.g. [Umeno, 2000, 1999, 1998]

Numerical Experimentation

PMC / CMC / QMC	N = 10	N = 10^2	N = 10^3	N = 10^4
$\int_0^1 e^x dx$	1.54464560	1.60391781	1.67767504	1.70929627
	1.80241182	1.38343424	1.61749711	1.70409600
$\cong$ **1.718281828459045**	1.59818977	1.75556959	1.71391782	1.71511801
$\int_0^1 \int_0^1 e^{xy} dxdy$	1.25060309	1.26440568	1.29981050	1.31472135
	1.34889037	1.02949566	1.21680953	1.31020325
$\cong$ **1.317902151454404**	1.22984290	1.34958187	1.31403513	1.31556883
$\int_0^1 \int_0^1 \int_0^1 e^{xyz} dxdydz$	1.11214566	1.11346090	1.13706131	1.14409612
	0.94684567	0.95059967	1.12476063	1.16363260
$\cong$ **1.146499072528643**	1.09371629	1.16785666	1.14383721	1.14510484

Analysis of Simulations

Pseudorandom Monte Carlo (PMC) Approach:

- **Using Mathematica 7 random number generator**
- **Very good results in general of $O(N^{-1})$**

Chaotic Random Monte Carlo (CMC) Approach:

- **Using Logistic Map with corresponding density correction**
- **Results generally comparable to pseudorandom results**

Quasirandom Monte Carlo (QMC) Approach:

- **Using π digits, these specific results are surprisingly good…**
- **In general, more investigations are required to confirm this!**

Variance Reduction

In general:

- **Uniformity appears generally more important than randomness**
- **PMC and CMC results can often be improved thru variance reduction**

Importance Sampling Strategy:

- **Variable of integration may be transformed for better results**
- **Significant improvements are possible with complex problems**

Stratified Sampling Strategy:

- **Domain of integration may be partitioned for different sampling**
- **Small sample means often contribute to better overall estimates**

Geodetic Applications

Direct Problem: Gravimetric terrain corrections at the origin:

$$\delta g(0,0,0) = G\bar{\rho}\int_{-L}^{L}\int_{-L}^{L}\int_{0}^{H(x,y)} \frac{z\,dz\,dy\,dx}{(x^2+y^2+z^2)^{3/2}}$$

then for N small prisms over an area A,

$$\delta g(0,0,0) \approx G\bar{\rho}\ A \left\langle \int_0^h \frac{z\,dz}{(d^2+z^2)^{3/2}} \right\rangle \approx \frac{G\bar{\rho}A}{N}\sum_{i=1}^{N}\left(\frac{1}{d_i} - \frac{1}{\sqrt{d_i^2+h_i^2}}\right)$$

which is very appropriate for LIDAR and similar dense terrain data [Blais, 2010]

Inverse Problem: Recovery of ocean bathymetry from surface gravity data and disturbances using Simulated Annealing computations [Blais et al, 2008]

Concluding Remarks

- **Pseudorandom Monte Carlo simulations generally give results of $O(N^{-1})$**
- **Quasirandom Monte Carlo results using digits from π are most surprising!**
- **Chaotic Monte Carlo limited experimentation shows no better than $O(N^{-1})$**
- **Makila [2004] results imply that the Logistic Map is not always appropriate**
- **Variance reduction strategies can improve results from $O(N^{-1})$ to $O(N^{-2})$**
- **Gravimetric terrain corrections using LIDAR data are very efficient & useful**
- **Research and computational experimentation are continuing for gravity terrain corrections and uncertainty characterization especially for climate change applications such as e.g. in hydrology [Mutulu and Blais, 2010]**