A Review of the Merits of the Stochastic Dynamic Equations and the Monte Carlo Approach in Modeling and Understanding Systems

Rex J. Fleming

Global Aerospace, LLC

1. Introduction

One can deal with the uncertainty in initial conditions of a numerical model by two different approaches. The Monte Carlo (MC) approach use a randomly chosen set of initial states, usually obtained from a random Gaussian distribution, to evaluate the impact of uncertainty in initial conditions for the time integration of a physical system. The Stochastic Dynamic Equations (SDE) approach begins with an infinite ensemble of initial states represented simply by the variance of the initial conditions. The MC is thus an approximation of the SDE method. However, the SDE has a closure issue. Time derivatives of second moments also requires knowledge of third moments, time derivatives of third moments involves fourth moments, and so on.

This review will provide a simple example of the two methods for the original Lorenz chaos equations. This equation set can produce a simple fixed point solution or a complex chaotic solution – examples of both solutions with both methods will be demonstrated.

There is a method of using **both approaches t**o fully dissect the nonlinearity of a system of equations. This is rarely discussed, but demonstrates the power of the stochastic dynamic equation set. That example will be reproduced here.

2. The Monte Carlo Approach

For large model applications such as weather and climate models, the MC method is the only practical approach. The proper programming of Monte Carlo models can make maximum use of the current multi-processor computers. A simple demonstration of the MC method for the Lorenz (1963) equations:

$$\mathbf{X^{\bullet} = P\,(Y - X)} \qquad \mathbf{X(1)^{\bullet} = P\,(X(2) - X(1))} \qquad \mathbf{(1)}$$

$$\mathbf{Y^{\bullet} = -X\,Z + R\,X - Y} \qquad \mathbf{X(2)^{\bullet} = -X(1)\,X(3) + R\,X(1) - X(2)} \qquad \mathbf{(2)}$$

$$\mathbf{Z^{\bullet} = X\,Y - B\,Z} \qquad \mathbf{X(3)^{\bullet} = X(1)\,X(2) - B\,X(3)} \qquad \mathbf{(3)}$$

where a (•) denotes a time derivative, and where **X, Y, Z** have been replaced with **X(1), X2), X(3)** to make subsequent notation somewhat easier to follow; and where **P = 10, B = 8/3,** and where stable fixed point (FP) solutions occur for **R < 24.74 and chaos occurs for R $\geq$ 24.74.**

Numerical integration of **(1) – (3)** with a 4th order Runge-Kutta scheme requires evaluation of the right hand side (rhs) four times per single time step. Before each evaluation of the rhs, the **current X's are put in XT** and the p**redicted values put in XP.** For the MC solutions of **(1) – (3)** in FORTRAN one has:

```
DO 100 J = 1, K      { where K may be 100, 1000, 40,000, or some larger integer }
XP(J,1) = P * (XT(J,2) - XT(J,1))
XP(J,2) =  - XT(J,1) * XT(J,3) +  R * XT(J,1) - XT(J,2)
XP(J,3) = XT(J,1) * XT(J,2) - B * XT(J,3)
100 continue
```

If one wanted to use 40,000 slightly different initial conditions to examine the error growth in these equations, then with as many as 40,000 processors, one could be performing the same calculations simultaneously. Thus, the MC approach is the only practical way to deal with uncertainty in large scale grid models. Future multi-processor computers will allow even larger sample sizes.

Figure 1 shows a **deterministic** solution of **(1) – (3)** with **R = 14** (stable FP solution) with **no uncertainty** in the initial conditions of [**0, 1, 0] for [X1, X2, X3]. The results match the theory that X(3) = R - 1 = 13 and the theory says that X(1) = X(2) = ± 5.888; but the result here is - 5.888.**

3. The SDE Approach

The SDE approach and the 2nd moments were introduced into atmospheric science by Epstein(1969). The 3rd moment equations were introduced and discussed by Fleming (1971). The general form of a deterministic set is given by:

$$X_i^{\bullet} = \sum_{p,q} a_{ipq} X_p X_q - \sum_{p} b_{ip} X_p + c_I$$

The SDE equations follow:

$$\mu_i^{\bullet} = \sum_{p,q} a_{ipq} (\mu_p \mu_q + \sigma_{pq}) - \sum_{p} b_{ip} \mu_p$$

$$\sigma_{ij}^{\bullet} = \sum_{p,q} a_{ipq} (\mu_p \sigma_{jq} + \mu_q \sigma_{jp} + T_{jpq}) + a_{jpq} (\mu_p \sigma_{iq} + \mu_q \sigma_{ip} + T_{ipq}) - \sum_{p} (b_{ip} \sigma_{jp} + b_{ip}\sigma_{ip})$$

$$T_{jkl}^{\bullet} = \sum_{p,q} a_{jpq} (\mu_p T_{klq} + \mu_q T_{klp} - \sigma_{pq} \sigma_{kl} + \lambda_{klpq}) + a_{kpq} (\mu_p T_{jlq} + \mu_q T_{jlp} - \sigma_{pq} \sigma_{jl} + \lambda_{jlpq}) + a_{lpq} (\mu_p T_{jkq} + \mu_q T_{jkp} - \sigma_{pq} \sigma_{jk} + \lambda_{jkpq}) - \sum_{p} (b_{jp} T_{klp} + b_{kp} T_{jlp} + b_{lp}T_{jkp})$$

where the μ(i) = the mean of X(i), σ (i,j) = the covariance of X(i) X(j) , T(i,j,k) = the 3rd moment about the mean, and λ (i, j, k, l) is the 4th moment about the mean.

The SDE approach has one significant **advantage.** The method allows a perfect blend of true physics and mathematical statistics. The SDE method can quantify the evolution of uncertainty in physical system in terms of the energetics of the physical system and the 'uncertain energy' of error growth (see Fleming, 1971).

The SDE approach has a significant **disadvantage**. The number of equations multiplies rapidly. Only a few stochastic equations will be displayed here, but the general form above can be used to arrive at the precise stochastic form for any deterministic system.

Just a few of the SDE equations for **(1)-(3)** are:

$$\mu(1)^{\bullet} = P\ (\mu(2) - \mu(1)) \qquad (4)$$

$$\mu(2)^{\bullet} = -\mu(1)\,\mu(3) - \sigma(1,3) + R\,\mu(1) - \mu(2) \qquad (5)$$

$$\mu(3)^{\bullet} = \mu(1)\,\mu(2) + \sigma(1,2) - B\,\mu(3) \qquad (6)$$

$$\sigma(1,1)^{\bullet} = 2\,P\,(\sigma(1,2) - \sigma(1,1)) \qquad (7)$$

$$\sigma(2,2)^{\bullet} = -2\,[\,\mu(1)\,\sigma(2,3) + \mu(3)\,\sigma(1,2) + T(1,2,3) - R\,\sigma(1,2) + \sigma(2,2)\,] \qquad (8)$$

$$\sigma(3,3)^{\bullet} = 2\,[\,\mu(1)\,\sigma(2,3) + \mu(2)\,\sigma(1,3) + T(1,2,3) - B\,\sigma(3,3)\,] \qquad (9)$$

4. Comparison of MC, SD2, and SD3 with Lorenz set with R = 14

If one applies the **SD2** equation set (**above equations, but with third moments and above assumed to be zero)** and uses uncertainty in the variables given by an initial variance of **σ(1,1) = σ(2,2) = σ(3,3) = 0.1, then one gets the result shown in Figure 2.** The results show that **X(3) = R -1 = 13** is correct, but the results for **X(1) = X(2) = 0.0,** not **- 5.888. The theory says ± 5.888, so perhaps the SD2 result is correct.**

The MC result with the same variance used in **SD2** is shown in **Figure 3**. The sample size is 40,000 so this **is** the correct statistical result. The results for **X(3) = 13** and for **X(1) = X(2) = -2.5** are correct. Thus, the **SD2** result has failed to capture the negative shift.

Figure 4, shows the **SD3 r**esult with 3rd moments included and a proper closure. The results are correct for both the **X(3) = 13 and X(1) = X(2) = -2.52.** There is another important attribute of the stochastic dynamic equations. For R = 14, a fixed point solution in the Lorenz equations, the corresponding **SD3** equations have the time tendency of the statistical moments go to zero, so from **(4) μ(1) = μ(2)**, from **(7)** one obtains **σ(1,1) = σ(1,2),** and combining these results with **(6)** one obtains the variance of **X(1) = σ(1,1) =** $\mathbf{B\,X(3) - X(1)^2 = (8/3)\,(13) - (2.52)^2 = 28.32}$. **Figure 5** indicates X(3) versus the variance of X(1) and indeed this answer is correct.

Since **X(3)** becomes a constant **= R -1 ,** all 2nd moments and higher with a "3" as an index become zero. The eq. for $\sigma(1,2)^{\bullet} = 0$ (not shown) gives **σ(2,2) = σ(1,1) = 28.32. Figure**

6 indicates the comparison of the MC and **SD3** results for the variance of **X(2) = σ(2,2)** versus time. One sees a very close agreement between the two approaches – both for **σ(2,2)** and for the **explosive randomness** that occurs even in this fixed point solution.

5. Comparison of MC and SD3 with Lorenz set with R = 28

Figure 7 shows the deterministic solution of **(1) – (3) for R = 28.** The strange attractor appears to be a formidable challenge for the SDE. **Figure 8 indicates two problems that appear in the MC results for this important third moment.** Shown is **T(1,2,3)** (equal to zero with **R = 14**) versus time (number of iterations).**T(1,2,3**) is the 3^{rd} moment about the mean involving all three variables. One first sees the explosive randomness, far greater than in the fixed point solution, then a modulation about zero for 1500 iterations, then an excursion upward to a modulation of about 300 at 2000 iterations, and finally a **turbulent jostling** about 200 from 3000 to 8000 iterations. **There is the first problem of the explosive randomness, then the determination of the final result.**

The **SD2** cannot pass the first problem of explosive randomness. This solution will blow up. Thus the **SD3** will be used to check this and other variables in the set. The details of the closure for SD3 are included in a second paper at this Epstein Symposium. It will simply be stated here that a closure can be found. **X(3) is no longer a constant but an integral part of the chaos.** The mean values of **X(1) and X(2)** are both equal to zero, and their statistics are symmetric. The symmetry implies that moments (2^{nd}, 3^{rd}, or 4^{th}) with an **odd** number of 1's and 2's in the indices will be equal to zero. Those with an even number will be non-zero and likely quite large.

The results shown in **Figure 9** are for **X(3)** as a function of time. The initial explosive randomness is quite large in the MC run. The **SD3** has captured that initial randomness and the final result of the mean value for **X(3) = 23.55** achieved by both solutions.

6. A complete statistical picture of the Lorenz strange attractor

The results of **Figure 9** are encouraging, but the initial explosive randomness which appears in **Figs. 8 and 9** appears to present a challenge to completely understand the statistical properties of the strange attractor. While moments can be computed via the MC approach, how are the moments related to each other?

When a system is bounded and dissipative as the Lorenz system, all trajectories eventually tend toward some bounded set of zero volume in phase space. Knowing this, it was decided to solve the equations with a MC approach, flooding the attractor with a very large sample size. To handle the **turbulent jostling** seen in **Figure 8,** it was determined that a suitable time averaging of the MC samples would also be required.

Table 1 indicates the results of varying the sample size and varying the number of consecutive iterations used in a time average. The variable shown is **T(1,1,3)** which has a time mean of **400.6.** Beyond a critical time averaging period, there was little difference found in the time variance for a **very large sample size.** The results to be shown are for the case of a sample size of 40,000 and a time average over the last 4000 iterations of a run to 16,000 iterations.

All the statistical moments (e.g., up through fourth moments) can be computed from such a MC calculation at each time step. Though many calculations are required, those moments after the time averaging, are shown in Table 2 and 3. They become stable and unchanging. There is another advantage of using **both methods** to arrive at fundamental conclusions concerning this system. A very important fact is that the full **SDE** equations -- written out up through third moments, with the **fourth moment terms left in the right hand side (without any assumptions about those moments,) provide exact relationships.** With moments unchanging, the left hand side of a full moment equation is equal to zero. There are derivable dynamic relationships between the moments that are **exact!** These are listed in Tables 2 and 3 as **calculated MC** values and **computed SDE** values. The values from the MC averages in those Tables are approximate and could be made even closer with a larger sample size.

One need not have made all those calculations! Armed with just two results from the MC calculations, **X(3) = 23.55 and its variance σ(3,3) = 74.34** one has everything required to derive all the moment values shown in Tables 2 and all but two of the 4^{th} moment values in Table 3 – **one needs only paper, pencil, and a hand calculator.**

Only a few examples are permitted here. From **(4)** set to zero one has **μ(1) = μ(2)**; from **(5)** set to zero one has: **μ(1) [μ(3) + R -1] = 0, but μ(3) ≠ R - 1 (unlike for R = 14), thus μ(1) = μ(2) = 0.**

From **(7) and (6) set to zero,** one has **σ(1,1) = σ(1,2) = B X(3) = (8/3) (23.55) = 62.8.**
From **(9) set to zero**, one has **2 [μ(1) σ(2,3) + μ(2) σ(1,3) + T(1,2,3) - B σ(3,3)] = 0 or T(1,2,3) = B σ(3,3) = (8/3) (74.34) = 198.2.**

All of these values agree with the Table results. One can continue to use one result to build upon another, and many other interesting relationships can be found. Another non-anticipated result is the relation involving all three variances (which can be deduced from the equations and verified from the Table results):
σ(2,2) = σ(1,1) [R - μ(3)] – B σ(3,3).

7. Summary

The MC approach is valuable today and will become even more so in the future with further advances in multi-processor computers. Such MC predictions will be more accurate, and have the capability to roughly measure the uncertainty associated with that prediction. Our socio-economic system must deal with both accuracy and uncertainty.

The full SDE represent a perfect blend of physics and statistical relationships. They offer greater insight as a research tool. Together with the MC approach, this allows one to delve more deeply into the nonlinear nature of physical systems. Thank you, Edward S. Epstein for your contributions!

References

Epstein, E. S.,1969: Stochastic dynamic prediction. *Tellus*, **21**, 737-757.

Fleming, R. J., 1971: On stochastic dynamic prediction: I. The energetics of uncertainty and the question of closure. *Mon. Wea. Rev.*, **99,** 851-872.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20,** 130-142.

Tables

Size	# of consecutive iterations used in time average		
	250	1000	4000
100	6931	8302	9470
200	1812	1765	1856
500	1011	831	859
1000	453	388	416
2000	258	248	221
5000	89	91	96
40000	8	10	11

Table 1. Time variance of T(1,1,3) as a function of sample size and number of consecutive iterations used in time average

Moment	Var #	**MC Value**	**SD3 Value**	Moment	Var #	**MC Value**	**SD3 Value**
μ(1)	**X(1)**	**-.001**	**.000**	**T(1,1,1)**	**X(10)**	**.060**	**.000**
μ(2)	**X(2)**	**-.001**	**.000**	**T(1,1,2)**	**X(11)**	**0.060**	**.000**
μ(3)	**X(13)**	**23.55**	**23.55**	**T(1,1,3)**	**X(12)**	**400.6**	**400.6**
σ(1,1)	**X(4)**	**62.8**	**62.8**	**T(1,2,,2)**	**X(13)**	**.040**	**.000**
σ(1,2)	**X(5)**	**62.8**	**62.8**	**T(1,2,3)**	**X(14)**	**198.2**	**198.2**
σ(1,3)	**X(6)**	**-.005**	**-.000**	**T(1,3,3)**	**X(15)**	**-.080**	**-.000**
σ(2,2)	X(7)	**81.20**	**81.20**	**T(2,2,2)**	**X(16)**	**.009**	**.000**
σ(2,3)	X(8)	**.001**	**-.000**	**T(2,2,3)**	**X(17)**	**84.83**	**84.83**
σ(3,3)	X(9)	**74.34**	**74.34**	**T(2,3,3)**	**X(18)**	**-.060**	**-.000**
				T(3,3,3)	**X(19)**	**132.4**	**132.4**

Table 2. Calculated MC values and computed SD3 values from full equations for R = 28

Moment	Var #	**MC Value**	**SD3 Value**	Moment	Var #	**MC Value**	**SD3 Value**
λ(1,1,1,1)	X(20)	9060.1	9060.1	**λ(1,2,2,3)**	**X(27)**	**-.15**	**0.0**
λ(1,1,1,2)	**X(21)**	**9060.2**	**9060.1**	**λ(1,2,3,3)**	**X(28)**	**5021.5**	**5021.6**
λ(1,1,1,3)	**X(22)**	-.14	**0.0**	**λ(1,3,3,3)**	**X(29)**	**-1.2**	**0.0**
λ(1,1,2,2)	**X(23)**	**10735**	**10735**	λ(2,2,2,2)	X(30)		
λ(1,1,2,3)	**X(24)**	**-.003**	**0.0**	λ(2,2,2,3)	X(31)		
λ(1,1,3,3)	**X(25)**	**6712.5**	**6713.0**	λ(2,2,3,3)	X(32)		
λ(1,2,2,2)	**X(26)**	**13774**	**13774**	λ(2,3,3,3)	X(33)		
				λ(3,3,3,3)	X(34)		

Table 3. Calculated MC values and computed SD3 values for R = 29.
The 4th moments X(20) and X(30 through X(34) are not required in the SD3.

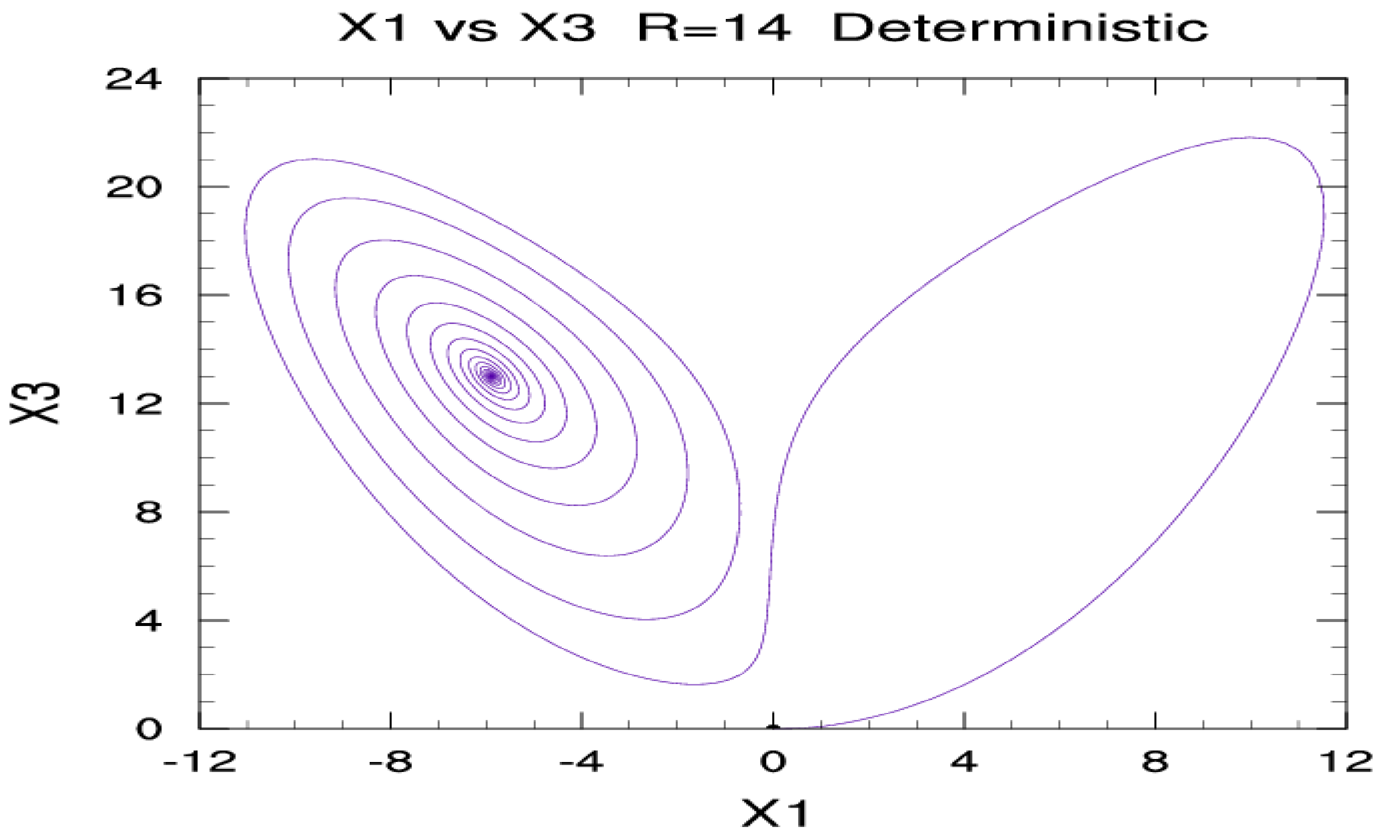

Figure 1. Initial conditions are [0,1,0] for [X1, X2, X3] R = 14

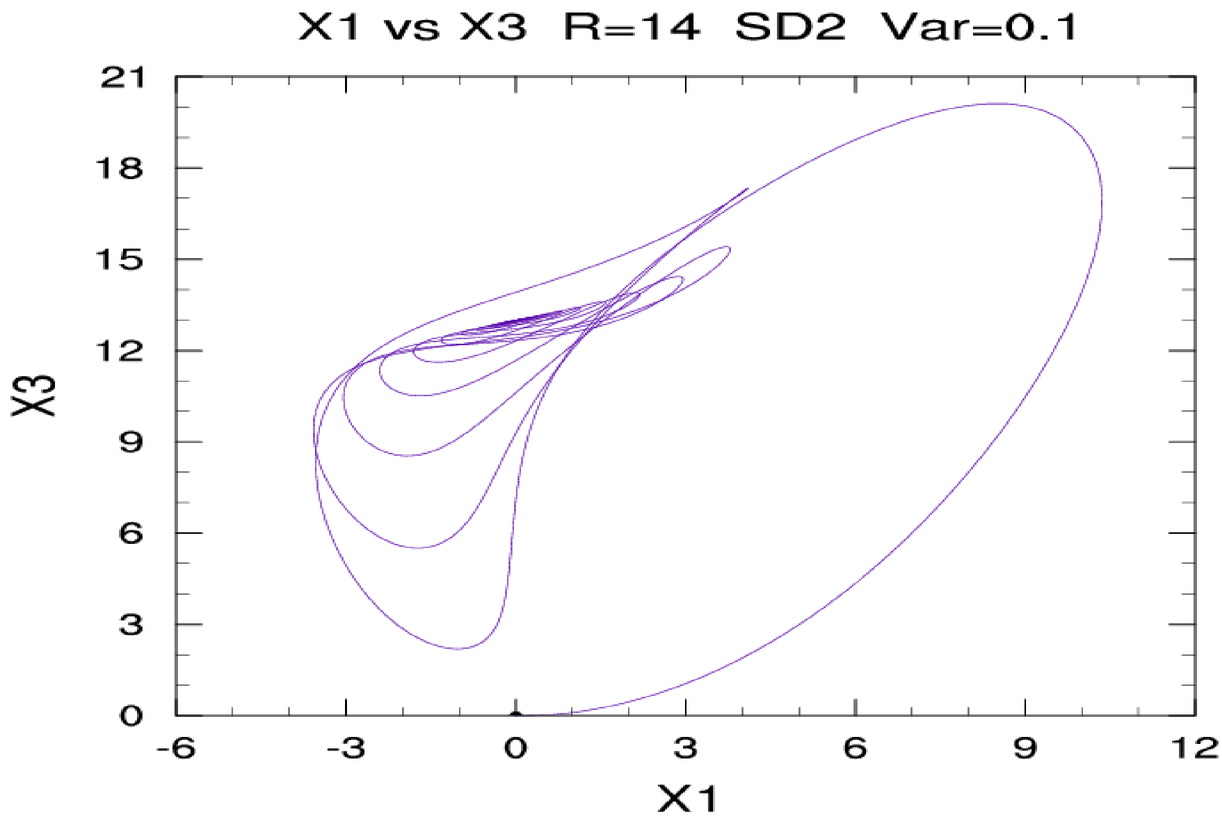

Figure 2. SD2 initial conditions as Fig 1; variance of each variable = 0.1

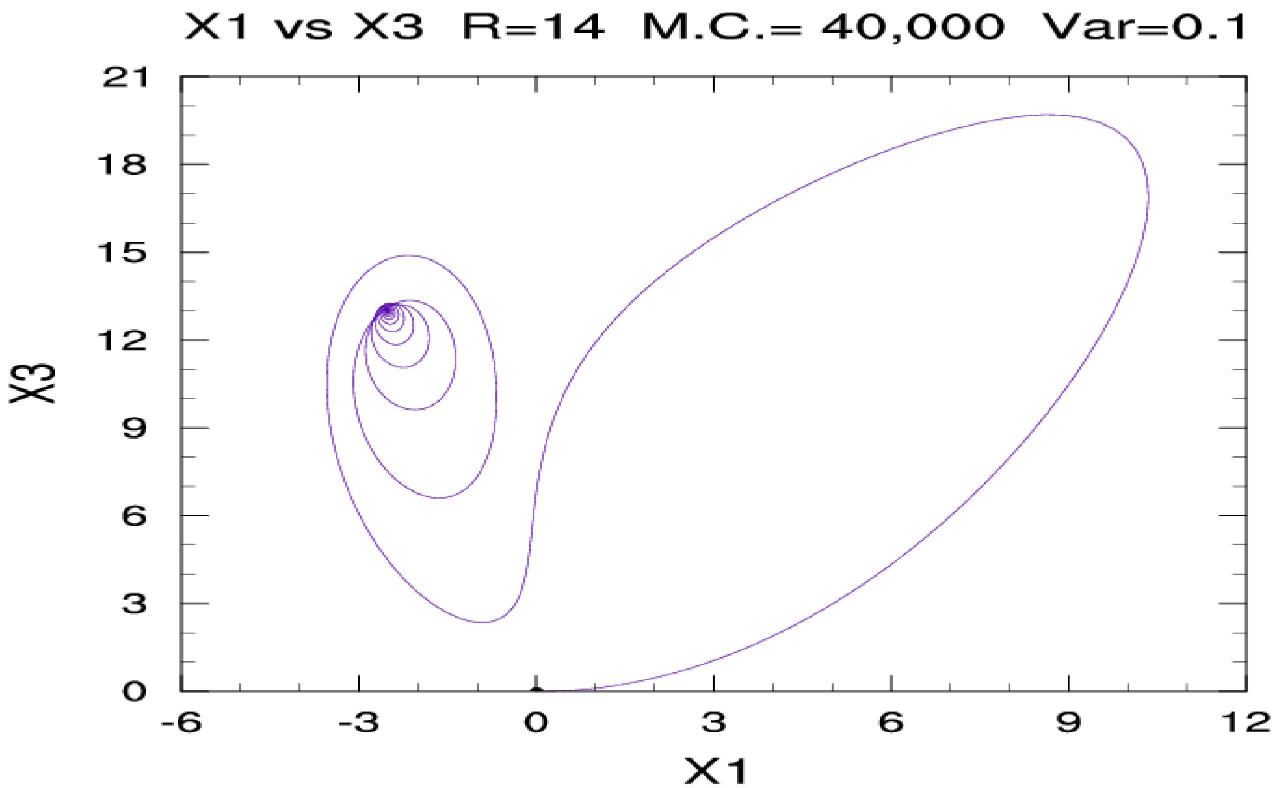

Figure 3. MC: same initial conditions as SD2; X1 = X2 = – 2.52

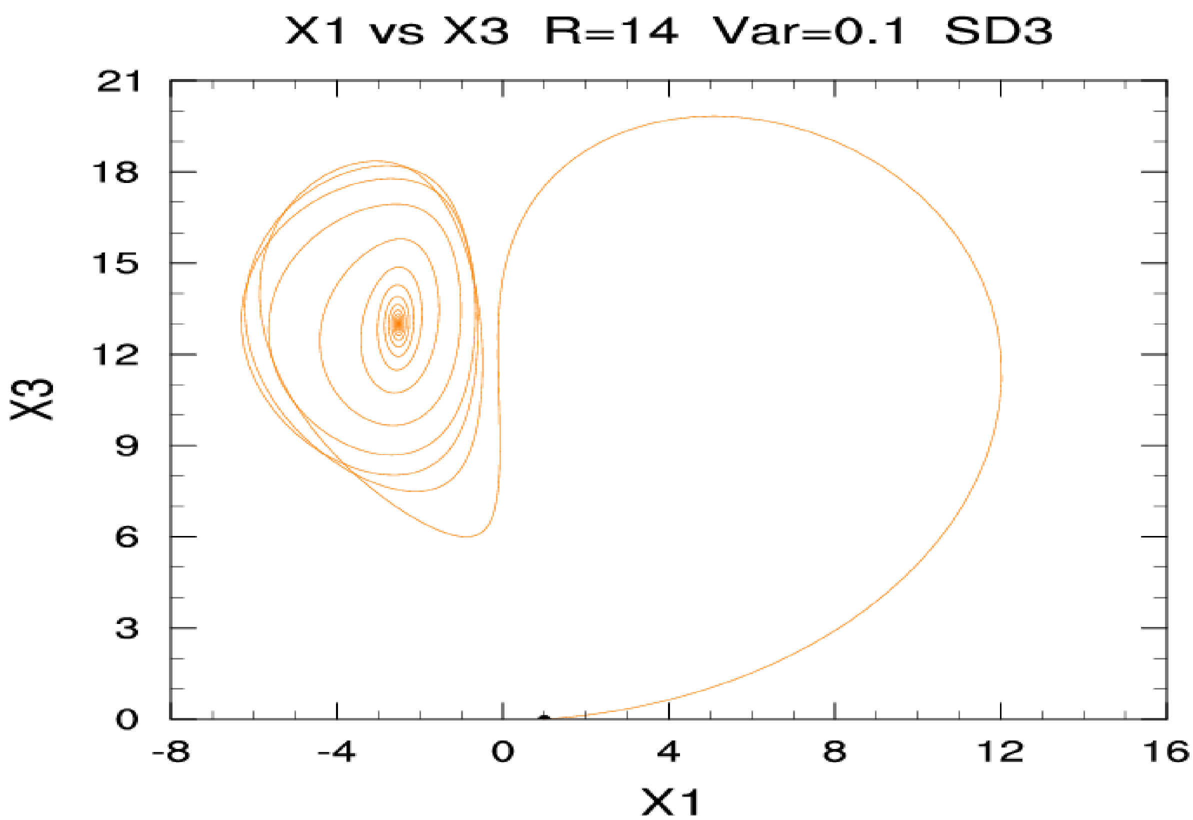

Figure 4. SD3: correct X3 = 13; also correct for X1 = X2 = – 2.52

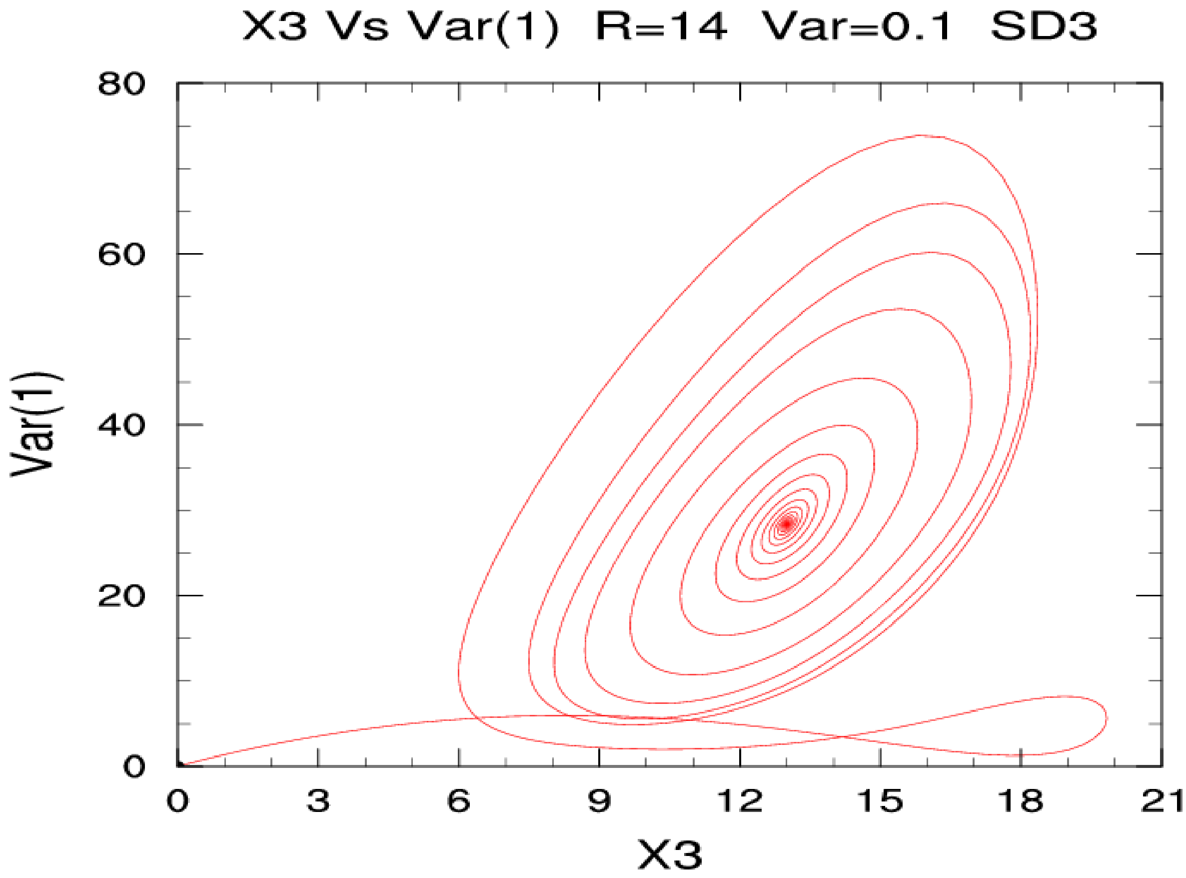

Figure 5. Correct variance of X1 = 28.32; matches SD3 derived value

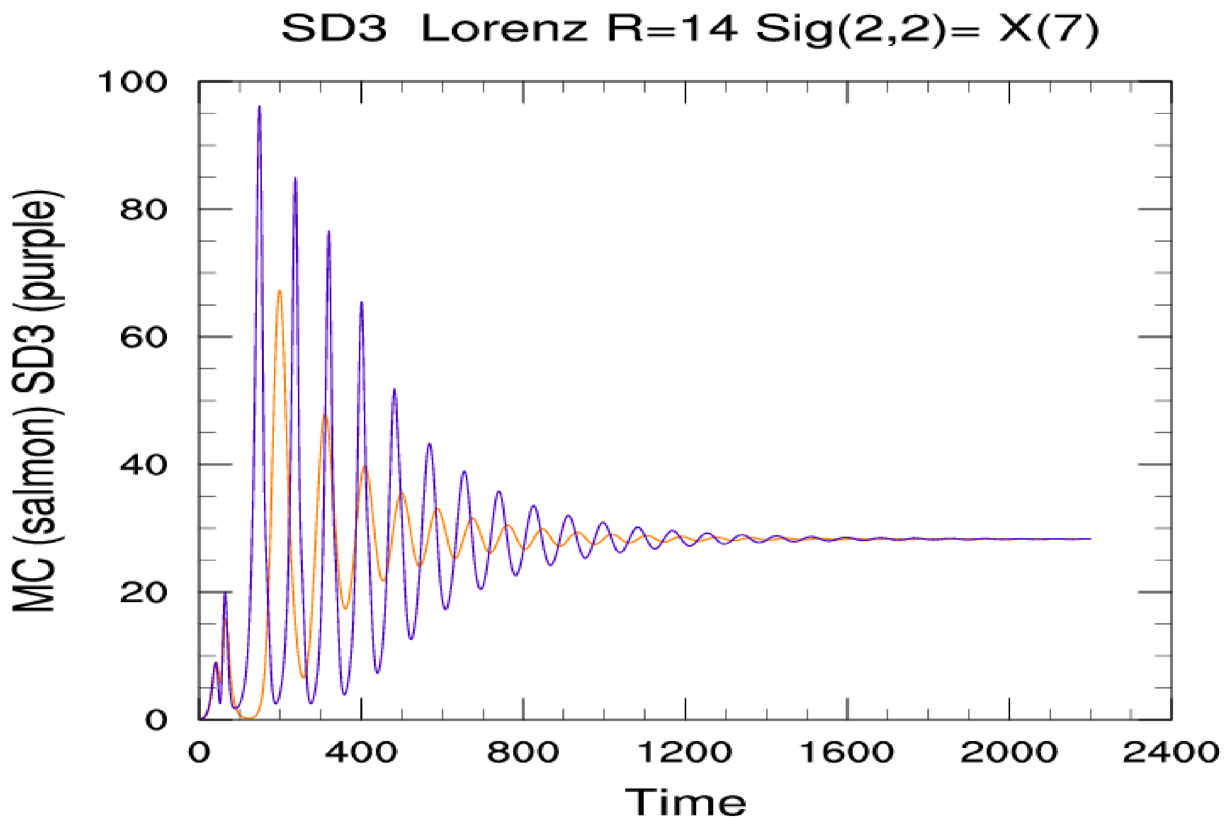

Figure 6. MC and SD3 agree on both initial "explosive randomness" and final value of σ(2,2) = 28.32

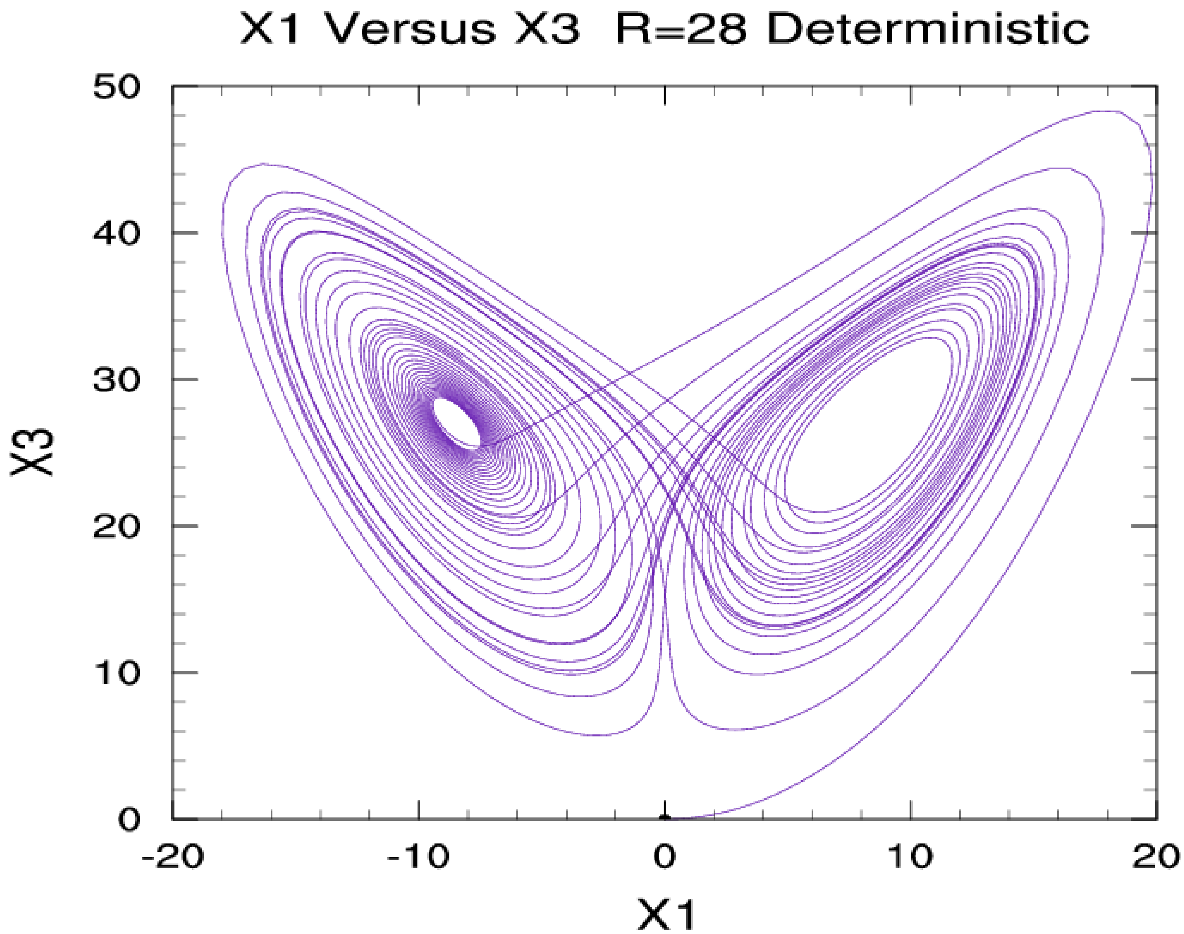

Figure 7 Deterministic solution [0,1,0] R = 28. A formidable challenge!

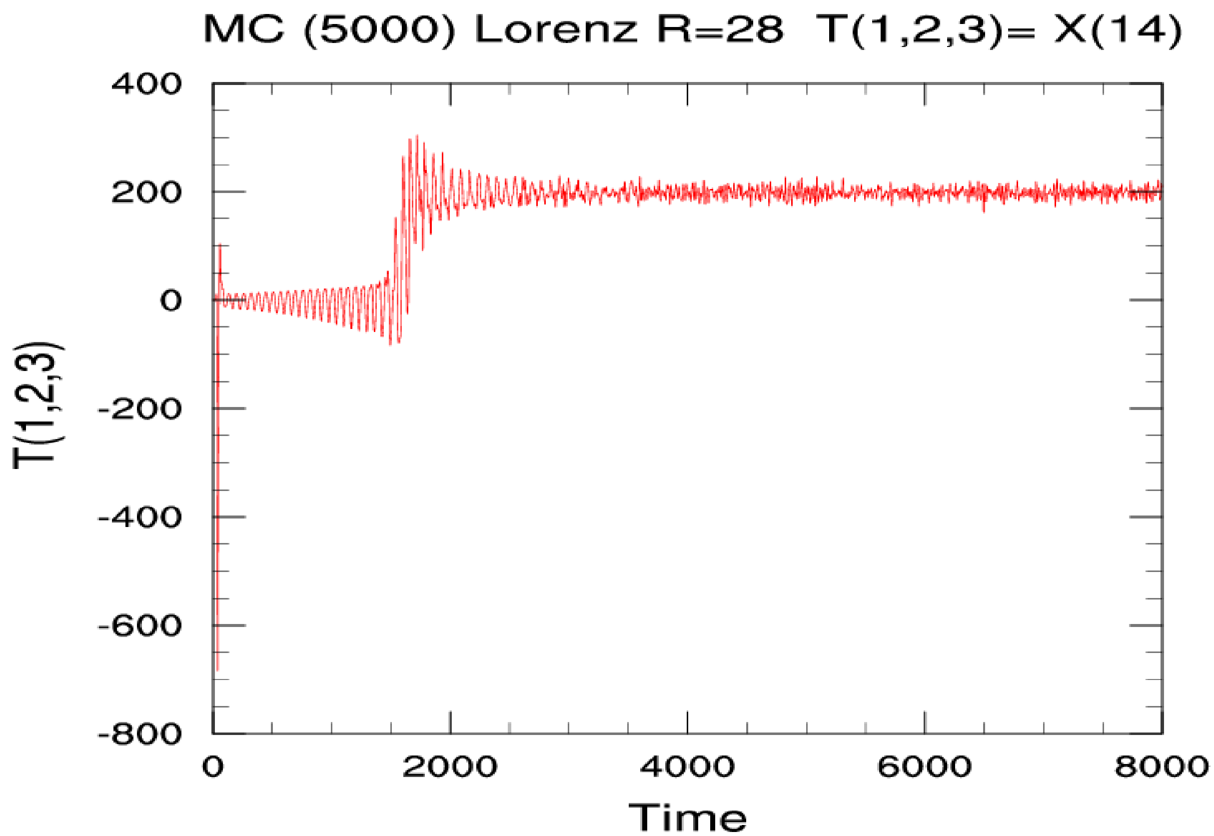

Figure 8. T (1,2,3) jostling around 200 with R = 28 [This = 0.0 for R 14]

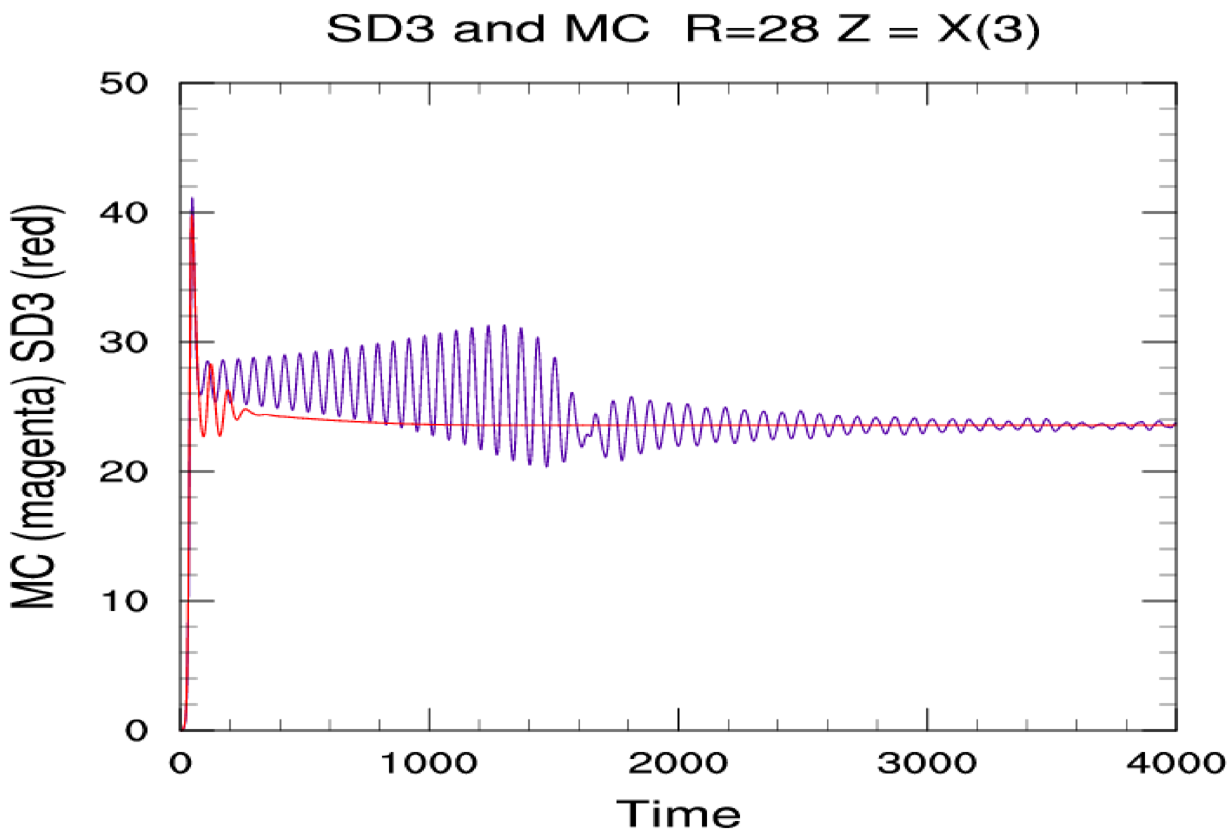

Figure 9. X(3) as a function of time: MC and SD3 in excellent agreement