Estimation of the Hurst Exponent for the Burdekin River using the Hurst-Mandelbrot Rescaled Range Statistic[1]

Sarah Millen and Rodney Beard
s370766@student.uq.edu.au and r.beard@economics.uq.edu.au
School of Economics
The University of Queensland

Abstract

The issue of the impact of agricultural run-off in north Queensland on nutrient accumulation in the Great Barrier Reef lagoon has recently been the subject of increased scientific, media and policy debate. Models of river flow and of nutrient plumes at the mouth of Queensland rivers are an essential prerequisite for studying the impact of agriculture on the Great Barrier Reef ecosystem. Hurst found long-term dependency in fluctuations at the mouth of the Nile River the theoretical basis of which was later found by Mandelbrot to involve fractional Brownian motion. Approximately one quarter of the inflow into the Barrier Reef lagoon comes from the Burdekin and Fitzroy region. Using a data set collected by Isdale et al. that covers the period 1644-1980, that is based on reconstructing river flow data for the Burdekin river by analysing luminescence lines in coral reefs at the river mouth, we have estimated the Hurst exponent of the Burdekin river for the purpose of developing a river flow simulation model based on fractional Brownian motion.

Introduction

The impact of terrestrial run-off on the Great Barrier Reef, arising from both present and past land uses, has received increased attention in both the literature and in the media. From studying corals on both the Pandora and Havannah reefs it has been established that the transfer of sediments to the reef has increased substantially as river catchments have been modified to allow grazing, agriculture and related activities since European settlement (McCulloch, Fallon, Wyndham, Hendy, Lough and Barnes, 2003). There is great interest in the effects of eutrophication and sedimentation on the macro-algae-coral community, and its

[1] This is a revised version of a paper presented at the First Queensland Statistics Conference, The University of Southern Queensland, Toowoomba, Australia, October 1-3 2003. We would like to thank participants for their comments.

impact on the reef and associated industries such as fisheries and tourism. The largest external source of nutrients and sediments is from terrestrial runoff (Williams, 2001: 4 and PC, 2003: 25). Studies have shown that exposure to high levels of nutrients results in changes to coral growth and calcification, disrupts the reproduction process and changes the settlement success of planulae. Additionally, an increase in sediments leads to changes in the coral community structure and creates a less favourable habitat for hard corals, zooxanthellate soft corals and calcareous coralline algae, calcareous coralline algae being a critical component in reef building (Williams, 2001: 5).

Nutrients and sediments are deposited out of the mouths of rivers, the amount of nutrients and sediments is dependent on the water flow, which can vary depending on rainfall patterns. The Great Barrier Reef has a monsoonal climate with a winter dry season and a summer wet season, 70% of rainfall occurring in summer. River flow is highly variable and the estimated input of fresh water ranges from 20 to 80 cubic kilometres (Lough et.al., 2002: 333). It has been established that the flood events after the dry season have an increased suspended sediment load (McCulloch et. al., 2003) and cyclones during the summer season cause major increases in nutrient and sediment inputs through both the generation of cyclone related floodwaters and the re-suspension of nutrient rich sediments (Russ and McCook, 1999: 249).

Therefore, to evaluate the different policy options, and determine the potential impact of agriculture on the Great Barrier Reef and associated industries, it is necessary to develop a model that takes into account the changing levels of runoff and hence nutrient and sediment accumulation. Other studies have also attempted to develop river-flow models of the Burdekin and other North Queensland rivers. For example a model based on calibration of a three-dimensional plume has been developed by the CRC Reef Research for the Burdekin, Herbert, Tully and Johnstone rivers (King, et al., 2002). That model differs from ours in a number of respects. Firstly, the data set employed by King et al. involved daily data from 1951-1998. Secondly, the model parameters were determined by calibration rather than by statistical estimation and finally the model was built in order to predict spatial spread of river plumes. The type of model developed by King et al. does not lend itself easily to integration with bio-economic models hence we have opted for an alternative approach that is more amenable to integration within a bio-economic modelling framework.

It is well established that stream flow is persistent; an indication of the level of persistence can be gained from Hurst's work. Hurst found long-term dependency in fluctuations at the mouth of the Nile River, the process and theoretical foundations for determining this was later developed by Mandelbrot. This paper calculates the Hurst Exponent for the Burdekin River in North Queensland. Subsequently, the Hurst exponent will be used to simulate river flow, which will be later incorporated into a bio-economic model (Millen, 2003)[2]. This model will facilitate an assessment of the impact of coastal agriculture on fisheries industries in North Queensland under different government policies.

The paper is organised into five sections, the remainder of this section will describe the origins of the data and the method of collection, the second section will discuss the Methodology behind the Hurst exponent, the third section will discuss the process used to estimate the Hurst exponent, the fourth section presents the results of the estimation and finally the paper concludes by discussing the application of the Hurst exponent and future work.

The Calculation of the Hurst exponent makes use of reconstructed runoff data for the Burdekin River. The data set spans from 1644 to 1980 and was obtained from a study of the occurrence and intensity of luminescent lines in corals by Isdale, P.J., B.J. Stewart, K.S. Tickle, and J.M. Lough (2001). These luminescent lines provide record of the periods in which the corals were exposed to river plumes and other mainland runoff; the intensity of the lines is dependent on the percentage distance of a reef across the shelf and to the average water depth between the reef and the mainland (Lough et.al., 2002: 333).

The study found that the variation in luminescence intensity was significantly correlated with instrumental measurements of annual river flow; hence the luminescence lines provide a record of the mainland influence on reefs along the length of the Great Barrier Reef. Using the intensity of the luminescence measured in two coral cores from Havannah and Pandora Reefs, which lay in the Burdekin River plume, it was possible to reconstruct the annual runoff of the Burdekin River.

[2] University of Queensland Honours thesis 2003 in preparation

Methodology

In 1951 Hurst found long-term correlations in the fluctuation of outflows from the Nile River. He discovered that hydrological time series show longer periods of droughts and floods than were to be expected if the processes had both finite memory and variance (Van de Giesen and Mata, 2002). His work was based on the methods used for reservoir design. The design of a reservoir seeks to determine the optimum capacity that will allow ideal performance over a range of years. The capacity that allows the reservoir to produces a uniform outflow and never empty or overrun is determined by adjusting a rivers cumulative discharge for the year by the sample average for the set of years in question, n. The difference between the maximum and minimum of these adjusted value is denoted by $R(d)$, and is the optimum capacity for the reservoir.

Using $R(\Delta t)$, otherwise known as the range, as a tool to investigate actual behaviour of river discharge records Hurst found that the cumulative fluctuations in outflow satisfied the following power law (Hastings and Kissell, 1998: 83):

$$range\{y(s) | t \leq s \leq t+n\} = const \times n^{H}$$

Where the range is defined as the maximum deviation of cumulative actual behaviour from average behaviour within a sample. Mandelbrot and Wallis later confirmed it was valid to model geophysical records using fractional noises (Mandelbrot and Wallis, 1969: 321).

The Hurst exponent, H, is a self-similarity parameter that measures the long-range dependence in a time series, and provides a measure of long-term nonlinearity. The expected values of H lie between 0 and 1. For H = 0.5 the cumulative behaviour is a random walk and the process produces uncorrelated white noise. However in the cases where the Hurst exponent is either greater than or less than 0.5 there are underlying non-linear dynamics in the system. H < 0.5 represents anti-persistent behaviour and H > 0.5 is fractional Brownian motion with increasing persistence strength as H approaches 1 (Lange, 2003).

Fractional Brownian motion may be defined as follows (Falconer, 1990). Given a probability space (Ω, Σ, P) where Ω is a sample space and Σ a σ-algebra and P an associated probability measure, then a random process may be defined as a mapping $X : \Omega \times [0, \infty) \rightarrow \Re$, so that a sample function $t \mapsto X(\omega,t)$, where $\omega \in \Omega$ results from a random process. Fractional Brownian motion with index *H*, where $H \in (0,1)$ may then be defined as random process $X : \Omega \times [0,\infty) \to \Re$ such that

i. $X(t)$ is continuous and $X(0) = 0$ with probability 1

ii. For all $t \geq 0$, $\Delta t > 0$ the increment $X(t+\Delta t) - X(t) \sim N(0, \Delta t^{2H})$

So that $P(X(t+\Delta t) - X(t) \leq x) = (2\pi)^{-1/2} \Delta t^{-H} \int_{-\infty}^{\infty} \exp(-u^2 / 2\Delta t^{2H}) du$.

Fractional Brownian motion is a long-memory process and is a generalisation of Brownian motion. It has been widely studied and its properties are well understood. On studying the dependence of $R(d)/S(d)$ upon n Hurst found that $R/S \propto n^H$, Mandelbrot subsequently found that this cumulative process would be Gaussian with a vanishing delta expectation and delta variance equal to $|\Delta t|^{2H}$ (Mandelbrot, 1983: 250).

Based on Hurst's hydrological analysis R/S analysis was proposed by Mandelbrot and Wallis (1969) as a method to estimate the Hurst exponent. The Rescaled range statistic makes no prior assumptions about the shape of the probability distribution that is being studied, as it is non parametric; therefore it is ideal to be used when studying nonlinear stochastic and deterministic systems (Peters, 1994: 53). Hence, in the subsequent section R/S analysis will be used to estimate the Hurst exponent for the Burdekin River.

Hurst Exponent Estimation

The process for estimating the Hurst exponent can be found in both Weron (2002) and Peters (1994). Firstly the time series must be divided into d contiguous sub-series of length n, where $d \times n = N$, the total length of the time series. For each of theses sub-series m, where $m = 1,...,d$:

(1) Determine the mean, E_m, of each sub-series.

(2) Determine the standard deviation, S_m, of each sub-series.

(3) Normalise the data ($Z_{i,m}$) by subtracting the mean from each data point:

$$X_{i,m} = Z_{i,m} - E_m \text{, for } i = 1,..,n.$$

(4) Using the normalised data create a cumulative time series by consecutively summing the data points:

$$Y_{i,m} = \sum_{j=1}^{i} X_{j,m} \text{, for } i = 1,..,n.$$

(5) Using the new cumulative series find the range by subtracting the minimum value from the maximum value:

$$R_m = \max\{Y_{1,m},...,Y_{n,m}\} - \min\{Y_{1,m},...,Y_{n,m}\}$$

(6) Rescale the range, R_m / S_m by dividing the range by the standard deviation.

(7) Calculate the mean of the rescaled range for all sub-series of length n:

$$(R/S)_n = \frac{1}{d}\sum_{m=1}^{d} R_m / S_m$$

(8) The length of n must be increased to the next higher value, where $d \times n = N$ and d is an integer value. Steps 1 to 7 are then repeated, these steps should be repeated until $n = N/2$.

(9) Finally, the value of H is obtained using an ordinary least squares regression with $\log(n)$ as the independent variable and $\log(R/S)_n$ as the dependant variable, the slope of the resulting equation is the estimate of the Hurst exponent. The regression is run over values of n greater than 10, as small values of n produce unstable estimates when sample sizes are small (Peters, 1994: 63).

Results

Using the method outlined above the Hurst exponent can be estimated by performing the linear regression of:

$$\log(R/S)_n = \log c + H \log n$$

The regression produces the following estimates, which are also represented graphically in figure 1:

$$y = 0.7527x - 0.315 .$$

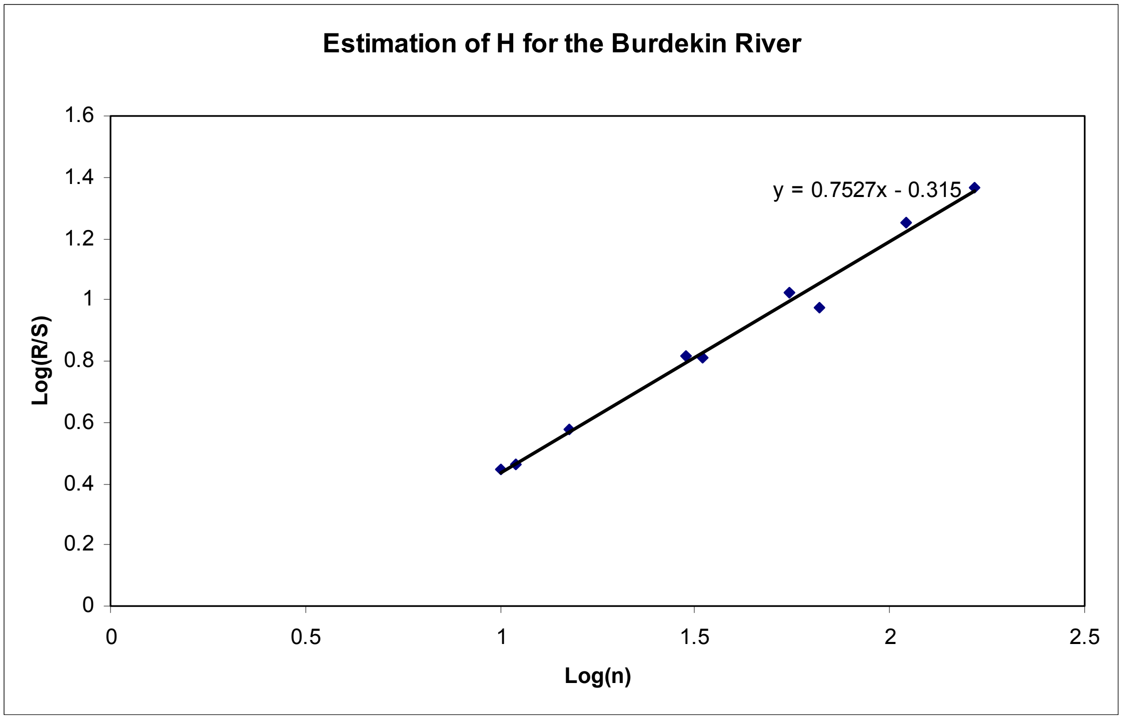

Figure 1

The estimated value of the Hurst exponent for the Burdekin River is 0.7527, for full regression output refer to Appendix A. The value of H is greater than 0.5 indicating that the annual discharge from the Burdekin River is persistent, that is, there are extended periods in which the system deviates from the long term mean. These results correspond with the previous research of others. Values of the Hurst exponent that lie between 0.5 and 1 are essentially Black noise processes, these processes are seen in nature, and are in long-run cyclical records such as river levels (Peters, 1994: 183).

This persistence that exists within the system is possibly related to the Southern Oscillation, which is linked to both the weather and climate in Australia. The Southern Oscillation is measured by an index of pressure difference between Tahiti and Darwin, and provides a measure of the strength of the Walker circulation. The Walker circulation is a major circulation cell that moves air zonally between the Eastern and Western sides of the South Pacific (Sturman and Tapper, 1996: 369).

The Southern Oscillation index is correlated with Australia's average surface pressure, rainfall, lake levels, tropical cyclones and the Australian monsoon. The Southern Oscillation is significant for Australian rainfall; its role is enacted via sea surface temperatures in tropical Australasia, although the nature of the relationship is not clear. Research has shown that there exists a statistical relationship between the Southern Oscillation Index and rainfall and it is correlated with the rainfall data in a number of ways. Specifically, there are simultaneous and lagged correlations using annual and seasonal data, and the phase of the Southern Oscillation Index is related to monthly rainfall. It has been shown that during negative phases of the Southern Oscillation Index, there is a tendency for higher surface pressures over Australia that is associated with reduced rainfall across the continent. Conversely, positive phases are associated with lower surface pressures and increased rainfall (Sturman and Tapper, 1996: 375).

Specifically there are implications for Queensland associated with the Southern Oscillation Index. Annual rainfall over Eastern Australia is significantly correlated with Southern Oscillation indices, that is, a rapid increase in the Southern Oscillation Index is related to high rainfall totals in Eastern Australia. Additionally, a relationship exists between extreme phases of the Southern Oscillation and tropical cyclone activity in Australia. It is during the high phases of the Southern Oscillation Index that cyclones develop closer to the coast of North-east Australia, impacting on Queensland. Finally, the Australian monsoon that dominates the weather and climate, and controls the rainfall regime in tropical Australia from December to March each year is closely linked to the Southern Oscillation Index (Sturman and Tapper, 1996: 375-379).

The Hurst exponent can now be used to simulate water outflow, which can then be integrated into the bio-economic model. H is generalised so it can range from 0 to 1 for fractional Brownian motion. Simulation of fractional Brownian sample paths is difficult (Falconer, 1990: 247). There are a number of possible procedures. Random walk approximation, directly generating the noise and approximating the path using a randomized Weierstrass function. Both the direct method and the randomized Weierstrass method are compared here.

Using the direct method one generates a fractional Brownian sample path in continuous-time $Z = X(t+\Delta t) - X(t) \sim N(0, \Delta t^{2H})$ by the inverse transform technique (Banks et al., 1996: 322). This is then re-scaled using the formula $Y = \mu + \frac{\sigma}{\Delta t^{H}} Z$ where μ is the mean level of the river-flow and σ is the standard deviation of the river-flow. Alternatively, the randomized Weierstrass method may be employed. This approximates fractional Brownian sample paths using the function:

$Z(t) = \sum_{k=1}^{\infty} C_k \lambda^{-Hk} \sin\left(\lambda^k t + A_k\right)$, where $\lambda > 1$ and $C_k \sim N(0,1)$ and the A_k are independent random variables uniformly distributed on $[0,2\pi)$ (Falconer, 1990: 247). Again this must be re-scaled using $Y = \mu + \frac{\sigma}{\Delta t^H} Z$ to obtain forecasts of the appropriate scale. As noted below in the section on model validation, improved results may be obtained by replacing the mean by the mode.

A scatter plot of the reconstructed runoff data is shown below (figure 2). The average annual runoff is 99.1867 mm and the standard deviation is 96.05442. Using these figures the simulated fractional Brownian sample path was rescaled and plotted for the period 1930-1980 to produce the second graph shown below (figure 3).

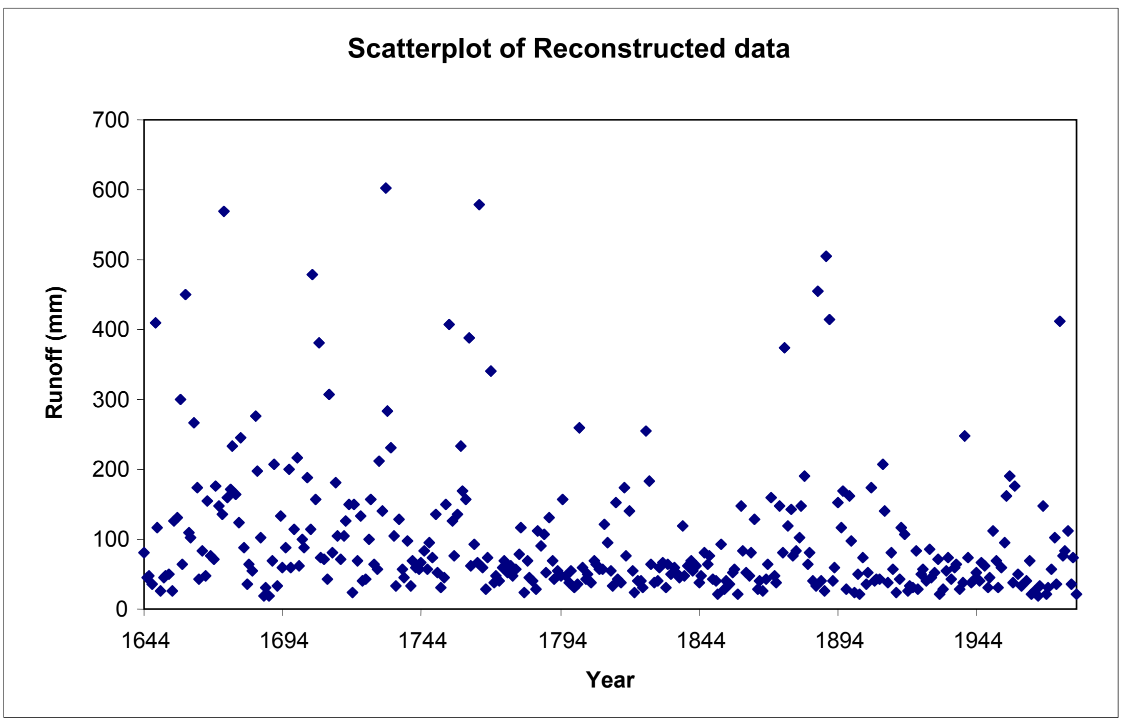

Figure 2

From figure 3 it can be seen that although there is not much difference in the results of the two simulation methods, the randomized Weierstrass approach is slightly better at capturing the extreme tails of the distribution.

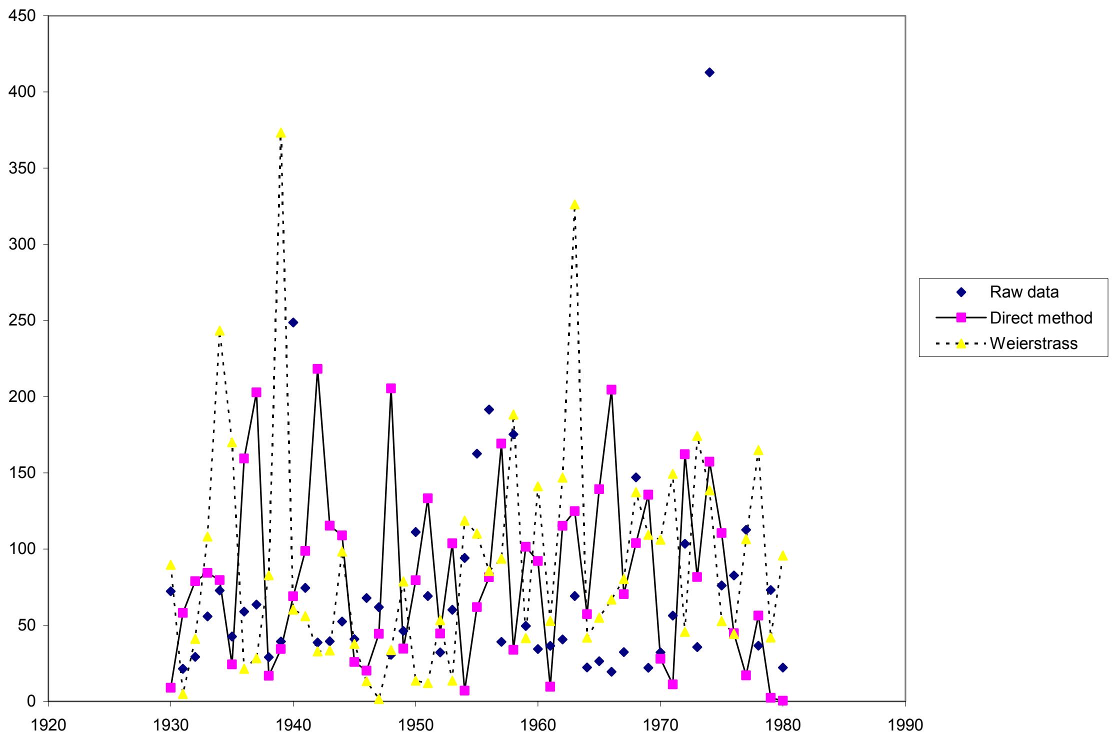

Figure 3: Model Predictions against reconstructed river-flow data 1930-1980

Model Validation

In order to develop an effective river-flow simulation model it is necessary to validate the model predictions against actual data. To do this we compared the empirical distribution against the predicted distribution from the simulated fractional Brownian motions based on both the direct method and the randomized Weierstrass method. Because river-flow cannot take on negative values it is necessary to constrain the stochastic process to be positive valued. This can be done in one of two ways. Either a reflecting or absorbing barrier can be employed. It was found that a reflecting barrier gives the best results. Furthermore more robust results are obtained by replacing the mean by the modal value of the empirical data set for the purpose of rescaling the simulated fractional Brownian sample paths. The model predictions are characterized by convergence in distribution not point-wise convergence[3], so we used Kernel density estimators to compare predicted and actual distributions

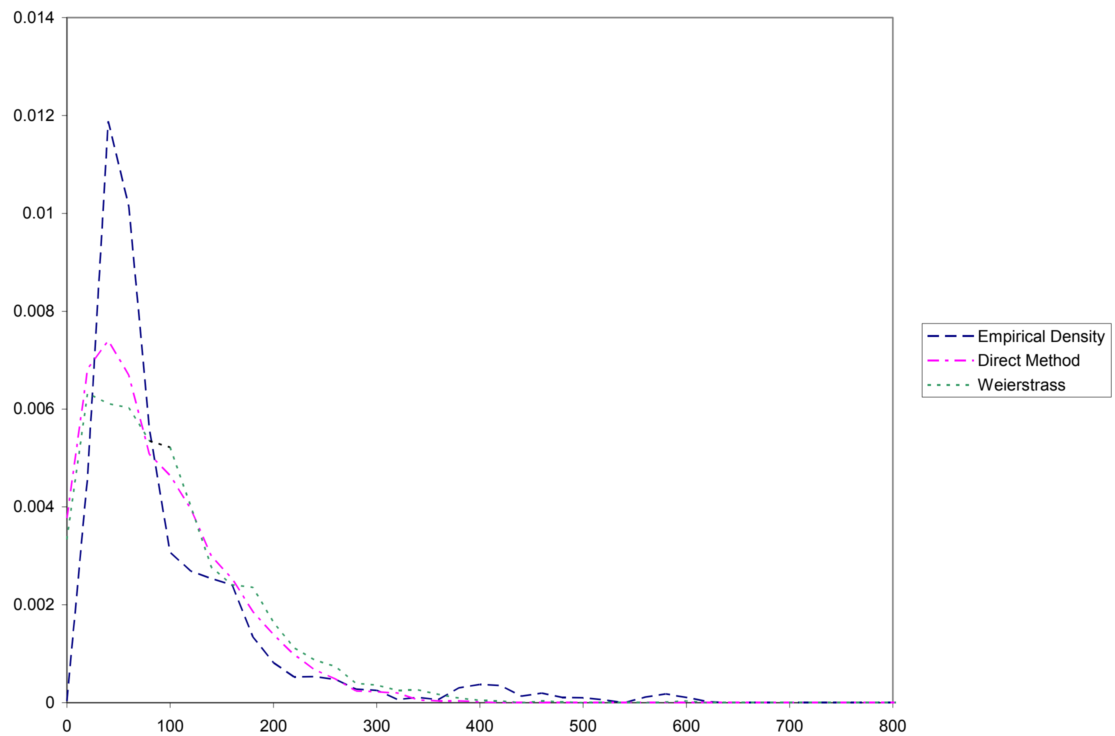

Figure 4: Comparison of the distribution of model predictions against the distribution of reconstructed river-flow data.

We employed the optimal Epanechnikov Kernel for the purposes of the comparison with a fixed window width of h=20. It is likely that a variable window width estimator would give an improved fit. The predicted distribution shows that the empirical data set has high kurtosis 8.286621932 compared with the predicted data set based on direct simulation 0.333538998 and based on the randomized Weierstrass method 1.245310378. Both approaches to simulation are underpredicting the range and amount of variation in the data but the predicted mean river-flows are reasonably close to the actual. We conclude that this approach shows

[3] We would like to thank Olena Kravchuk for this observation.

some promise as a means of simulating river-flow but that we still need to validate the model in a more thorough manner.

Conclusion

This paper has shown that using the Hurst-Mandelbrot rescaled range statistic the estimated value for the Hurst exponent is 0.7527 for the Burdekin River. It has been shown that it is possible to simulate values for future runoff. Such a simulation is of great value when analysing the impact of agriculture on the Great Barrier Reef. As nutrients are delivered in plumes at the mouth of the rivers it is necessary to have a model of river flow to account for extreme fluctuations in the delivery of nutrients.

The model of runoff will be integrated into the development of a bio-economic model that will assist in the analysis of the policy options available to the government. The development of such a framework will aid in determining the course of action that will be most effective in reducing the affects of coastal agriculture on the Great Barrier Reef and associated industries such as Queensland fisheries.

Appendix A

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.994599485
R Square	0.989228135
Adjusted R Square	0.987689297
Standard Error	0.036389198
Observations	9

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.851233194	0.851233194	642.841032	3.79121E-08
Residual	7	0.009269216	0.001324174		
Total	8	0.86050241			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	- 0.314963229	0.047848799	-6.582468885	0.000309312	-0.428107578	-0.20181888	-0.428107578	-0.20181888
X Variable 1	0.752706577	0.02968752	25.35430993	3.79121E-08	0.682506798	0.822906356	0.682506798	0.822906356

Table 1

Empirical Data Set		*Direct simulation*		*Weierstrass*	
Mean	99.1867656	Mean	83.32344275	Mean	95.38958
Standard Error	5.23241755	Standard Error	1.984919813	Standard Error	2.41778
Median	64.09	Median	71.91235772	Median	77.55243
Mode	44.24	Mode	#N/A	Mode	#N/A
Standard Deviation	96.0544177	Standard Deviation	62.80005231	Standard Deviation	76.49514
Sample Variance	9226.45117	Sample Variance	3943.84657	Sample Variance	5851.506
Kurtosis	8.28662193	Kurtosis	0.333538998	Kurtosis	1.24531
Skewness	2.67349496	Skewness	0.894484153	Skewness	1.167269
Range	583.86	Range	300.0249171	Range	435.0204
Minimum	18.47	Minimum	0.12364149	Minimum	0.473278
Maximum	602.33	Maximum	300.1485586	Maximum	435.4937
Sum	33425.94	Sum	83406.76619	Sum	95484.97

References

Banks, J., Carson, J.S. II and Nelson, B.L. 1990. *Discrete-Event System Simulation,* 2nd Edition, Prentice-Hall, Upper Saddle River, N.J.

Falconer, K. 1990. *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley & Sons Chichester.

Hastings, H. M. and Kissells R. (1998) Is the Nile Outflow Fractal? Hurst's Analysis Revisited. *Natural Resource Modeling* 11[2]: 83-93.

Hurst, H.E. 1951. Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers.* 116: 770-808

Isdale, P.J., Stewart, B.J., Tickle, K.S. and Lough, J.M. 1998. Palaeohydrological variation in a tropical river catchment: a reconstruction using fluorescent bands in corals of the Great Barrier Reef, Australia. *The Holocene* 8: 1-8.

Isdale, P.J., B.J. Stewart, K.S. Tickle, and J.M. Lough, 2001,Burdekin River Annual Runoff Reconstruction, IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series #2001-025. NOAA/NGDC Paleoclimatology Program, Boulder CO, USA.

King, B., McAllister, F., Done, T., 2002. Modelling the Impact of the Burdekin, Herbert, Tully and Johnstone River Plumes on the Central Great Barrier Reef. CRC Reef Research Centre Technical Report No. 44, CRC Reef Research Centre, Townsville.

Lange, H. 2003. Time Series Analysis of Ecosystem Variables with Complexity Measures. http://www.bitoek.uni-bayreuth.de/~Holger.Lange/iccs2/lange.htm. Visited, March 11th, 2003.

Lough, J.M., Barnes D.J. and McAllister F.A. (2002) Luminescent Lines in Corals from the Great Barrier Reef Provide Spatial and Temporal Records of Reefs Affected by Land Runoff. *Coral Reefs* 21: 333-343.

Mandelbrot, B.B. 1983. *The Fractal Geometry of Nature*. New York: W.H. Freeman and Company.

Mandelbrot, B.B. and Wallis, J.R. 1969. Some Long-Run Properties of Geophysical Records. *Water Resources Research*. 5[2]: 321-340.

McCulloch, M., Fallon, S., Wyndham, T., Hendy, E., Lough, J. and Barnes, D. 2003. Coral Record of Increased Sediment Flux to the Inner Great Barrier Reef since European Settlement. *Nature.* 421: 727-730.

Peters, E.E. 1994. *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*. Brisbane: John Wiley and Sons Inc.

Productivity Commission (PC). 2003. *Industries, Land Use and Water Quality in the Great Barrier Reef Catchment*. Research Report, Canberra.

Russ, G.R. and McCook, L.J. 1999. Potential Effects of a Cyclone on Benthic Algal Production and Yield to Grazers on Coral Reefs across the Central Great Barrier Reef. *Journal of Experimental Marine Biology and Ecology.* 235: 237-254.

Sturman, A.P. and Tapper, N.J. 1996. *The Weather and Climate of Australia and New Zealand*. Melbourne: Oxford University Press.

Van de Giesen, N and Mata, L.J. 2002. Comparison of the Hurst exponents of Historical and GCM rainfall time series. *Hydrology Days*.

Weron, R. 2002. Estimating long-range dependence: finite sample properties and confidence intervals. *Physica A*. 312: 285-299.

Williams, D. McB. 2001. *The Impact of Terrestrial Run-Off on the Great Barrier Reef World Heritage Area*, CRC Reef Research Center, Australian Institute of Marine Science.