Using fractals for modelling volatile data and estimating the value of electricity and weather contracts

info@futuremath.co.nz

Fat-tailed distributions and clustered volatility

Fat-tailed probability distributions exhibit a higher number of extreme events than are expected using more well-known probability functions such as normal and lognormal distributions.

Many kinds of data exhibit fat-tailed probability distributions, particularly data generated by complex systems, to the degree that fat-tailed distributions are a near universal property of complex systems. Yet despite this, the use of fat-tailed distributions to model phenomena such as weather, electricity prices and financial markets is a relatively recent innovation.

Certainly these methods do not appear to have been adopted by the standard modelling packages such as @risk.

Another important point is that most of these same systems exhibit the property of clustered volatility. This means that the greatest changes in the value of the data tend to be clustered together in time.

The only method that can properly characterise these two behaviours is fractal analysis.

Fractals are mathematical objects that have found wide use in recent times. The key characteristic of fractals is that they display “scaling” structure, that is to say, that zooming in to examine small sections of fractal data simply resembles the larger-scale data, no matter how closely one zooms in. Thus fractals are often said to be self-similar objects, since the look the same on all scales.

Though for real objects this scaling property only holds for a finite range of scales, it has been shown that the statistical properties of systems such as weather and financial markets can only be described with fractals.

Using these fractal methods, it is possible to come up with theoretically robust schemes for pricing various kinds of derivative products based on weather or financial market behaviour. Much research has been conducted to explore these methods in recent times [1], though a standard approach based on fractals is yet to supersede other more traditional methods.

For example, analysis of financial markets by many researchers has revealed probability distribution functions that display fat tails and that follow power laws. Mandelbrot was the first to show this, using daily prices for cotton in 1963 [2].

Despite this, most quantitative finance methods for pricing derivatives are based on Brownian motion, owing to the work of Batchelier [3] in 1900. This approach drastically underestimates large amplitude fluctuations that occur during periods of severe volatility.

The widespread use of these models is attributable to the Black-Scholes equation [4], which is applied to model the fair value for derivatives trading. For some kinds of application, the Black-Scholes equation can be made to work well enough that there is little value in using a fractal model instead. Therefore there is significant institutional inertial to replacing such approaches with new methods based on fractals.

However, the Black-Scholes approach has significant weaknesses, and is likely to be of limited use for very volatile data such as that underpinning weather derivatives. Its weakness lies in two key assumptions. The first is that the movements from one day to the next are independent of each other. The second is that they follow a normal distribution with a well-defined standard deviation. This assumes large spikes in prices are extremely rare events that almost never happen.

However, in practice one observes spikes on a regular basis, often as frequently as once per month [5]. Furthermore, while the size of the daily fluctuations can remain roughly constant for long periods, the variability may suddenly jump for a briefer period. Such periods are often referred to as examples of “clustered volatility”. Thus the size of price movements are not independent, but display a degree of correlation in time. This property can be used to model the behaviour of the volatility function into the future, something that current methods do not allow.

In the following sections, it will be shown how fractal-based models can use the properties of clustered volatility to make more accurate estimates for the prices of derivatives, and how these methods may be particularly suitable to volatile markets and weather derivatives.

Many methods based on fractals have been developed for analysing volatile data, particularly in the area of geophysics. For example, experiments show that turbulence data displays fractal characteristics, but the solutions to the classical Navier-Stokes differential equations of fluid dynamics do not reproduce these. Fractal models, on the other hand, can beautifully reproduce the experimental data, even though they only "phenomenologically" model the physics.

It just happens that turbulence data has statistical properties very similar to sharemarket data. Although there is still a great deal of debate about the reasons for this, it appears that complex systems with a large number of parts interacting in a non-linear way universally display these sorts of properties. Data from these systems is best described in terms of "fractal dimensions", which are parameters that describe the statistical character of the data. There does not appear to be any other method that captures these properties, and no way of simulating such data except via fractal methods.

Spikes, Fat-tailed Distributions and Clustered Volatility

In terms of characterising data associated with various kinds of derivatives, there are principally two types of data. The first resembles turbulence data (such as currency markets), the second rain data (the NZ electricity market is a good example).

In this section we will examine the first kind, using the cross-rate between the Kiwi and United States dollars as an example.

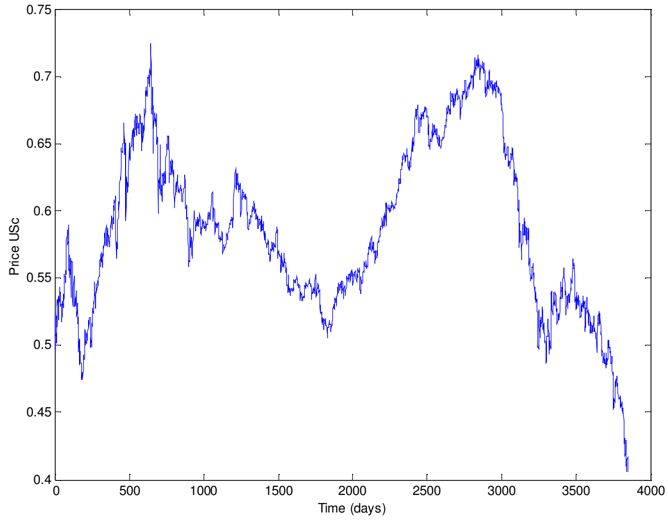

Figure 1: Cross rate between the Kiwi and US dollars.

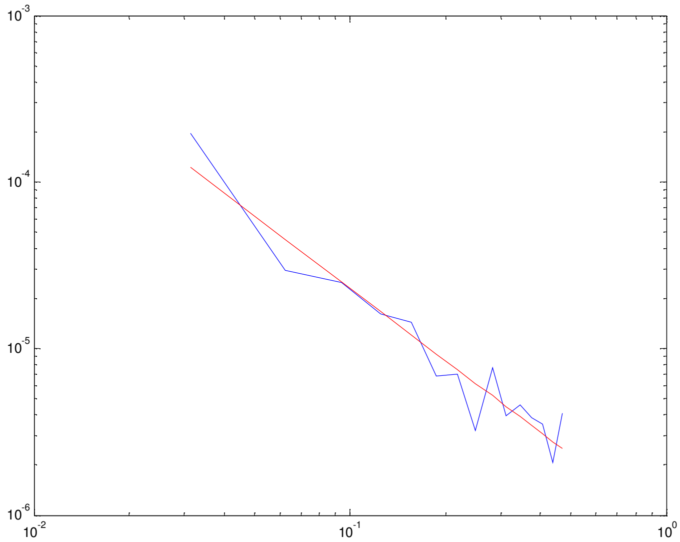

Figure 2: The spectrum of the Kiwi dollar data shown in Figure 1.

Figure 1 shows changes in the value of the Kiwi dollar for 1000 data points during the late 1980s and early 90s. As discussed in the introduction, if this data is fractal then it displays scaling properties. One way to characterise this is to examine the Fourier Spectrum of the data. If the data is scaling, then we would expect the spectrum to show no preferred scales, so there will be no large peaks in the spectral plot representing preferred periods.

Figure 2 shows the spectrum. It can be seen that the spectrum roughly obeys a power law (i.e. a straight line on a log-log plot), therefore does not exhibit a preferred scale. The power-law exponent describing the slope on this graph is –1.44, and is one fractal dimension of the data. This provides us with our first necessary parameter with which the data can be characterised.

However, there are further parameters that need to be obtained that are associated with the property of clustered volatility. For the first type of data (i.e. turbulence-like), the property of clustered volatility is best analysed by examining the behaviour of the changes in price, rather than the price itself. Figure 3 shows the changes in the price of the Kiwi dollar, compared with a normally distributed set of price changes. The two sets of data are qualitatively quite different. One might speculate that the proper probability distribution for the changes in the Kiwi dollar more closely resembles a lognormal distribution (after taking absolute values). However, the actual distribution has a greater proportion of extreme values than can be accounted for by a distribution of lognormal form. Such distributions are often said to exhibit “fat tails”. The distribution is most accurately described by a Levy-Stable type distribution, of which the Cauchy distribution is an example. Fractals can easily be made to obey probability distributions of this type.

To further appreciate the advantages of fractal methods it is important to understand that the standard deviation of the data is not a good measure of risk. This is because the standard deviation changes dramatically from one subset of data to another. This is symptomatic of clustered volatility.

The way to characterise this property is to analyse how the standard deviation of the data depends on the resolution at which it is examined. For example, Figure 4 shows the change-of-price data for the New Zealand dollar squared. From the figure, it is apparent that the largest changes are clustered together in time.

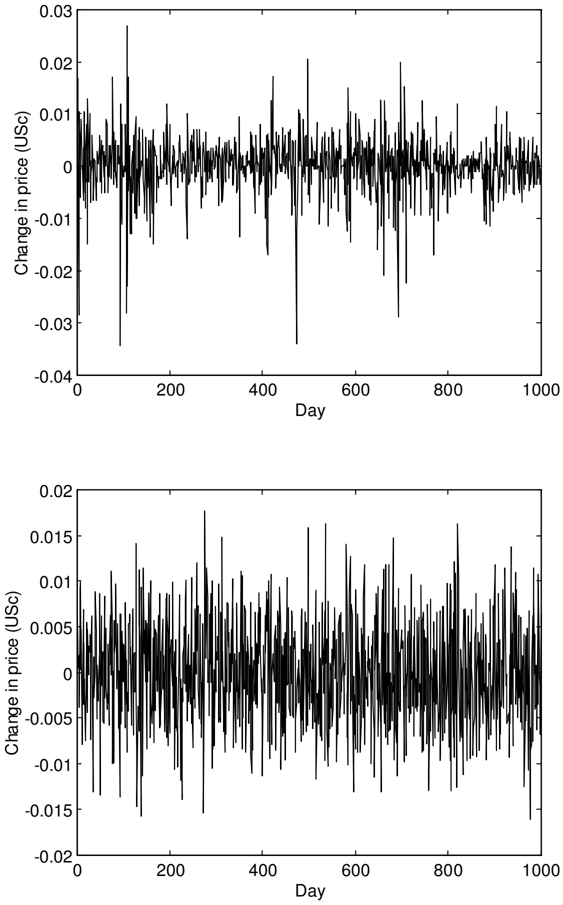

Figure 3: Comparison of actual changes in the value of the Kiwi dollar (top) with a normal distribution of changes (bottom).

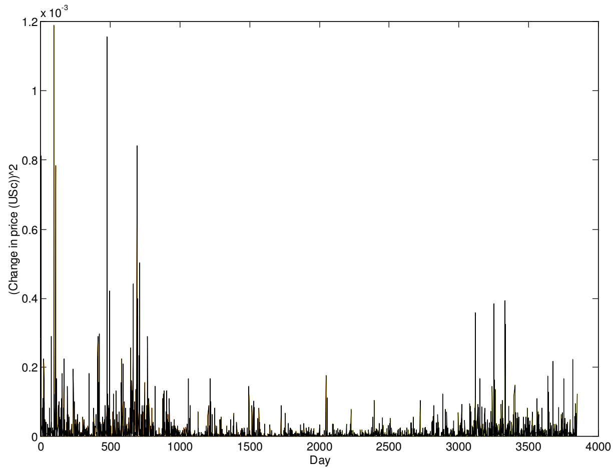

Figure 4: Squaring the changes in the value of the Kiwi dollar reveals that the largest changes are clustered together in time. This property is called clustered volatility.

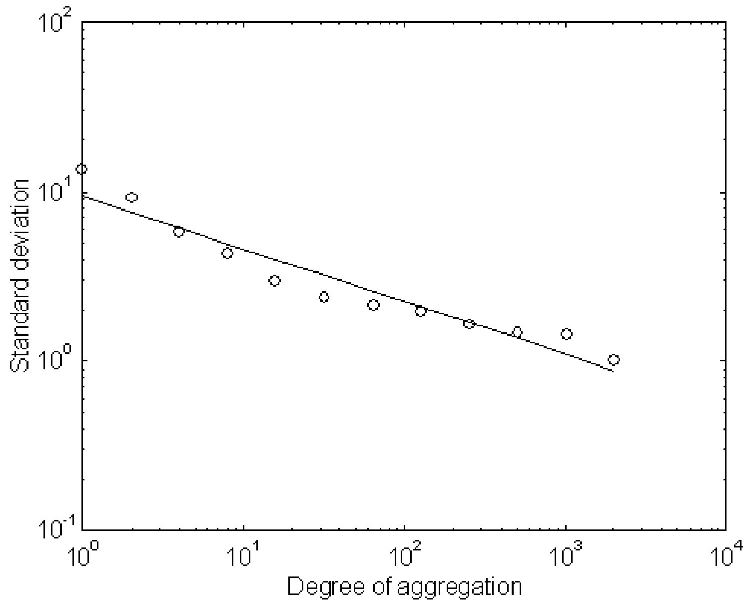

Figure 5: Examining the value of the standard deviation of the data plotted in Figure 4 shows that it depends on the resolution at which the data is examined.

If we measure the standard deviation of this data, we find that it depends on the degree of aggregation. This dependence is plotted in Figure 5. When plotted like this on a log-log scale, the standard deviation is found to obey a power-law (i.e. straight line on a log-log plot) dependence on the data resolution. The slope of the straight line is effectively a fractal dimension associated with the standard deviation (i.e. second-order statistical moment). Here the slope is –0.27. Broadly speaking, the steeper the slope, the greater the degree of clustered volatility exhibited. For comparison, sharemarket data usually exhibits a steeper slope than that shown for the Kiwi dollar, and rain data a steeper slope still.

Similar plots can be made for the other statistical moments, thus allowing a "spectrum" of fractal dimensions to be obtained, which characterise the statistical nature of the data completely. This type of characterisation is called "multifractal" analysis [6]. Unlike the standard deviation, this fractal dimension characterisation is independent of the resolution of the data.

We can use this information to model how the volatility is likely to behave in the near future given a certain level of volatility in the present.

Fractal Simulations

Before looking at the other type of data (rain-like), we will quickly discuss how fractal data can be simulated. Once a characterisation of a data set using fractal dimensions has been obtained, a properly parameterised fractal model can generate synthetic data virtually indistinguishable from the real data. Thus this simulated data exhibits the desired clustered volatility.

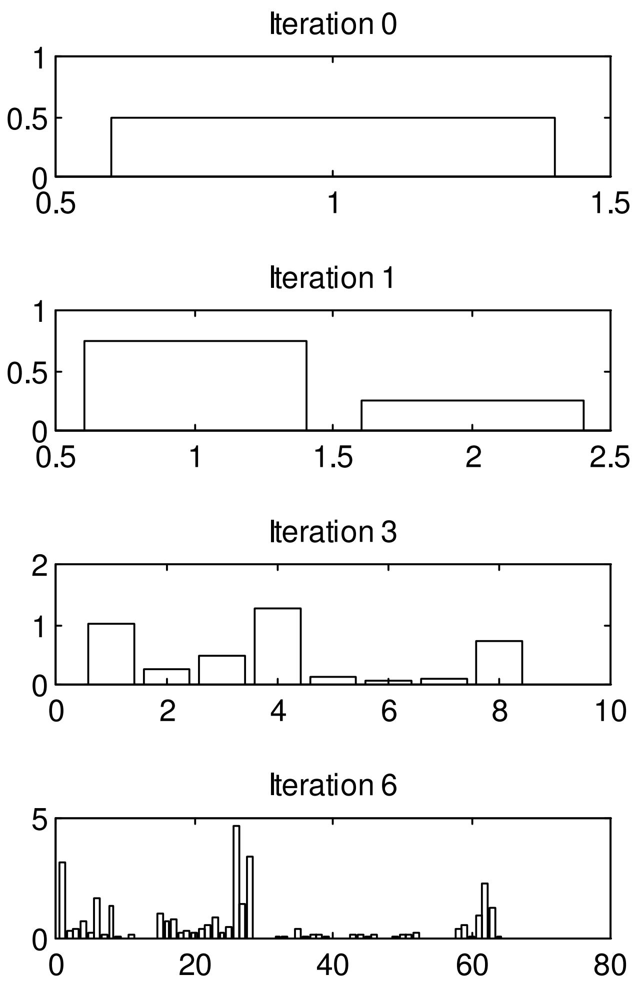

Figure 6: Schematic demonstration of a fractal cascade generator.

One such model is a fractal cascade. The idea of a fractal cascade is that it takes an interval over which a function is uniformly distributed, splits that interval into some fraction (say half), and reassigns new values to that section according to some random generating function. This process is then reiterated on each new interval, and so on, until sufficient new iterations have occurred to generate a convincing facsimile. The process is illustrated schematically in Figure 6.

How well the simulated data display the same multifractal properties as the actual data depends on the form of the generator used. Mandelbrot suggests a simple generator for financial data in [5]. However, designing and choosing the appropriate generator is quite a tricky business, and these simple approaches do not generally produce the correct statistical behaviour of all types of data.

Example of pricing a derivative contract: I) Electricity price.

Having obtained a method for generating synthetic data which displays the same statistical properties as the real data, it is possible to use this model to calculate the price that should be paid for hedging instruments.

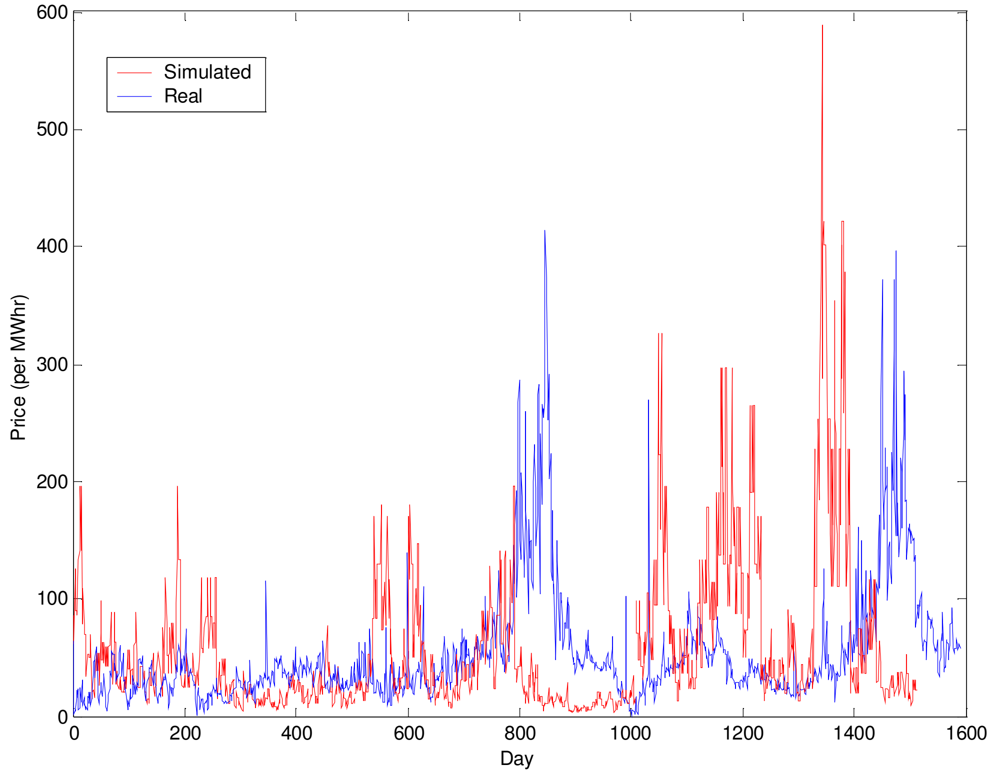

Figure 7: Simulated daily electricity price data constructed using a multifractal, compared with real data. The purpose of this graph is to demonstrate the similarity in the appearance of real and simulated data.

In the case of the electricity market data shown in Figure 7, it is relatively straightforward to find the fractal dimensions of the data (the data plotted is the average daily price). This type of data is markedly different from the currency market data examined in the earlier section. It is much more spiky, making it more important to characterise the volatility than the mean trend (for shorter-term estimates, at least). Such data resembles rainfall data, and so is “rain-like”. The principal difference between this data and the “wind-like” data is that we can model the price directly rater than the volatility.

Having obtained the fractal parameters, synthetic data can be produced with the same statistical properties as the real data (also shown in Figure 7). The synthetic data is scaled to match the real data by a constant so that it has the same long-term mean as the real data. All other statistical information, such as the volatility, are already described by the fractal dimensions.

Since we can generate as much of this data as we like, it is easy to generate statistics on the likelihood of certain future outcomes. Furthermore, because the degree of volatility for such data is so prominent, the behaviour of the volatility function in the future has a strong degree of correlation with its behaviour in the immediate past.

Thus we can use the synthetic data to answer questions about the probability of certain events happening conditional on current conditions.

An example of this is shown in Figure 8. The graph is the cumulative probability distribution for the mean electricity price for the month April 6 to May 6.

Two cases are shown.

The first (Blue) represents the probability distribution that one would generally expect (i.e. ignoring the peculiarities facing the market at that time). The second (Red) is the probability distribution given that the mean price for the previous month was $160 per MWhr.

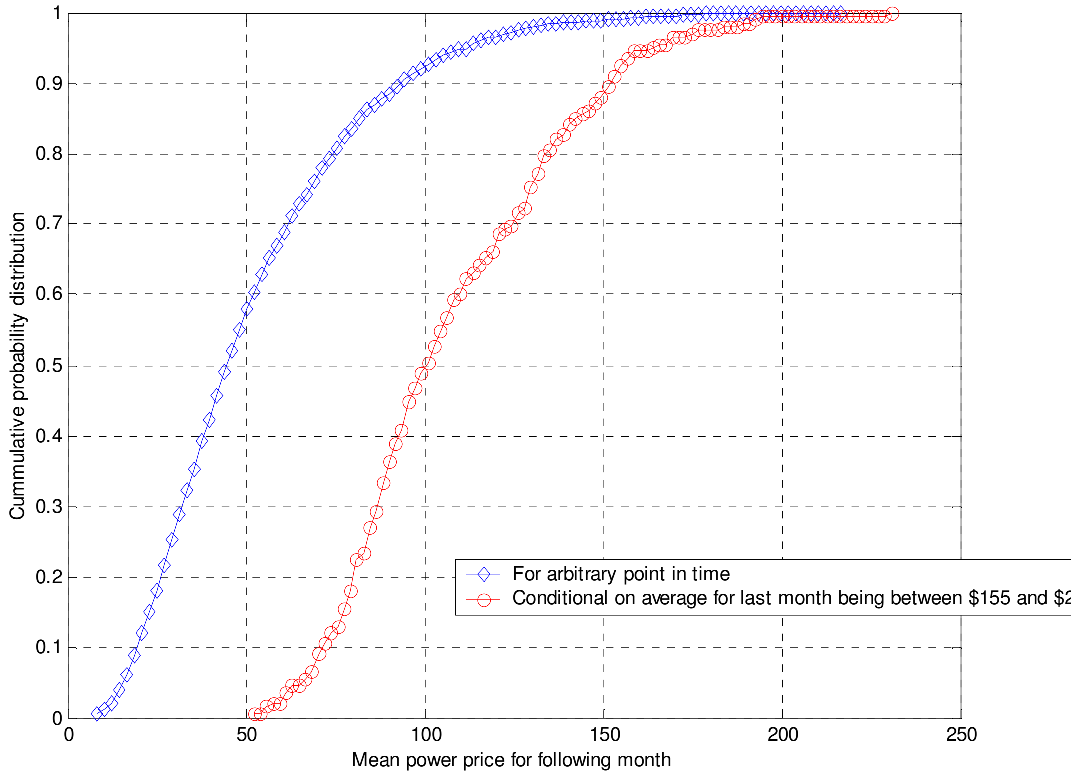

Figure 8: The cumulative probability distribution for the mean price of electricity for a given month, both generally and for the specific case where the mean price for the previous month was between $155 and $200 per MWhr.

Reading from this graph allows us to generate the following estimates:

	$0-$50	$50-$100	$100-$150	$150-$200
Case at start of April	58%	35%	6%	1%
Current case	<1%	50%	38%	12%

The table shows the probability that the mean electricity price will fall between certain ranges in the next month. This in turn allows one to estimate the fair value for electricity contracts.

For example, suppose that a contract was struck at the start of the year to supply 1MWh of electricity over the period from April 6 to May 6 for a price of $50 per MWh. To estimate the likely savings a company could make by owning this contract as opposed to buying on the spot market, we calculate the expectation values for the electricity price from the probability distribution graphs above.

I.e. the expectation value is:

$$\text{Expectedcost for next month} = \int (x - \$50) f(x)\, dx$$

where $f(x)$ is the probability of the electricity price being x.

According to our estimates, the holder of this contract could in the coming month save \$2.77 per MWhr in the general case, and \$58.24 per MWhr in the current situation. Note that expected savings means what would be expected on average, and for a particular realisation could be much more or less.

The same calculations can be made for various periods, since it is easy to generate new $f(x)$ functions for these in the same way. The table below demonstrates this.

	Term of contract		
Start date	1 month	6 months	12 months
May 6	10.7c	6.6c	5.1c
June 6	8.6c	6.0c	4.9c

Furthermore, because the model provides a probability distribution of outcomes, it is possible to look at the likelihood of various final results. An example is shown in the table below, using public domain data on electricity usage by various companies. If we assume that each company bought all of its electricity needs through a 1-year contract for a fixed price of 5c/kWhr, the table below shows the probability of a profit or loss relative to the spot market.

Company	Best 25% of cases	Bottom half of best cases	Top half of worst cases	Worst 25% of cases
Pan Pac	Better than \$6.6m	\$6.6m to -\$0.55m	-\$0.55m to -\$6.1m	Worse than -\$6.1m
Comalco	Better than \$60m	\$60m to -\$5.0m	-\$5.0 to -\$55m	Worse than -\$55m
AIA	Better than \$0.28m	\$0.28m to -\$0.023m	-\$0.023 to -\$2.6m	Worse than -\$2.6m
NZ Refining	Better than \$2.8m	\$2.8m to -\$0.24m	-\$0.24 to -\$0.26m	Worse than -\$0.26m

Hedging example II: Rainfall derivatives.

The same procedure as described above can be used to calculate the value of rainfall derivatives. Here we use the example of monthly rainfall for Melbourne. Such data is shown in Figure 9.

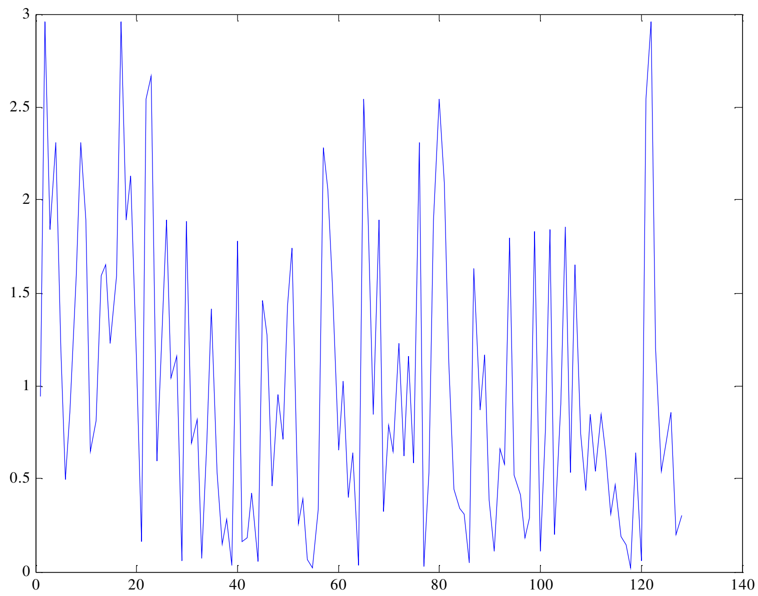

Figure 9: Sample of normalised monthly rainfall data.

Although it does not look much like the electricity data from the previous section, this data displays a high degree of temporal correlation that can also be used to make conditional probability distributions. An example is shown in Figure 11 for the level of rain expected for the next 6 months from July 2003. Clearly the level of rain from the preceding six-month period has had a significant impact on the distribution for the next six months.

Note that rainfall has a seasonal dependence, so in generating such six-monthly estimates it is necessary to introduce an adjustment to the estimated ranges, to reflect the fact that the historical average for the first six months of the year may be different from the average for the second six months of the year.

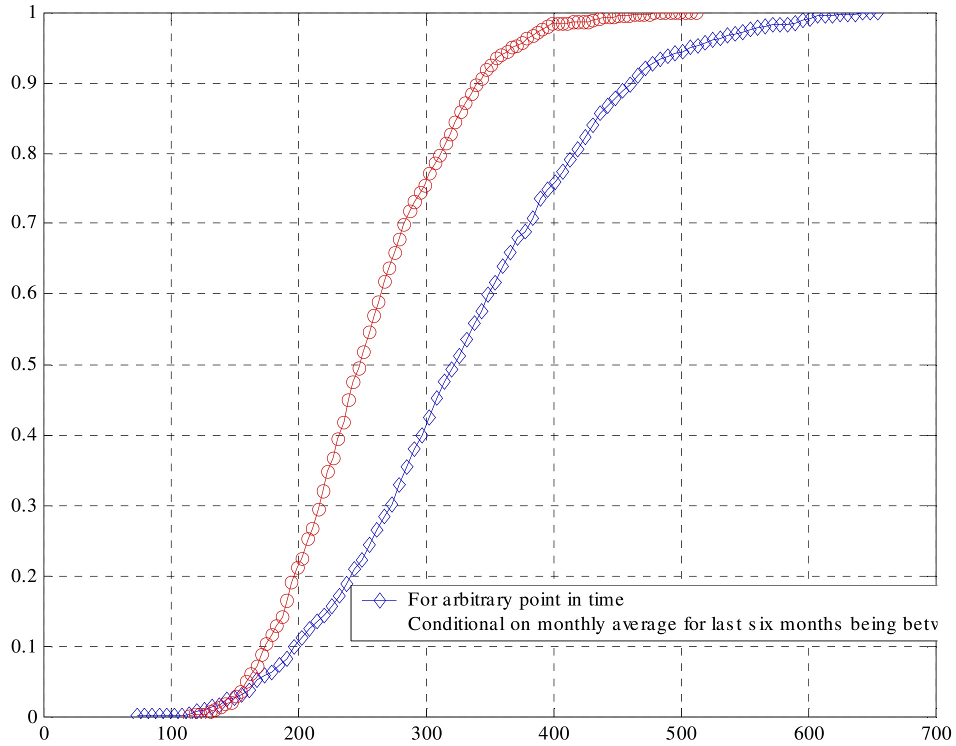

Figure 10: Cumulative probability distributions for the total rainfall for six months from the end of June 2003.

Using these distributions, the value of a put option (which pays out money if the actual rainfall falls below some strike value) can be calculated using the equation:

$$w = e^{-rt} \int_{\infty}^{c} (x-c) f(x)\, dx$$

where w is the theoretical value of the option, x is the rain total, t is the time to maturity, r is the risk-free interest rate, c is the strike value, and f is probability distribution function used to generate the cumulative distribution in Figure 11.

Methods such as these provide a robust way of estimating the fair value for derivatives based on the electricity market, and more generally, other very volatile derivatives, such as weather-based contracts. Their existence makes it practical for such a market to operate without the overdue risk to the participants that extreme volatility brings.

References

[1] Farmer, J. D. 1999. “Physicists Attempt to Scale the Ivory Towers of Finance”, *Computing in Science and Engineering*, Nov-Dec, 26-39.

[2] Mandelbrot, B. B. 1963. “The Variation of Certain Speculative Prices”, *Journal of Business,* **36**(4), 394-419.

[3] Batchelier, L. 1900. “Theorie de la speculation”, *Annales Scientifiques de l’Ecole Normale Superieure*, 21-86.

[4] Black, F. and M. Scholes. 1973. “The Pricing of Options and Corporate Liabilities”, *Journal of Political Economy*, **81**, 637-654.

[5] Mandelbrot, B. B. 1999. “A Multifractal Walk Down Wall Street”, *Scientific American*, Feb, 50-53.

[6] Davis A., Marshak A., Wiscombe W., and Cahalan R. 1994. “Multifractal Characterisations of Nonstationary and Intermittency in Geophysical Fields: Observed, Retrieved, or Simulated”, *J. Geophys. Res.* **99** (D4), 8055-8072.