

A horizontal banner with a dark red background. On the left, the text "School of Economics and Finance" is written in a white, serif font. Below it, in a smaller white font, is "Te Kura Ohaoha Pūtea". The right side of the banner features a collage of images, including a building, a map of New Zealand, and a close-up of a person's face.

School of Economics and Finance  
Te Kura Ohaoha Pūtea

SEF Working paper: 08/2011  
June 2011

---

**Modelling New Zealand electricity  
prices from a risk management  
perspective**

Caroline Moy and Leigh Roberts

The Working Paper series is published by the School of Economics and Finance to provide staff and research students the opportunity to expose their research to a wider audience. The opinions and views expressed in these papers are not necessarily reflective of views held by the school. Comments and feedback from readers would be welcomed by the author(s).

Further copies may be obtained from:

The Administrator  
School of Economics and Finance  
Victoria University of Wellington  
P O Box 600  
Wellington 6140  
New Zealand

Phone: +64 4 463 5353

Email: [alice.fong@vuw.ac.nz](mailto:alice.fong@vuw.ac.nz)

**Working Paper 01/2011**  
**ISSN 2230-259X (Print)**  
**ISSN 2230-2603 (Online)**

# Modelling New Zealand electricity prices from a risk management perspective

Caroline Moy\*      Leigh Roberts†

2 June 2011

## Abstract

A direct approach is taken to modelling New Zealand electricity prices, in which extreme value theory is used to augment a basic time series model. Despite its simplicity, the resulting model is suitable for answering fundamental questions of interest to risk managers, who might not find it worthwhile to apply a more sophisticated and complex approach to statistical modelling.

---

\*Tower Insurance, Auckland, New Zealand

†School of Economics and Finance, Victoria University, Wellington.  
email: leigh.roberts@vuw.ac.nz

# 1 Introduction

While modelling market behaviour is rarely straightforward, electricity markets seem to present the modeller with greater challenges than other markets. One reason is that it is virtually impossible to store electricity, so that conventional economic theories of market behaviour do not apply; another is that transmission of electricity to the national grid entails substantial leakages, which is a particular problem in a large sparsely populated country such as New Zealand. A third point is that the strategic importance of electricity means that the industry is subject to manifold constraints.

Analysis of the electricity markets when they were operated as state sponsored monopolies was hardly easy either; but in one sense one could then approach the problem as a large scale operations research or decision sciences problem. Now that the electricity industry in developed countries is largely deregulated, the situation for the modeller has become vastly more intricate. Even at the most basic level of understanding the outworkings of a market, it is often not easy to comprehend institutional arrangements of an industry in which participants operate at several levels, with various stages of vertical integration, some of which are encouraged and others prohibited. In New Zealand, for instance, generators and retailers are frequently jointly owned, but are forced to be independent of lines companies (eg see Macrae and Wolak (2009)).

Faced with complicated market structure and behaviour, a modeller's reaction naturally tends toward greater complexity, whether at the statistical level; or at the level of the competitive games played in submitting bids in the 24 hour forward auctions; or at the level of the regulator in attempting to control aspects of the industry for social ends. It is frequently the case however that risk managers are not so interested in elaborate mathematical modelling of a scenario. Their time horizon is often shorter rather than longer, and an easily produced simple forecast may be quite adequate for their purposes, especially in explaining to their superiors what they and their traders in the market place are up to.

Bearing this in mind, our paper deliberately aims at simplification of the mathematical approach to modelling electricity prices. We adopt a basic time series approach, and collateral information is ignored, although covariates could easily be introduced into our model. Decision makers would normally wish to consider several perspectives of a situation before reaching a decision; and the methodology suggested here is intended to complement other methodologies for quantifying risk, not to supplant them.

Restricting our attention to a single representative node of the national electricity grid, and to a time window of 32 months of daily average prices from 1 January 2006, we model the price time series as a seasonal ARIMA (SARIMA) model. The residuals from this fitted model are heavy tailed, and we fit a Pareto distribution to these residuals. This enables us to obtain approximate answers to questions of interest to risk managers, viz. concerning the level of Value at Risk or Expected Shortfall; or finding the profit or loss to be expected from a forward contract, *inter alia*. We also investigate the cyclical behaviour of this time series by fitting wavelets, giving the risk manager clues as to which frequencies he should focus on in his modelling, and which contracts he might use to hedge.

The first section describes the data, and comments on the initial approaches to modelling, looking at residuals obtained by differencing the original data; inspecting autocorrelation functions; and looking at the data through wavelet decomposition.

The next section describes the preferred model for the data, and we obtain a series of normalised residuals from that model. This series is heavy tailed, and a Pareto distribution is fitted to the positive tail. The Pareto model allows us to consider some basic quantities of interest to the risk manager; and a short conclusion terminates the paper.

Finally, all computations were effected in R, particularly using the 'evir', 'POT' and 'wavelet' libraries. Abbreviations utilised include ACF and PACF for (partial) autocorrelation function; and cdf for (cumulative) distribution function, pdf for probability density function. The acronyms MSE stand for mean square error, and MLE for maximum likelihood estimation; ARIMA

and GARCH have their usual meaning, and DC is direct current.

## 2 Modelling the data

### 2.1 A first look at the data

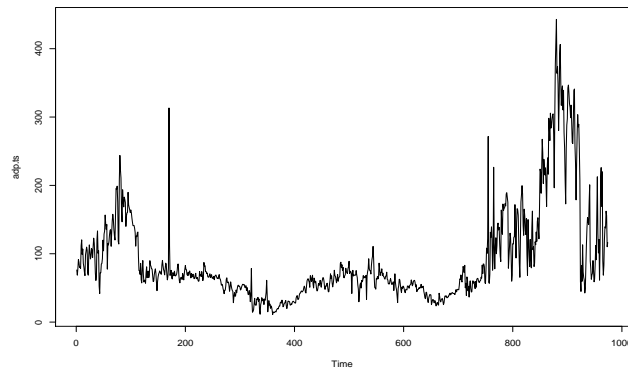


Figure 1: Average daily electricity prices from 1 January 2006 - 31 August 2008, at Haywards grid point

The electricity price data shown in Figure 1 is that for Haywards Hill, an important node in the national grid in New Zealand because it lies at the north end of the DC cable connecting the North and South Islands. The daily data shown is an average of the original half hourly price data. There are slight hiccoughs in the data shown because of two additional (half hourly) readings taken when daylight saving ends, each year about April; and two fewer readings when daylight saving is introduced, generally in October each year. Accounting for the slight inconsistencies in the data at these times did not materially impact the conclusions of this paper.

Average daily prices for two further grid points over the same time window are compared with those for Haywards in Figure 2. Benmore is the grid

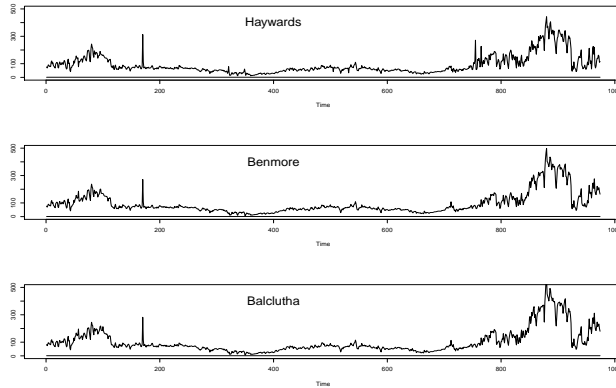


Figure 2: Average daily electricity prices from 1 January 2006 - 31 August 2008, for Haywards, Benmore and Balclutha

point at the southern end of the inter-island DC cable, near Christchurch; and Balclutha is the grid point nearest the aluminium smelter at the bottom of the South Island, the biggest electricity consumer in New Zealand. At the admittedly coarse level of this diagram the prices at the three nodes display similar features. There are some 300 grid points in New Zealand, with those displayed being three of the more important, and we choose to model the Haywards prices in this paper.

## 2.2 Fitting an ARMA type model

It is clear from Figure 1 that some differencing will be required before one can hope to fit ARMA or GARCH type models, since these classes correspond to stationary time series.

A first look at the data would entail fitting models by differencing, first by applying filters of the form  $1 - L$ , where  $L$  is the lag or back shift operator; then by applying filters of the form  $1 - L^s$ , where  $s$  is the periodicity of a seasonal effect; and finally by combining these filters, viz.  $(1 - L)(1 - L^s)$ .

The latter two models incorporate a moving average through the data:

$$1 + L + L^2 + L^3 + \dots + L^{s-1} = \frac{1 - L^s}{1 - L}$$

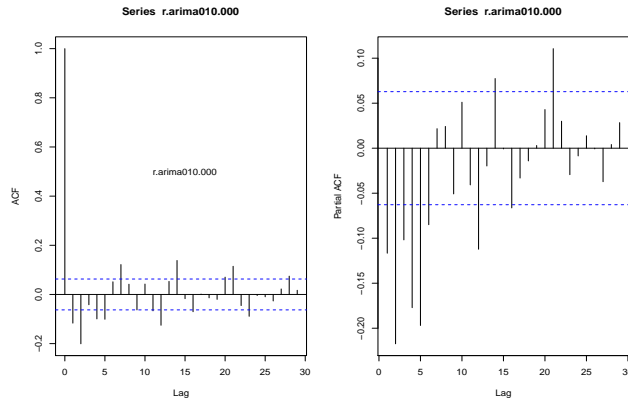


Figure 3: ACF and PACF for residuals of the simple differencing model

Figure 3 shows the autocorrelation and partial autocorrelation functions (ACF, PACF) of the residuals when simple differencing is applied to the data. The spikes at every 7th reading indicate the presence of a cycle of period 7 days: this is expected on prior grounds, and most modelling of electricity pricing and/or demand contains a weekly cycle. So we take  $s = 7$  in the above.

The three models produced by differencing in the above ways produce residuals plotted in Figure 4. None of these three models is sufficient in itself for satisfactory modelling: we shall clearly need to fit additional parameters. According to Moy (2010, p. 26), the SARIMA/EVT model outperforms the GARCH/EVT model on this dataset, in that the former provides a lower MSE and provides a better fit, judging visually. We therefore restrict ourselves to the ARMA type models in this paper, which assume the following form:

$$\Phi(L^s)\phi(L)(1-L^s)^D(1-L)^d X_t = \Theta(L^s)\theta(L)\epsilon_t$$

which may be denoted as  $\text{ARIMA}(p, d, q)(P, D, Q)_s$ , in a customary notation.



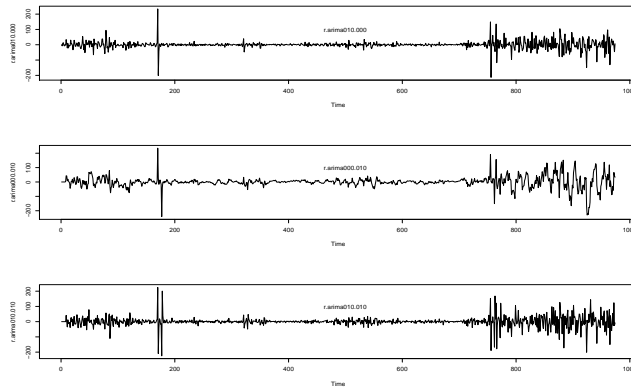


Figure 4: Residuals from fitting the three basic differenced models

Our preferred model for this dataset is  $ARIMA(2, 0, 1)(0, 1, 1)_7$ , the fitting of which produces the coefficients

|      | ar1  | ar2  | ma1  | sma1 |
|------|------|------|------|------|
|      | 1.52 | -.53 | -.81 | -.94 |
| s.e. | .05  | .05  | .04  | .01  |

The fit of the coefficients is sharp, meaning that the standard errors are relatively small; and the resultant ACF and PACF appear satisfactory, to judge from Figure 5. The model is reasonably parsimonious, and is preferred on the basis of the Akaike Information Criterion over other simpler models considered.

Our fitted model is shown in red in Fig 6, superimposed on the original data. Also shown are the conventional 95% forecasts for a month out of sample (September 2008), the blue lines being two standard errors away from the central green estimate. The influence of the weekly cycle is clear in the forecasts.

Before proceeding to fit a Pareto distribution to the tails of the residuals

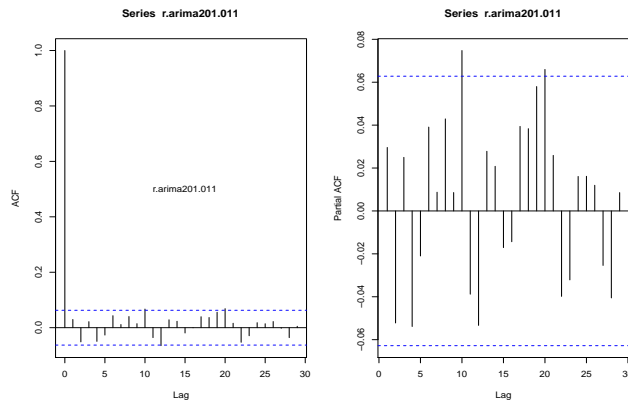


Figure 5: ACF and PACF for residuals of the preferred SARIMA model

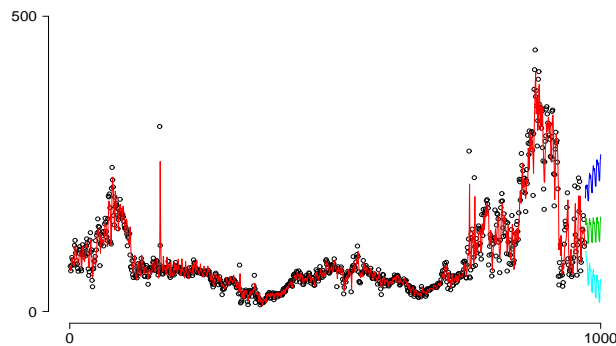


Figure 6: Fitted (red line) and actual daily electricity prices from 1 Jan 2006 - 31 August 2008, Haywards grid point. The green lines are the point forecasts for September 2008; and the dark and light blue lines the conventional confidence interval of width four standard errors.

from our fitted model, we examine a wavelet decomposition of the data.

### 2.3 Wavelet analysis

A wavelet decomposition is essentially a fitting of cycles of various frequencies on a roving basis throughout the data. One advantage of such an approach is to allow the strength of a cycle of a particular frequency to vary over time, in contrast to the situation with a conventional Fourier analysis of the data. On the other hand, the association with frequency, via 'time scales', is looser than in Fourier analysis, basically because one is not fitting sine curves.

We have used the conventional discrete wavelet transform (DWT), with seven levels of frequencies, or rather seven time scales. The data used is from March 2006 to August 2008, giving a multiple of 7 lots of  $128 = 2^7$  days, viz. from the 79th to the 974th day inclusive:  $974 - 79 + 1 = 7 \times 128$ . One could have used up to 9 levels,  $2^9 = 512$  still being less than 974, the length of our price time series; but 7 levels seemed reasonable for our purposes. A standard reference for wavelets is Percival and Walden (2000).

Principal features of the wavelet fit are shown in Figures 7 and 8: the first graph shows the wavelet smooths, the second the wavelet details. The vertical scale of the graphs has been adjusted for consistency as we compare a subgraph with those above and below it, although there is no such consistency between the two graphs.

At first glance, the lower frequencies for the smooths (towards the top of Figure 7) do little more than confirm our impressions from the data itself in Figure 1 on p. 4: there is greater activity when it is expected, viz. during the first third of the data; to a lesser extent in the middle third; and to a far greater extent in the final third. The higher frequencies have also a discernible pattern of the same form. The persistence of this pattern of activity across frequencies is a somewhat indirect indication of the usefulness of white noise as a modelling tool for this data, although in itself this observation provides little validation of the stationarity assumed in our fitted ARMA models.

The wavelet spectrum indicates the overall strength of the signal at the given time scale. Scaling these to sum to unity, the relative strengths of the signals at the various time scales over more or less the whole period is indicated in Figure 9. There is substantial strength at the 2-4 month time scale, corresponding to the 6th level (cycles of length  $2^6 = 64$  days to  $2^7 = 128$  days). This is presumably a 3 month or seasonal cycle, and was not picked up from the ARIMA modelling previously carried out, although our ACFs were not extended so far. Nor would this cycle easily be modelled, for 3 months is an indeterminate number of days. The strength of the signal at the 6th level is also indicated in the detail in Figure 8.

The consumption of electricity can be expected to have a seasonal influence; but that is not to say that the price is necessarily subject to the same influence. The risk manager may have contracts which depend on both quantity supplied to the grid as well as price; and the message is perhaps to signify the potential usefulness of 90 day contracts for hedging, whether options, futures and forward contracts, or contracts for difference (CFDs), or whatever other type of derivative, are involved.

With the benefit of hindsight, after having seen the spectrum of the wavelets over nearly the whole period in Figure 9, it is possible to discern the presence of rough 3 month cycles in the data: perhaps most easily in the levels D4 and D5 in Figure 8.

The weekly cycle is hardly salient, to judge from the spectrum of the wavelet fitting: it belongs within time scale 2 (cycles of length  $2^2 = 4$  days to  $2^3 = 8$  days). The time scale 2 contributed a little more to the overall signal than did time scales 1 and 3, to judge from Figure 9; but there is not much in it. This impression is corroborated by a further glance at the ACF in Figure 3 on p. 6, in which there are significant correlations at periodicities other than weekly.

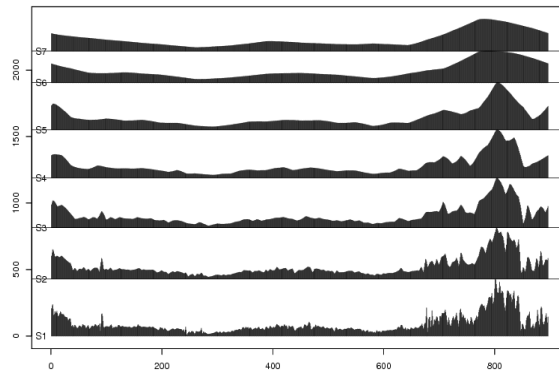


Figure 7: wavelet smooths

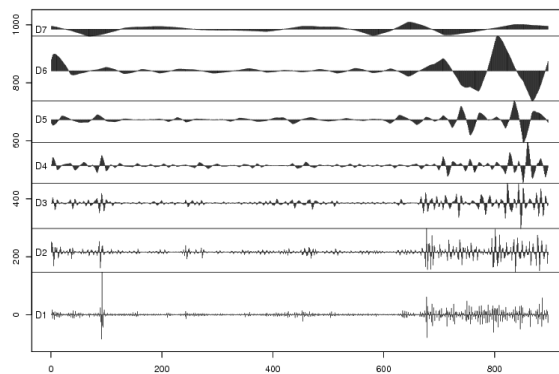


Figure 8: wavelet details

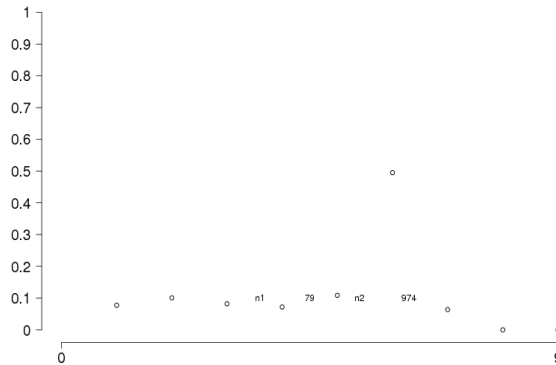


Figure 9: wavelet spectrum for March 2006 - August 2008

### 3 Fitting the Pareto distribution

One could fit a Pareto distribution to the unadjusted residuals from the fitted model, but it is convenient to standardise those residuals first, largely for ease of comparison with standard normal random variates, the statistical cornerstone of our SARIMA model.

In standardising those residuals, we use the mean and variance of the earlier unstandardised residuals, ignoring the residuals occurring later in the time series. We do this partly because the innovations in ARMA or GARCH type models arise from the immediate past, and are naturally not impacted by the future; to some extent we are mimicking the operation through time of fitting the assumed underlying model from the available data, except that we are using the parameter values obtained from fitting the model over the whole period.

Estimating the mean and variance from the preceding residuals in order to standardise residuals produces the standardised residuals in Figure 10. It is clear that one could well be interested in fitting Pareto distributions to

the positive and negative tails separately, thereby modelling innovations in the underlying SARIMA model as a mixture of positive and negative Pareto distributions and a normal distribution.

We assume that the risk manager is worried about prices being too high rather than too low, and fit a Pareto distribution to the positive tail only. Innovations in the underlying SARIMA model are therefore considered as a mixture of a positive Pareto distribution and a normal distribution.

One reason for eschewing the fitting of a Pareto distribution to the negative tail is that, while we have a reasonable amount of data for time series purposes, it is not much data for fitting extreme value distributions. A more pragmatic reason is that in the forecast values shown in Figure 6, our approach is basically to increase the upper limit of the forecasting interval in Figure 6, as indicated in Figure 12; extending the interval downwards may well take us to zero and beyond, which is not useful in predicting prices. We assume that extending the forecasting interval upwards is of more interest to the risk manager; in the contrary case one could try the same approach with logarithms, and model the negative tail with a Pareto distribution.

Our approach is to take the greatest 100 standardised residuals, use the lowest of these as a threshold point, and take this as the basic data to which we wish to fit a Pareto distribution. The other 874 points are taken to be normally distributed. The weights in the mixture are chosen as 1/10 and 9/10 instead of the more precise 100/974 and 874/974.

Having fixed the weighting probabilities and fitted the distributions to be used in the mixture, ideally one would fit a SARIMA model to the data with innovations drawn from the mixture; or at least one would simulate the SARIMA model already fitted (and thus using the parameters obtained from the original fit) to simulate values, both within the basic period and beyond. We do not do either of these: we simply adjust the forecast intervals in Figure 6 by increasing the upper limit in accordance with the Pareto distribution fitted to the upper tail of the standardised residuals.

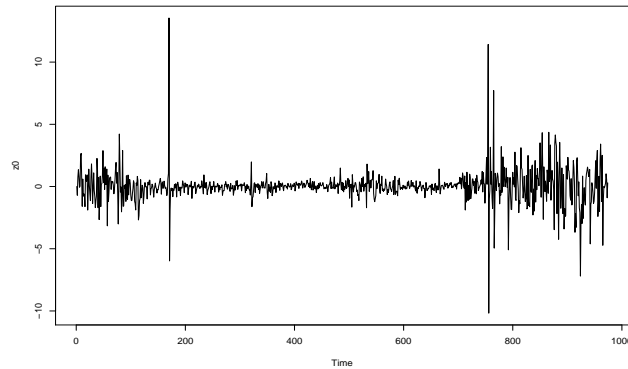


Figure 10: Standardised residuals of the preferred SARIMA model

Following McNeil, Frey and Embrechts (2005, p. 281), the mean excess plot of residual lifetimes for those top 100 points are plotted in Figure 11.

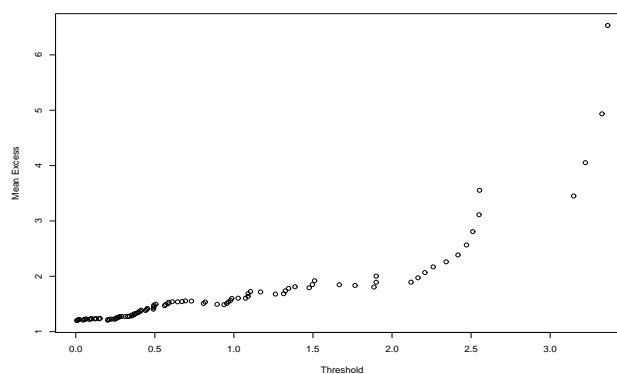


Figure 11: Mean excess plot

In theory, the parameters of the fitted gpd are to be found from the latter part of this graph, the final straight line (see McNeil et al. (2005, p. 281)). This would give a threshold of about 3. Depending on what one considers to be a 'straight' line, one could model the GPD by choosing the threshold for fitting to be 2 or 3, or anywhere in-between. We tried 2, 2.5 and 3,



and chose the lowest of these figures, partly because the shape parameters for the higher two values had standard errors almost as large as themselves, whereas the estimate at 2 was rather sharper; and partly because as a matter of practicality, the shape parameter in the first case was less than one, which is necessary if the Pareto distribution is to possess a finite mean value. The lower the fitting threshold the better for the number of points used for the fit; but also the harder to justify the fitting of an extreme value distribution.

Using 2 as the fitting threshold yielded the following MLE parameter fits:

|      | scale | shape |
|------|-------|-------|
|      | .82   | .66   |
| s.e. | .34   | .37   |

This Pareto distribution is to be applied to residuals exceeding unity in the standardised residuals, since the 100th greatest standardised residual is close to unity. The remaining 874 lower values of standardised residuals produced a mean and standard deviation of  $-.215$  and  $.879$  respectively, but we do not need these values for our purposes, since we are not intending to simulate a SARIMA model with a mixture for the innovations.

The cdf for this so-called Generalised Pareto distribution assumes the form

$$G(x) = 1 - \left(1 + \frac{\xi(x-u)}{\beta}\right)^{-1/\xi} \quad (1)$$

for

$$1 + \frac{\xi(x-u)}{\beta} > 0 \quad \text{and} \quad x > u$$

see McNeil et al. (2005, p. 276). The corresponding pdf is

$$g(x) = \frac{1}{\beta} \left(1 + \frac{\xi(x-u)}{\beta}\right)^{-1-1/\xi}$$

and the mean residual lifetime (Hogg and Klugman (1984, p. 58)) or mean excess (McNeil et al. (2005, p. 277)) is given, for  $u < v$ , by

$$e(v) = E(X - v | X \geq v) = \frac{\beta + \xi(v-u)}{1-\xi} \quad (2)$$

The label of 'Generalised Pareto' for the distribution given in (1) is a little misleading. Models arising in extreme value theory would generally be expected to produce a positive shape parameter  $\xi$ , in which case this 'Generalised' Pareto distribution reduces to the ordinary Pareto distribution (eg, Hogg and Klugman (1984, p. 222)).

One genesis of the Pareto distribution is that of a ratio of independent exponential and gamma random variates; and the usual meaning ascribed to the Generalised Pareto distribution by North American actuaries is when that ratio has a general gamma random variate in the numerator rather than the exponential: see Hogg and Klugman (1984, p. 54) and Panjer and Willmot (1992, p. 121), *inter alia*. The usage of the Generalised Pareto distribution in R follows that in McNeil et al. (2005, p. 276). The inconsistency of the two definitions is unfortunate.

Denoting the standardised residuals by  $X$ , and the components of the mixture comprising  $X$  by  $X_1$  for the normal and  $X_2$  for the Pareto; and setting  $x_0$  to be the 90th percentile of  $X_2$ , one has

$$P(X < x_0) = \frac{9}{10} P(X_1 < x_0) + \frac{1}{10} P(X_2 < x_0) = \frac{9}{10} + \frac{1}{10} \times .9 = .99$$

The 99th percentile of  $X$  is therefore the 90th percentile of the Pareto random variate  $X_2$ .

Referring to (1),

$$\left(1 + \frac{\xi(x_0 - u)}{\sigma}\right)^{-1/\xi} = \frac{1}{10}$$

rearrangement of which yields

$$1 + \frac{\xi(x_0 - u)}{\sigma} = 10^\xi$$

For the Pareto fitted to this data, the scale parameter is  $\beta = .82$  and the shape parameter  $\xi = .66$ , while the lower threshold  $u = 1$ . Inserting these values into the last equation yields

$$x_0 = \frac{\beta}{\xi} \times (10^\xi - 1) + u = 5.44.$$

Forecasts for a month out of sample (September 2008) are shown in Figure 12. The expected point values are in green, and surrounding these central values are 95% confidence intervals using normal residuals in the SARIMA modelling: the limits are in blue, and the width is 4 standard errors. The influence of the weekly cycle in the forecasts is apparent, with four obvious cycles. The upper mauve forecast is of the 99% VaR point, substantially above the upper limit of the conventional confidence interval, in dark blue.

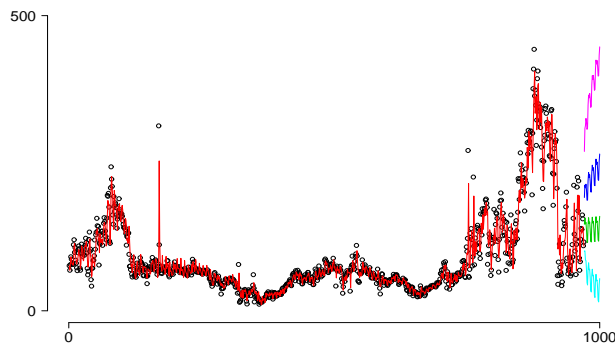


Figure 12: forecast upper quantiles

## 4 From a risk manager's perspective

The highest 'forecasts' in Figure 12, viz. the mauve coloured squiggles on the top right corner, are in fact the 99th percentiles of the out of sample price according to our model. One can go further and calculate the expected price conditional on its exceeding that 99th percentile, given that our value of the shape parameter is less than unity: from (2) on p. 15, the residual lifetime of a Pareto random variate at value 5.44 is 11.0; in other words the expected value of a Pareto random variable in our model conditioning on its exceeding 5.44 is 16.4. Transferring this to the Figure 12 means an expected shortfall of the forecast mean + a massive 16.4 times the forecast standard error, whereas the Value at Risk (VaR) is the forecast mean + 5.4

times the forecast standard error. For a 7 day ahead forecast, the Value at Risk translates into \$220, and the expected shortfall into \$615; for a 30 day horizon, these become respectively \$356 and \$936.

To be using our model in this way may be useful to give some ideas of more realistic Values at Risk, but one clearly needs to be careful when interpreting the expected shortfalls calculated in the above manner. To compound the uncertainty, the standard errors for the above calculations are assuming that there is no uncertainty in the parameters fitted to the basic model. More significantly, our fitting of a Pareto distribution was also highly uncertain in itself: quite reasonable fitting choices, for instance, resulted in Pareto distributions with infinite mean.

That said, any model for the expected shortfall has to be extremely uncertain. To be working with the upper reaches of a fat tailed distribution when that distribution is itself highly uncertain is of course going to be fraught with difficulties in interpretation of the results.

The above calculations also assume that the risk manager is only concerned about one particular contract maturing on a given day. They will in fact have hedged their positions in the various contracts, and be wishing to take a holistic view of their financial commitments.

## **5 Conclusion**

We have fitted a seasonal ARIMA (SARIMA) model to New Zealand electricity prices at a particular node in the national grid, and then modelled the resulting outliers using extreme value theory (EVT) methods. The double dipping approach of using EVT after the prior application of a previous methodology of modelling is not new (see McNeil and Frey (2000)), but seems not to have been applied to electricity markets, nor to ARMA type models.

We have ignored other possibly relevant factors, such as water levels in the dams which provide much of New Zealand's electricity supply, quantity of electricity demanded each day, and the weather. It is easily possible to insert explanatory variables into our method, but we chose not to. It is one thing to model the behaviour of electricity pricing as a function of covariates over a time period; and another matter to provide satisfactory forecasts in the future when those covariates need themselves to be forecast. Our approach eschewing collateral information is intended as a serviceable guide to the risk manager. As and when more elaborate mathematical models are fitted, the risk manager should certainly incorporate the extra information into his decision making.

In truth, the SARIMA and EVT models are not quite elementary either, although they can readily be used to produce up to date forecasts; but our stripping away any of the complications bedevilling the electricity industry is appealing, and the use of the Pareto distribution allows us to formulate approximate answers to some basic risk management questions quite neatly.

## References

- Hogg, R. V. and Klugman, S. A. (1984). *Loss Distributions*, Wiley.
- Macrae, S. D. and Wolak, F. A. (2009). How do firms exercise market power: evidence from a bid-based wholesale electricity market. from the internet, dated 13 March 2009.
- McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach, *Jnl. of Empirical Finance* **7**(3-4): 271–300.
- McNeil, A. J., Frey, R. and Embrechts, P. (2005). *Quantitative risk management*, Princeton University Press.
- Moy, C. (2010). *Statistical analysis of new zealand electricity prices: a risk manager's perspective*, Master's thesis, School of Mathematics, Statistics and Operations Research. Submitted.

Panjer, H. H. and Willmot, G. E. (1992). *Insurance Risk Models*, Society of Actuaries.

Percival, D. B. and Walden, A. T. (2000). *Wavelet methods for time series analysis*, CUP.