Are Daily and Weekly Load and Spot Price Dynamics in Australia's National Electricity Market Governed by Episodic Nonlinearity?

Phillip Wild*, Melvin J. Hinich** and John Foster***

In this article, we use half hourly spot electricity prices and load data for the National Electricity Market (NEM) of Australia for the period from December 1998 to February 2008 to test for episodic nonlinearity in the dynamics governing daily and weekly cycles in load and spot price time series data. We apply the portmanteau correlation, bicorrelation and tricorrelation tests introduced in Hinich (1996) to the time series of half hourly spot prices and load demand from 7/12/1998 to 29/02/2008 using a FORTRAN 95 program. We find the presence of significant third and fourth order (non-linear) serial dependence in the weekly load and spot price data in particular, but to a much more marginal extent, in the daily data.

* Corresponding author. School of Economics and ACCS, University of Queensland, St Lucia, QLD, 4072, Australia, Tel: +61 7 3346 9258, email: p.wild@uq.edu.au.

** Applied Research Laboratories, University of Texas at Austin, Austin, TX 78712-1087, Tel: +1 512 232 7270, email: hinich@austin.utexas.edu..

***School of Economics, University of Queensland, St Lucia, QLD, 4072, Australia, Tel: +61 7 3365 6780, email: j.foster@uq.edu.au.

1 INTRODUCTION

The Australian electricity market as a whole encompasses both supply and demand side interactions encompassing generation, transmission, distribution and retail sale activities. The predominant market in Australia is the National Electricity Market (NEM) which is structured as a gross pool arrangement. The NEM commenced operation as a de-regulated wholesale market in New South Wales, Victoria, Queensland, the Australian Capital Territory (ACT) and South Australia in December 1998. In 2005, Tasmania joined as a sixth region. Operations are essentially based on six interconnected regions that broadly follow state boundaries (NEMMCO 2005, 4).

A number of 'stylised' facts are widely accepted as applying to load demand and spot price dynamics in the market. The first relates to broader cyclical character of load demand. Specifically, observed load demand patterns in the market tend to vary from region to region depending upon such factors as population, temperature and industrial and commercial needs. Electricity demand tends to be cyclical in nature, with demand being lower in the spring and autumn than in summer and winter.

The second 'stylised' fact relates to the load curve having a weekly and daily cycle. The peak hourly load in Australia has two distinct peaks that are generated by domestic activity. Demand tends to be low in the early morning hours and begins to increase, with a first peak period occurring between 7.00 am and 9.00am, before tending to drop off and flatten out between 11.30 am to 1.30pm before starting to climb once again. The second peak occurs between 4.00pm and 7.00pm. Demand also follows a weekly cycle and tends to be higher on weekdays than during weekends.

The third ‘stylised’ fact relates to spot electricity price dynamics that are viewed as exhibiting both the properties of high volatility (i.e. a lot of price spikes) and strong mean-reverting behaviour (volatility clustering followed by sustained periods of ‘normality’). The numerous spot price spikes act as outliers producing significant deviations in the empirical distribution function from Gaussianity. In fact, the spot price data displays the same predominant empirical ‘leptokurtosis’ feature of most high frequency asset price data – the tails of the empirical distribution functions are much fatter than those associated with normal distribution implying large fourth order cumulants.

In Foster, Hinich and Wild (2008), the extent of and stability of daily and weekly cycles in both load and spot price time series data was investigated. A major finding of that article was that the mean properties of both the load and spot price data for the NEM states considered were periodic. The most important periodicities for both datasets were found to contain significant but imperfect signal coherence suggesting that some ‘wobble’ existed in the waveforms of load and spot price data. This was determined by applying the Randomly Modulated Periodicity Model introduced in Hinich (2000) and Hinich and Wild (2001) to the data.

It was also originally postulated in Hinich (2000) and Hinich and Wild (2001) that the generating mechanism for an RMP process would be essentially nonlinear in character. Therefore, a natural research question is whether the mechanism responsible for generating both daily and weekly data exhibits some type of nonlinearity, and if so, whether this nonlinearity is ‘episodic’ in character. The rationale for the likely existence of episodic nonlinearity is that this type of behaviour would seem to be required if the commonly accepted 'stylised' fact of strong mean reversion in spot electricity prices, for example, is to eventuate.

Another reason why the finding of the presence of nonlinearity would be important is because this finding would effectively rule out many classes of linear models as candidates for modelling both load and spot price dynamics. Instead, the finding would suggest that attempts to fully model both daily and weekly dynamics would have to encompass models that could possibly generate nonlinear 'bursting' (in the case of spot prices) to model the episodic nonlinear serial dependence evident in the underlying data.

The article is organized as follows. In Section 2 we briefly discuss the data used and highlight some transformations that were made to the spot price electricity data in order to implement the tests considered in this article. In Section 3 we outline the portmanteau correlation, bicorrelation and tricorrelation tests employed in this article. These tests will be used to test for second-order (linear), third- and fourth-order (nonlinear) serial dependence, respectively. In Sections 4 we will briefly state the well-known Engle LM ARCH test that will be used to test for the presence of pure ARCH and GARCH structures in the daily and weekly waveforms. In Section 5, the empirical results for both the daily and the weekly waveforms will be presented. In Section 6, some concluding comments will be offered.

2 DATA AND ASSOCIATED TRANSFORMATIONS

In this article, we use half hourly spot electricity prices and load data for the period from 7/12/1998 to 29/02/2008.[1] This produced a sample size of 161786 observations.

[1] The half hourly load and spot price data were sourced from files located at the following web addresses: http://www.nemmco.com.au/data/aggPD_1998to1999.htm#aggprice1998link,

We apply the tests to time series load and spot price data from New South Wales (NSW), Queensland (QLD), Victoria (VIC) and South Australia (SA).

In applying the various tests outlined in this article, we convert all data series to continuous compounded returns by applying the relationship

$$r(t) = \ln\left(\frac{y(t)}{y(t-1)}\right) * 100, \tag{1}$$

where:

. $r(t)$ is the continuous compounded return for time period "t"; and

. $y(t)$ is the source price or load time series data.

In order to apply (1), $y(t)$ cannot take negative or zero values. However, it was evident that for Queensland, Victoria and South Australia, there was the occasional occurrence of negative spot prices.

In the presence of negative prices, some transformations had to be made to the respective price series to remove negative prices before we were able to apply (1) to convert the data to returns. This transformation involves two steps. First, any values which were negative or zero are set to the previous non-negative value using the following decision rule:

$$if \quad y(t) = \begin{cases} \leq 0, & x(t) = y(t-1) \\ else, & x(t) = y(t) \end{cases}, \tag{2}$$

---

http://www.nemmco.com.au/data/aggPD_2000to2005.htm#aggprice2000link, and
http://www.nemmco.com.au/data/aggPD_2006to2010.htm#aggprice2006link.

where $y(t)$ is the source time series data and $x(t)$ is the transformed data series. The second step involved applying a linear interpolation routine to the transformed series $x(t)$ obtained by using the following decision rule:

$$if \quad y(t) = \begin{cases} \leq 0, \; z(t) = \left\{ \dfrac{\left[x(t-1) + x(t+1) \right]}{2} \right\}, \\ else, \qquad z(t) = x(t) \; \left[= y(t) \right] \end{cases} \tag{3}$$

where $z(t)$ is the new transformed data (also see Foster, Hinich and Wild (2008)).

3 THE PORTMANTEAU CORRELATION, BICORRELATION AND TRICORRELATION TEST STATISTICS IN MOVING TIME WINDOWS FRAMEWORK

We utilize the framework originally proposed in Hinich and Patterson (1995), (now published as Hinich and Patterson (2005)) which seeks to detect epochs of transient serial dependence in a discrete-time pure white noise process (i.e. *i.i.d* random variates). A common approach to processing time series with a periodic structure is to partition the observations into non-overlapping frames where there is exactly one waveform in each sample (data) frame. This methodology involves computing the portmanteau correlation, bicorrelation and tricorrelation test statistics (denoted as C, H and $H4$ statistics) for each frame to detect linear and nonlinear serial dependence respectively.

Let the sequence $\{x(t)\}$ denote the sampled (and transformed) data process in (3), where the time unit 't' is an integer. The test procedure employs non-overlapped time frames (windows), thus if n is the frame length, then the k-th window is defined as $\{x(t_k), x(t_k+1), \ldots, x(t_k+n-1)\}$. The next non-overlapped window is

$\{x(t_{k+1}), x(t_{k+1}+1), ..., x(t_{k+1}+n-1)\}$, where $t_{k+1} = t_k + n.$ Define $Z(t)$ as the sequence of standardized observations given by

$$Z(t) = \frac{x(t) - m_x}{s_x} \tag{4}$$

for each $t = 1,2,...,n$ where m_x and s_x are the sample mean and standard deviation of the sample frame. As such, the data in each sample frame is standardised on a frame-by-frame basis.

The null hypothesis for each sample frame is that the transformed data $\{Z(t)\}$ are realizations of a stationary pure white noise process. Therefore, under the null hypothesis, the correlations $C_{ZZ}(r) = E[Z(t)Z(t+r)] = 0, \forall\ r \neq 0$, the bicorrelations $C_{ZZZ}(r,s) = E[Z(t)Z(t+r)Z(t+s)] = 0, \forall\ r, s$ except when $r = s = 0$, and the tricorrelations $C_{ZZZZ}(r,s,v) = E[Z(t)Z(t+r)Z(t+s)Z(t+v)] = 0, \forall\ r, s, \mathit{and}\ v$ except when $r = s = v = 0$. The alternative hypothesis is that the process in the sample frame has some non-zero correlations, bicorrelations or tricorrelations in the set $0 < r < s < v < L$, where L is the number of lags associated with the length of the sample frame. In other words, if there exists second-order (linear) or third- or fourth-order (nonlinear) serial dependence in the data generating process, then $C_{ZZ}(r) \neq 0, C_{ZZZ}(r,s) \neq 0, \mathit{or}\ C_{ZZZZ}(r,s,v) \neq 0$ for at least one r value or one pair of r and s values or one triple of $r, s\ \mathit{and}\ v$ values, respectively.

The r sample correlation coefficient is

$$C_{ZZ}(r) = \frac{1}{\sqrt{n-r}} \sum_{t=1}^{n-r} Z(t)Z(t+r). \tag{5}$$

The C statistic is designed to test for the existence of non-zero correlations (i.e. second-order linear dependence) within a sample frame, and its distribution is

$$C = \sum_{r=1}^{L} \left[C_{ZZ}(r)\right]^2 \approx \chi_L^2. \tag{6}$$

The (r,s) sample bicorrelation coefficient is

$$C_{ZZZ}(r,s) = \frac{1}{n-s} \sum_{t=1}^{n-s} Z(t)Z(t+r)Z(t+s), \text{ for } 0 \le r \le s. \tag{7}$$

The H statistic is designed to test for the existence of non-zero bicorrelations (i.e. third-order nonlinear serial dependence) within a sample frame, and its corresponding distribution is

$$H = \sum_{s=2}^{L} \sum_{r=1}^{s-1} G^2(r,s) \approx \chi_{L(L-1)/2}^2 \tag{8}$$

where $G(r,s) = \sqrt{n-s}\, C_{ZZZ}(r,s)$.

The (r,s,v) sample tricorrelation coefficient is

$$C_{ZZZZ}(r,s,v) = \frac{1}{n-v} \sum_{t=1}^{n-v} Z(t)Z(t+r)Z(t+s)Z(t+v), \text{ for } 0 \le r \le s \le v. \tag{9}$$

The $H4$ statistic is designed to test for the existence of non-zero tricorrelations (i.e. fourth-order nonlinear serial dependence) within a sample frame and its corresponding distribution is

$$H4 = \sum_{v=3}^{L} \sum_{s=2}^{v-1} \sum_{r=1}^{s-1} T^3(r,s,v) \approx \chi_{L(L-1)(L-2)/3}^2 \tag{10}$$

where $T(r,s,v) = \sqrt{n-v} \times C_{ZZZZ}(r,s,v)$.

Since it is conceptually difficult to quantify how much of any 'significant' autocorrelation can be attributed to thin trading volume or spot price limits, this investigation focuses instead on whether load and spot prices data contain predictable nonlinearities after removing all linear dependence. The autocorrelation structure in each sample frame is removed by an autoregressive AR(p) fit, where 'p' is the number of lags that is selected in order to remove significant C statistics at some pre-specified threshold level.[2] It is worth noting that the AR fitting is employed purely as a 'pre-whitening' operation and not in order to obtain a model of 'best fit'. The portmanteau bicorrelation and tricorrelation tests are then applied to the residuals of the fitted AR(p) model of each sample frame, so that any rejections of the null hypothesis of pure white noise can be attributed to significant H or $H4$ statistics.

The number of lags L is defined as $L = n^b$ with $0 < b < 0.5$ for the correlation and bicorrelation tests and $0 < b < 0.33$ for the tricorrelation test, and where b is a parameter to be chosen by the user. Based on results of Monte Carlo simulations, Hinich and Patterson (1995, 2005) recommended the use of $b = 0.4$ (in relation to the bicorrelation test) which is a good compromise between: (1) using the asymptotic result as a valid approximation for the sampling properties of the H statistic for moderate sample sizes; and (2) having enough sample bicorrelations in the statistic to have reasonable power against non-independent variates.

[2] In the literature particularly dealing with long-term dependence, pre-filtering by means of an AR-GARCH procedure is often used to remove short-term autocorrelation and time-varying volatility. However, this procedure is unnecessary, in the current context, since the bicorrelation and tricorrelation tests rely on the property that the bicorrelation and tricorrelation coefficients equal zero for a pure noise process. As such, the null hypothesis is only rejected when there exists some non-zero bicorrelations or tricorrelations suggesting nonlinear serial dependence in the conditional mean (additive nonlinearity), and not the presence of conditional variance dependence (conditional heteroskedasticity).

Another element that must be decided upon is the choice of the frame length. In principal, there is no unique value for the frame length. The larger the frame length, the larger the number of lags and hence the greater the power of the test, but at the 'expense' of increasing the uncertainty of the event time when the serial dependence 'episode' occurs.

In this article, the data is split into a set of equal-length non-overlapped moving frames of 48 and 336 half hour observations corresponding to a frame of a day and a week's duration, respectively.[3] Our objective with these particular choices is to measure the extent to which any observed nonlinearity that is episodically present in the data appears to be operating on a daily or weekly time scale.

We can also use the correlation, bicorrelation and tricorrelation tests to examine whether a GARCH formulation represents an adequate characterisation of the data under investigation. This is accomplished by transforming the returns data into a set of binary data according to

$$\{y(t)\}: \begin{matrix} y(t)=1, \quad if\ Z(t)\geq 0 \\ y(t)=-1, if\ Z(t)<0 \end{matrix}. \tag{11}$$

If $Z(t)$ is generated by a pure ARCH or GARCH process whose innovations are symmetrically distributed with zero mean, then the binary data set $\{y(t)\}$ will be a stationary pure noise (*i.i.d*) Bernoulli sequence. In essence, while $Z(t)$ (a symmetric GARCH process) is a martingale difference process, the binary transformation outlined in (11) converts it into a pure noise process (Lim, Hinich and Liew (2005, 269-70)) which

[3] In principle, this window length needs to be sufficiently long enough to validly apply the bicorrelation and tricorrelation tests and yet short enough for the data generating process to have remained roughly constant (see Monte Carlo results in Hinich, 1996; Hinich and Patterson, 1995, 2005).

has moments that are well behaved with respect to asymptotic theory (Hinich (1996)). Therefore, if the null of pure noise is rejected by the C, H or H4 tests when applied to binary data determined from (11), this then signifies the presence of structure in the data that cannot be modelled by GARCH models. Moreover, while the rejections might be because of the presence of serial dependence in the innovations, this outcome still violates a critical assumption underpinning the formulation of GARCH models. Specifically, if the innovations are dependent (not *i.i.d*), then the statistical properties of the parameter estimates of ARCH/GARCH processes are unknown (Bonilla, Meza and Hinich (2007, p. 2531)).

To implement the test procedures on a frame-by-frame basis, define a frame as significant with respect to the C, H or H4 tests if the null of pure noise is rejected by each of the respective tests for that particular sample frame at some pre-specified (false alarm) threshold. This threshold controls the probability of a TYPE I error, - that of falsely rejecting the null hypothesis when it is in fact true.[4] For example, if we adopt a false alarm threshold of 0.90, this would signify that we would expect random chance to produce false rejections of the null hypothesis of pure noise in 10 out of every 100 frames. In a similar way, false alarm thresholds of 0.95 and 0.99 would signify that false rejections of the null hypothesis of pure noise in 5 out of 100 frames and 1 out of 100 frames respectively could be attributed to random chance.

Thus, according to the above criteria, if we secure rejections of the test statistics at rates (significantly) exceeding 10%, 5% and 1% of the total number of sample frames

[4] The false alarm threshold is to be interpreted as a confidence level, for example, a false alarm threshold of 0.90 is to be interpreted as a 90% confidence level. The level of significance associated with this confidence level is interpreted in the conventional way as 1 minus the threshold value. Therefore, for a threshold of 0.9, we get a corresponding significance level 0.1 – that is, a significance level of 10%.

examined, then this would signify the presence of statistical structure, thus pointing to the presence of (significant) second, third and fourth order serial dependence in the data set.

In principal, the tests can be applied to either the source returns data determined from application of (3)-(4) or to residuals from frame based autoregressive fits of this data. Recall that the latter can be viewed as a 'pre-whitening' operation and can be used to effectively remove second order (linear) serial dependence producing no significant 'C frames'. In this case, any remaining serial dependence left in the residuals must be a consequence of nonlinearity that is episodically present in the data - thereby, only significant H and H4 statistics will lead to the rejection of the null hypothesis of a pure noise process.

4 ENGLE LM ARCH TEST

In this article, we will also investigate the issue of parameter instability of GARCH models and the transient nature of ARCH effects. The 'well-known' Engle LM test for Autoregressive and Conditional Heteroscedasticity (ARCH) in residuals of a linear model was originally proposed in Engle (1982). This test should have power against more general GARCH alternatives, see Bollerslev (1986). The test statistic is based on the R^2 of the following auxiliary regression

$$x_t^2 = \beta_0 + \sum_{i=1}^{p} \beta_i x_{t-i}^2 + \xi_t \,, \tag{12}$$

where x_t^2 are typically squared residuals from a linear regression. Therefore, equation (12) involves regressing the squared residuals on an intercept and its own p lags.

Under the null hypothesis of a linear generating mechanism for x_t, $\left(NR^2\right)$ from the regression outlined in (12) is asymptotically distributed as χ_p^2, where N is the number of sample observations and R^2 is the coefficient of multiple correlation from the regression in (12).[5]

The ARCH testing procedure that is applied in this article involves applying the LM test to the squared data in each sample frame. As in the case of the application of C, H and H4 statistics on a frame-by-frame basis, this data will typically be the (squared) residuals from a frame-by-frame 'pre-whitening' AR(p) fit in the case of the ARCH LM test.

One key aspect of interest with this test procedure will be to determine whether there is a strong ARCH effect over all time periods (i.e. all sample frames) or whether ARCH is present only for short periods of time, for example, in a relatively small number of sample frames. It should also be noted that the same arguments made in the previous section in relation to false alarm thresholds and extent of rejections that can be attributed to random chance will continue to hold in this current case. As such, significant ARCH effects will arise if the percentage of framed based ARCH LM test rejections is significantly greater than the significance levels associated with the pre-specified false alarm threshold.

The ARCH test will only be applied to the spot price data. The load data does not exhibit any 'volatility clustering' affects that generate the conditional variance dependence (conditional heteroskedasticity) that the ARCH test is designed to identify.

---

[5] For the test to be valid, $E\left[x_t^8\right] < \infty$, that is, the eighth order moments must exist.

The spot price data, on the other hand, does display the type of patterns conventionally associated with conditional heteroskedasticity. This is why the authors argued in Foster, Hinich and Wild (2008) that a (periodic) mean plus volatility model for spot prices forecasting might have advantages over existing modelling approaches.

5 EMPIRICAL RESULTS

In Tables 1 and 2, the summary statistics of the NEM State load and spot price returns series are documented. It is apparent from inspection of both tables that the mean of the series are very small in magnitude. In Table 1, the mean 'returns' for the load data, on average, are all positive while the average returns for the four spot price returns series listed in Table 2 were negative over the complete sample. A 'difference in scale' can also be observed from an investigation of the maximum and minimum values of the respective returns series. For the load data, the maximum and minimum returns are in the ranges between 30 and 50 percent in absolute terms while the corresponding results for the spot price returns are of the order of 480 to 610 percent. Moreover, the differences in the values of the sixth order cumulants listed in both tables also reinforce the obvious difference in scale of the different series. Specifically, the sixth order cumulants of the spot price returns cited in Table 2 are much larger in magnitude than those listed in Table 1 associated with load returns.

It is also evident from inspection of both tables that the spot price returns are more volatile when compared with load data as indicated by the higher standard deviations documented in Table 2 when compared with those listed in Table 1. This indicates that the likely 'risk profile' of the load and spot price returns will be quite different.

Furthermore, volatility in both load and spot prices is slightly higher for SA than for the other three states considered - SA has the highest standard deviations for both load and spot price returns data.

All of the series except for SA spot price returns display positive or right skewness. All of the series also display evidence of leptokurtosis although this is a much more prominent feature in the case of the spot price returns data with excess kurtosis values in the ranges of 72 to 104 in magnitude. This implies that the tails of the empirical distribution functions of the spot prices returns in particular taper down to zero much more gradually than would the tails of the normal distribution (Lim, Hinich, Liew (2005, p.270)). Not unexpectedly, the Jarques-Bera (JB) Normality Test for all of the returns series listed in both tables indicates that the null hypothesis of normality is strongly rejected at the conventional 1% level of significance. This outcome reflects the strong evidence of both non-zero skewness and excess kurtosis listed in both tables.

Table 3 presents the results for the correlation ©, bicorrelation (H) and tricorrelation (H4) test statistics for the load returns data for a weekly sample frame of 336 (half hourly) observations. In all results reported in this section, bootstrapped threshold values were used because the sample properties of the test statistics for very small frame lengths do not necessarily closely approximate the theoretical thresholds especially when the underlying sample data contains both significant non-zero skewness and excess kurtosis, as in the current case.

The bootstrapped threshold values were determined in the following manner. Given the 'global' sample of 161785 returns for each respective series, a bootstrap 'sample frame' was constructed by randomly sampling 336 observations from the larger 'global population' and the various test statistic outcomes were calculated for that particular sample frame. This process was repeated 500000 times and the results for each test statistic were stored in an array. All test statistics entail application of the chi-square distribution and for each bootstrap replication, the chi square levels variable associated with each test statistic was transformed to a uniform variate which means, for example, that the 10% threshold corresponds to 0.90, the 5% threshold is 0.95, and the 1% threshold is 0.99 and the 'transformed' test statistic thresholds are now in the interval $(0,1)$. The arrays containing the bootstrap 'thresholds' for each respective test statistic (containing 500000 elements) from the bootstrap process was then sorted in ascending order and the bootstrap threshold was calculated as the 'quantile' value of the empirical distribution function of the various test statistics associated with the user specified 'false alarm' threshold value. For example, if the user set the 'false alarm' threshold value to 0.90, the bootstrap threshold value would be the 90% 'quantile' of the empirical distribution function of the relevant test statistic determined from the bootstrap process.

The number of frame based rejections for each test statistic is calculated by summing the number of frames over which rejections were secured at the calculated bootstrap threshold when the tests are applied on a sequential frame by frame basis to the actual returns data. A frame based rejection is secured if for an actual frame, the calculated threshold value exceeds the bootstrap determined 'false alarm' threshold. For example, suppose the bootstrap 'false alarm' threshold was determined to be 0.92 (say for a user specified 'false alarm' threshold value of 0.90), then if the calculated threshold value for the relevant test statistic exceeded 0.92, (say 0.94) then we would secure rejection of the null hypothesis for the test statistic for that frame at the 10% level of significance, thus securing a frame based rejection of that test statistic.[6] The percentage of frame rejections for each test statistic is calculated as the total number of frame based rejections computed as a percentage of the total number of frames.

The results for the weekly load returns data presented in Table 3 were determined after applying a 'global AR(340) fit' to the complete sample data.[7] This operation was employed purely to remove second order serial dependence and the AR lag length of 340 was deliberately chosen to exceed the weekly frame length of 336 observations. This regression can be viewed as essentially a type of weekly 'detrending' operation and operates to remove the mean weekly periodicity from the underlying data series. The residuals from this AR fit are then used to determine the bootstrap thresholds and underpin other empirical results obtained for the load returns data. To further eliminate second order serial dependence, an 'AR(10)' fit is applied on a frame by frame basis. The success of these combined 'prewhitening' operations can be seen when inspecting Table 3 by the fact that no significant C frames were found (see Column 4 of Table 3). This outcome noticeably contrasts with the much more significant number of H and H4 frames that were found to be significant – see Columns 5 and 6 of Table 3).

Recall that for the false alarm thresholds of 0.90, 0.95 and 0.99 respectively, we expect only 10%, 5% and 1% of the total number of frames to secure rejections that can be reasonably attributed to random chance. The fact that the actual number of rejections are much higher than 10%, 5% and 1% of the total number of frames for both the H and

---

[6] We term such frames 'significant' frames with respect to the relevant test statistic.

[7] For the frame length of 336, the number of lags employed for the C and H statistics were determined to be 10 and the number of lags for the H4 test was determined to be 6. The number of bicovariances and tricovariances used were determined to be 45 and 20 respectively.

H4 tests signify the existence of statistical significant third-order and fourth-order (nonlinear) serial dependence in the load returns data, thus confirming the presence of a nonlinear generating mechanism characterizing weekly load dynamics.

For example, in Table 3, for NSW, we secure rejections the null hypothesis of pure white noise in the case of the bicorrelation H statistic (see column 5) for 303, 234 and 135 frames at the bootstrapped determined significance levels of 10%, 5% and 1%, respectively. This, in turn, amounts to 62.99%, 48.65% and 28.07% of the total number of frames considered. These values substantially exceed the 10%, 5% and 1% rejection rates we can reasonably attribute to random chance, thus pointing to the presence of significant third order nonlinear serial dependence in the NSW load returns data. However, the fact that we do not secure rejections for all frames points to the third order nonlinear serial dependence being episodic in character - there are also many frames where the null hypothesis of pure noise cannot be rejected. Similar interpretations can be given to all other test results cited in Table 3.

It is also apparent from inspection of Table 3 that fourth-order serial dependence seems to be a more prevalent feature in the data than third-order serial dependence – the number of frame based rejections for the H4 test (column 6) generally exceeds the number of frame based rejections for the H test (column 5) at all three bootstrap false alarm thresholds reported in Table 3. It should also be noted that for the 0.99 threshold for the H4 statistic for the states of NSW and VIC, we had to set the false alarm threshold to 0.999999 because the bootstrapped values tended to be very high and ‘crowded out’ actual applications to the data.[8] The 'set' threshold value of 0.999999 is still very conservative when interpreted at the conventional 1% level of significance. As such, the significant number of frame based rejections is quite believable, thus confirming the presence of significant fourth order (nonlinear) serial dependence. Finally, we secure the least number of H and H4 test based frame rejections for the state of SA. This indicates that the nonlinear serial dependence is a less prominent feature of the load returns data associated with SA when compared with the other three states of NSW, QLD and VIC. This might reflect the combined effects of higher levels of overall load demand and resulting implications for power system security and reliability in the face of, for example, weather variations, unforeseen power generation outages and increased market interconnectedness. These factors might combine to produce power fluctuations capable of generating nonlinear events on an episodic basis that arise to a greater extent in NEM states other than SA which is a relatively smaller state and is more isolated within the NEM from the major NSW market than is QLD, for example.

In Table 4, the results for the three portmanteau tests, and additionally the LM ARCH test are presented for the spot price returns. In this case, no 'global prewhitening' was undertaken (in contrast with the load returns) although the frame by frame based 'AR(10) prewhitening' fit continued to be employed, thus suggesting a different type of 'dynamic' driving the mean periodicity of the spot price returns data when compared with the load returns data. It is evident from inspection of Table 4 (Column 4) that the 'prewhitening' operation has been successful – very few 'significant C frames' were discovered amounting to less than 1% of the total number of frames considered. However, there is a

[8] This result seems to be driven by outliers in the data and disappears when trimming is employed to reduce the impact of outliers without altering the qualitative conclusions made above about the presence of significant fourth-order nonlinear serial dependence.

lot of evidence of significant H and H4 based frame rejections reported in Columns 5 and 6) of the table. The nature of the rejections indicates that both third- and fourth-order nonlinear serial dependence is much more prominent in the spot price returns data than was the case for the load returns data – the extent of the frame based H and H4 rejections are in the range of 70%-95% for all states and all bootstrapped false alarm thresholds considered. This can be compared with the corresponding 20%-80% range for the load returns data displayed in Table 3. Furthermore, the H and H4 rejections are of similar orders of magnitude suggesting that one form of nonlinear serial dependence is not more prominent than the other form.

The frame by frame LM ARCH tests also signify the presence of pure ARCH/GARCH structure in the spot price returns data to a degree that exceeds what can be reasonably attributed to random chance. The order of magnitude, while significant, however, is of a lower order than associated with both H and H4 based frame rejections, particularly at the 5% and 1% levels of significance.

The results associated with the 'hard clipping' transformation applied to the residuals from the frame by frame 'AR(10)' fits applied to the actual spot price returns are documented in Table 5. These residuals are the same set of data that underpins the results cited in Table 4 except in the current case, the transformation in (11) was subsequently applied to the residuals prior to applying the portmanteau tests and with the ARCH LM test being dropped. Recall that the intention of this particular test framework is to see if 'non- GARCH' generating mechanisms are in operation in explaining weekly spot price returns dynamics. It is evident from inspection of Table 5 that the number of frame based rejections for the C, H and H4 statistics applied to the binary data sets are greater than the 10%, 5% and 1% rates we can reasonably attribute to random chance, thus pointing to the presence of a 'non-GARCH' generating mechanisms. Therefore, there is evidence pointing to structures in the weekly spot price returns data that cannot be modeled by a pure ARCH or GARCH model. However, it should be noted that the extent of frame based rejections is of a lower order of magnitude than those cited in Table 4, being in the range of 20%-60%. Furthermore, it is also apparent from Table 5 (Column 5) that the relatively larger number H statistic rejections indicates that the presence of third-order nonlinear serial dependence appears to be the most prominent type of nonlinear serial dependence present in the binary data.

It should also be noted once again that for all states listed in Table 4 that the false alarm threshold corresponding to 0.99 threshold for H, H4 and ARCH LM tests had to be set to either '0.9999' or '0.999999' because the bootstrapped values tended to be very high and 'crowded out' actual applications to the data.[9] The 'set' threshold values of '0.9999' or '0.999999' are very conservative when interpreted at the conventional 1% level of significance (associated with a threshold value of 0.99). As such, the significant number of frame based rejections is quite believable, thus confirming the presence of significant third and fourth order (nonlinear) serial dependence.

In this article, we also investigate the presence and nature of any nonlinear serial dependence evident in the dynamics of the daily load and spot price returns. This is accomplished by choosing an underlying frame length of 48 (half hours) which constitutes a time period of a day. The resulting analysis proceeds as before except that

[9] A similar practice had to be adopted for QLD 'C Statistic' results in Table 5 at the 0.95 and 0.99 threshold values. These results, once again, appear to be driven by outliers in the data and disappear when trimming is employed to reduce the impact of outliers without altering the qualitative conclusions made above about the presence of significant second, third or fourth-order serial dependence.

now the frame length is set to 48 instead of 336. This means that we get an increase in the total number of frames under investigation, increasing from 481 in the case of weekly returns data to 3370 frames in the case of the daily returns data. [10]

The results for the daily load returns are listed in Table 6. Once again, we employ a 'global AR(340) prewhitening fit' to remove the mean weekly periodicity. In performing this operation, we also remove the mean daily periodicity because this periodicity is a harmonic of the weekly periodicity. We also adopt a frame-by-frame 'AR(5) prewhitening' fit. These combined 'prewhitening' operations ensure that the number of 'significant C frames' that were very small in magnitude – less than 0.3 of one percent of the total number of frames considered - (see Column 4 of Table 6). This mirrors the results obtained in Table 3 in relation to the weekly data. As such, second-order (linear) serial dependence has been removed through the combined 'prewhitening' process and any further rejections of the null hypothesis of pure white noise will be attributable to either H or H4 based rejections indicating the presence of third- or fourth-order (nonlinear) serial dependence.

As in the case of the weekly returns results reported in Table 3, there is evidence of nonlinear serial dependence but now at a lower order of magnitude in the case of the daily load returns results reported in Table 6. Specifically, the frame based rejection for the H and H4 test statistics now occur at rates in the range of 5%-30% compared against the 20%-80% range associated with the weekly load returns data displayed in Table 3. Moreover, inspection of Table 6 also indicates that neither of the evident third- or fourth-

---

[10] For the frame length of 48, the number of lags employed for the C and H statistics were determined to be 5 and the number of lags for the H4 test was determined to be 3. The number of bicovariances and tricovariances used were determined to be 10 and 1 respectively.

order nonlinear serial dependence is the more prominent form – a situation that is different to the results reported in Table 3 for the weekly load returns. However, as in the case of the weekly load returns data, the results for the number of H and H4 test rejections is slightly smaller for SA when compared with the other three states. This broad result mirrors the same result discernible from Table 3.

The results for the daily spot price returns are reported in Table 7. We adopt the same ‘prewhitening’ scheme that was adopted for the weekly spot price returns – no global ‘prewhitening’ but a frame by frame based ‘AR(5) prewhitening’ fit – in order to remove second order (linear) serial dependence. This outcome can be seen by observing the very low number of ‘significant C frame’ reported in Column 4 of Table 7 – the number of significant C frames amounts to less than 0.2 of one percent of the total number of frames. There is also evidence of the presence of nonlinear serial dependence – the number of significant H and H4 frames significantly exceeds the 10%, 5% and 1% rates that can be reasonably attributed to random chance. The order of magnitude of the frame based rejections for H and H4 are in the range of 10%-55% which is smaller than the corresponding range in Table 4 of 70% to 95%. Thus, the presence of nonlinear serial dependence is a less prominent feature of the daily spot price returns data when compared with the weekly spot price returns.

Inspection of the last column of Table 7 also indicates the presence of significant GARCH structure although at a level that is much less prevalent when compared to the results associated with the weekly returns. The results associated with the ‘hard clipped’ transformation applied to the residuals of the frame-by-frame based ‘AR(5)’ fits are reported in Table 8.[11] It is apparent from inspection of this table that we cannot secure rates of rejection that point to the presence of 'non-GARCH' alternatives at the accepted significance levels.

Overall, the results suggest that nonlinear serial dependence plays a much less prominent role in explaining the evolution of daily spot price returns dynamics when compared against the results for the weekly returns. The ARCH LM test results cited in Table 7 indicate that GARCH effects play some role in explaining nonlinearity evident in daily spot price returns. This result is further confirmed by the 'hard clipping' results reported in Table 8 that appear to indicate that a lack of a presence of 'non-GARCH' alternatives in describing daily spot price return dynamics. This conclusion, however, should be tempered by the 'diminished' overall presence of nonlinear serial dependence operating at a daily time scale when compared, for example, with the weekly results.

Thus, a definite type of 'time scale' effect appears to be in operation. The prominence and role of nonlinear serial dependence appears to play a much greater role in explaining dynamics in both load and spot price returns dynamics over a weekly time scale rather than a daily time scale. This backs up the results reported in Brooks and Hinich (1998, p. 721) and Ammermann and Patterson (2003, p.188) in relation to the application of LM ARCH test of a frame-by-frame basis. Specifically, what we are seeing is that for very small frame lengths (i.e. of a day), there is increasingly long period during which there is no evidence of linear or non-linear serial dependence including ARCH effects in spot

---

[11] Note that the false alarm threshold corresponding to 0.99 for the QLD 'C statistic' had to be set to '0.999999' because the bootstrapped values tended to be very high and 'crowded out' actual applications to the data. As with the weekly data, this result appeared to be driven by outliers in the data and disappears when trimming is employed to reduce the impact of outliers.

price returns. Thus, the incidence of nonlinear serial dependence is very episodic at this particular time scale. However, as the frame length is aggregated (i.e. increased to a week), the effects become absorbed into periods containing both linear and nonlinear structures producing the increased incidence of frame based rejections of H, H4 and ARCH LM tests.

However, unlike the findings in Brooks and Hinich (1998) and Ammermann and Patterson (2003), the extent of aggregation from a day to a week is not that large within the context of the overall sample being considered and the extent of the relatively large number of frame based rejections cited in Tables 3-5, in particular, do indicate the significance presence of nonlinear serial dependence operating on a weekly time scale that perhaps has not been observed in other studies utilizing the test methods employed in this article. This suggests that while the nonlinear serial dependence is still episodic in character, it is less episodic in the current context when compared against similar studies undertaken using a wide assortment of high frequency finance based data as cited, for example, in Hinich and Patterson (1989, 2005), Brooks (1996), Brooks and Hinich (1998), Ammermann and Patterson (2003), Lim , Hinich and Liew (2003, 2004, 2005), Lim and Hinich (2005a, 2005b), Bonilla, Romero-Meza and Hinich (2007) and Hinich and Serletis (2007). Specifically, the extent of the nonlinear serial dependence observed for the weekly data on a frame-by-frame basis is unprecedented, pointing to a 'time scale effect' in the underlying data generating process whereby the nonlinearity strongly persists on a weekly time scale but to a much weaker degree on a daily time scale. It's effect most likely reflects the strong weekly periodicities found in both the load and spot price data cited in Foster, Hinich and Wild (2008).

6 CONCLUDING COMMENTS

In this article, an investigation was undertaken into whether nonlinear serial dependence was present in NEM State daily and weekly load and spot price data. This task was accomplished by applying the portmanteau correlation, bicorrelation and tricorrelation tests introduced in Hinich (1996) to the time series of half hourly spot prices and load demand from 7/12/1998 to 29/02/2008. The data corresponds to load and spot price time series data for the NEM states of New South Wales (NSW), Queensland (QLD), Victoria (VIC) and South Australia (SA).

These tests can be used to detect epochs of transient serial dependence in a discrete-time pure white noise process. The test framework involves partitioning the time series data into non-overlapping frames and computing the portmanteau correlation, bicorrelation and tricorrelation test statistics for each frame to detect linear and nonlinear serial dependence respectively. Furthermore, the presence of pure ARCH and GARCH effects in the spot price returns was also investigated by applying the Engle LM ARCH test and, additionally, a detection framework based upon converting a martingale difference process into a pure noise process and then testing for the presence of linear and nonlinear serial dependence in the transformed data.

Nonlinear serial dependence was found to be present in both daily and weekly load and spot price returns data considered in this article. However, a 'time scale' effect was found to be present. Specifically, nonlinear serial dependence was found to be a much more prominent feature in both the load and spot price returns dynamics over a weekly time scale rather than a daily time scale. At the daily time scale, the observed nonlinear serial dependence was found to be particularly episodic in nature - there is increasingly long periods during which there is no evidence of linear or non-linear serial dependence including ARCH effects in spot price dynamics followed by episodes on nonlinear dependence of limited duration. Moreover, GARCH effects appeared to be a more prominent feature in explaining the daily dynamics of spot price returns than was the case for the weekly dynamics.

At a weekly time scale, the results cited in Tables 3-5 indicate the significance presence of nonlinear serial dependence. While the nature of this dependence is still episodic in character, it is much less episodic (i.e. more universal) when compared, for example, to similar results from studies undertaken using an assortment of high frequency finance based returns data.

This finding also most likely reflects the strong weekly periodicities found in both the load ands spot price data and which were identified in Foster, Hinich and Wild (2008) using the RMP model. The finding of nonlinearity provides some added support for the proposition made in Hinich (2000) and Hinich and Wild (2001) that the generating mechanism for an RMP process would be essentially nonlinear in character. The added finding of episodic nonlinearity would also seem to be required if the commonly accepted 'stylised' fact of strong mean reversion in spot electricity prices, in particular, is to be obtained.

The findings of nonlinearity have implications for modeling weekly and daily load and spot price dynamics. Given the prevalence of both third and fourth-order nonlinear serial dependence in the data, it seems that time series models that employ a linear structure or assume a pure noise input such as GBM stochastic diffusion models would be problematic. In particular, the dependence structure would violate both normality and Markovian assumptions underpinning conventional GBM models.

An important research question is whether the nonlinear structure is a ‘deep structure’ – that is, whether or not it is driven solely by the presence of outliers. Trimming can be employed to investigate this issue. In particular, trimming can be used to control for the affects of outliers on the small sample properties of the various test statistics considered in this article as well as improving the small sample performance of the test statistics when viewed against the theoretical distribution defined across a wide assortment of quantiles. If this research found that the finding of nonlinear serial dependence was not sensitive to trimming scenarios, this would be indicative of deep structure that was not driven purely by the presence of outliers in the data. In this case, the validity of jump diffusion models which employ the ‘Poisson Process’ to model the probability of ‘outlier (i.e. jump) events’ will not be able to fully or adequately capture the nonlinearity present in the data. This research agenda would be important because both the GBM and jump diffusion models currently underpin accepted risk management strategies based on the ‘Black-Scholes Option Pricing Model’ that are employed in both the finance and electricity industry.

Finally, the episodic nature of the nonlinear serial dependence in the data also raises questions over what type of nonlinear time series model would be capable of generating this type of behavior, given that most nonlinear models posit a universal nonlinear generating mechanism.

REFERENCES

Ammermann, P.A. and D.M. Patterson (2003). "The cross-sectional and cross-temporal universality of nonlinear serial dependencies: Evidence from world stock indices and the Taiwan Stock Exchange." *Pacific-Basin Finance Journal*, 11:175-195.

Bollerslev, T. (1986). "Generalised autoregressive conditional heteroskedasticity." *Journal of Econometrics*, 31: 307-327.

Bonilla, C.A. Romero-Meza, R. and M.J. Hinich (2007). "GARCH inadequacy for modelling exchange rates: empirical evidence from Latin America." *Applied Economics*, 39: 2529-2533.

Brooks, C. (1996). "Testing for non-linearity in daily sterling exchange rates." *Applied Financial Economics*, 6: 307-317.

Brooks, C. and M.J. Hinich (1998). "Episodic nonstationarity in exchange rates." *Applied Economic Letters*, 5: 719-722.

Engle, R.F. (1982). "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation." *Econometrica*, 50: 87-1007.

Foster, J. Hinich, M.J. and P. Wild (2008). "Randomly Modulated Periodic Signals in Australia's National Electricity Market." *The Energy Journal*, 29: 105-130.

Hinich, M.J. and D.M. Patterson (1995)."Detecting Epochs of Transient Dependence in White Noise." *Mimeo,* University of Texas at Austin.

Hinich, M.J. (1996). "Testing for dependence in the input to a linear time series model." *Journal of Nonparametric Statistics*, 6: 205-221.

Hinich, M.J. (2000). "A Statistical Theory of Signal Coherence." *Journal of Oceanic Engineering*, 25: 256-261.

Hinich, M.J. and P. Wild (2001). "Testing Time-Series Stationarity Against an Alternative Whose Mean is Periodic." *Macroeconomic Dynamics*, 5: 380-412.

Hinich, M.J. and D.M. Patterson (2005)."Detecting Epochs of Transient Dependence in White Noise." In Belongia, M. and J. Binner, eds., *Money, Measurement and Computation*. London: Palgrave

Hinich, M.J. and A. Serletis (2007). "Episodic nonlinear event detection in the Canadian exchange rate." *Journal of the American Statistical Association*, *Applications and Case Studies*, 102 (477): 68-74.

Lim, K.P. and M.J. Hinich (2005a). "Cross-temporal universality of non-linear dependencies in Asian stock markets." *Economics Bulletin*, 7(1): 1-6.

Lim, K.P. and M.J. Hinich (2005b). "Non-linear market behavior: events detection in the Malaysian stock market." *Economics Bulletin*, 7(6): 1-5.

Lim, K.P. Hinich, M.J. and V.K.S. Liew (2003). "Episodic non-linearity and non-stationarity in ASEAN exchange rates returns series." *Labuan Bulletin of International Business and Finance*, 1(2): 79-93.

Lim, K.P. and V.K.S. Liew (2004). "Non-linearity in financial markets: evidence from ASEAN-5 exchange rates and stock markets." *ICFAI Journal of Applied Finance* 10(5): 5-18.

Lim, K-P. Hinich, M,J. and V. K. Liew (2005). "Statistical Inadequacy of GARCH Models for Asian Stock Markets: Evidence and Implications." *Journal of Emerging Market Finance,* 4(3): 263-279.

National Electricity Market Management Company Limited (NEMMCO) (2005). "An Introduction to Australia's National Electricity Market." National Electricity Market Management Company Ltd, June 2005. (http://www.nemmco.com.au/nemgeneral/000-0187.pdf).

Table 1. Summary Statistics for Load Returns Data

	NSW	QLD	VIC	SA
No of Observations	161785	161785	161785	161785
Mean	0.002	0.003	0.003	0.002
Maximum	36.80	42.3	40.2	32.5
Minimum	-30.90	-38.5	-41.3	-49.4
Std Dev	3.05	2.80	2.92	3.37
Skewness	1.01	0.83	0.96	0.29
Excess Kurtosis	1.54	1.77	1.45	1.27
6^{th} Order Cumulant	11.94	95.60	74.90	65.53
JB Test Statistic	43500.0	39400.0	38700.0	13100.0
JB Normality P-Value	0.0000	0.0000	0.0000	0.0000

Table 2. Summary Statistics for Spot Price Returns Data

	NSW	QLD	VIC	SA
No of Observations	161785	161785	161785	161785
Mean	-0.002	-0.0003	-0.002	-0.002
Maximum	545.0	591.0	497.0	597.0
Minimum	-572.0	-531.0	-488.0	-610.0
Std Dev	19.2	26.2	20.4	26.7
Skewness	0.49	0.29	0.36	-0.33
Excess Kurtosis	104.0	73.5	71.8	74.5
6^{th} Order Cumulant	41958.9	13650.2	21388.0	15187.0
JB Test Statistic	72900000.0	36300000.0	34700000.0	37300000.0
JB Normality P-Value	0.0000	0.0000	0.0000	0.0000

Table 3. Frame Test Results for Weekly Load Demand (Returns) Data
Specific Details: Removed 'Weekly Mean' By Global AR(340) 'Prewhitening' Fit; Applied Frame by Frame AR(10) 'Prewhitening' Fit to Remove Linear Dependence

Scenario / (State)	Total Num of Frames	False Alarm Threshold	Significant C Frames Num & (%)	Significant H Frames Num & (%)	Significant H4 Frames Num & (%)
NSW	481	0.90	0 (0.00%)	303 (62.99%)	400 (83.16%)
	481	0.95	0 (0.00%)	234 (48.65%)	345 (71.73%)
	481	0.99	0 (0.00%)	135 (28.07%)	259** (53.85%)
QLD	481	0.90	0 (0.00%)	261 (54.26%)	326 (67.78%)
	481	0.95	0 (0.00%)	196 (40.75%)	286 (59.46%)
	481	0.99	0 (0.00%)	101 (21.00%)	174 (36.17%)
VIC	481	0.90	0 (0.00%)	300 (62.37%)	428 (88.98%)
	481	0.95	0 (0.00%)	241 (50.10%)	379 (78.79%)
	481	0.99	0 (0.00%)	123 (25.57%)	286** (59.46%)
SA	481	0.90	0 (0.00%)	257 (53.43%)	337 (70.06%)
	481	0.95	0 (0.00%)	199 (41.37%)	271 (56.34%)
	481	0.99	0 (0.00%)	94 (19.54%)	139 (28.90%)

Notes:

** - false alarm threshold is set to 0.999999.

Table 4. Frame Test Results for Weekly Spot Price (Returns) Data
Specific Details: No Global AR 'Prewhitening' Fit;
Applied Frame by Frame AR(10) 'Prewhitening' Fit to Remove Linear Dependence

Scenario / (State)	Total Num of Frames	False Alarm Threshold	Significant C Frames Num & (%)	Significant H Frames Num & (%)	Significant H4 Frames Num & (%)	Significant ARCH Frames Num & (%)
NSW	481	0.90	1 (0.21%)	464 (96.47%)	452 (93.97%)	395 (82.12%)
	481	0.95	1 (0.21%)	406 (84.41%)	396 (82.33%)	326 (67.78%)
	481	0.99	1 (0.21%)	358* (74.43%)	358* (74.43%)	232** (48.23%)
QLD	481	0.90	4 (0.83%)	449 (93.35%)	478 (99.38%)	411 (85.45%)
	481	0.95	3 (0.62%)	380 (79.00%)	440 (91.48%)	327 (67.98%)
	481	0.99	2 (0.42%)	338** (70.27%)	401* (83.37%)	245* (50.94%)
VIC	481	0.90	1 (0.21%)	461 (95.84%)	461 (95.84%)	410 (85.24%)
	481	0.95	1 (0.21%)	438 (91.06%)	421 (87.53%)	334 (69.44%)
	481	0.99	1 (0.21%)	390* (81.08%)	386* (80.25%)	199* (41.37%)
SA	481	0.90	2 (0.42%)	453 (94.18%)	462 (96.05%)	413 (85.86%)
	481	0.95	1 (0.21%)	385 (80.04%)	424 (88.15%)	332 (69.02%)
	481	0.99	1 (0.21%)	316** (65.70%)	376* (78.17%)	250* (51.98%)

Notes:

* - false alarm threshold is set to 0.9999.
** - false alarm threshold is set to 0.999999.

Table 5. Frame Test Results for Weekly Spot Price (Returns) Data

Specific Details: No Global AR 'Prewhitening' Fit;
Applied Frame by Frame AR(10) 'Prewhitening' Fit to Remove Linear Dependence
Frame by frame Hard Clipping of Residuals

Scenario / (State)	Total Num of Frames	False Alarm Threshold	Significant C Frames Num & (%)	Significant H Frames Num & (%)	Significant H4 Frames Num & (%)
NSW	481	0.90	176 (36.59%)	272 (56.55%)	149 (30.98%)
	481	0.95	111 (23.08%)	212 (44.07%)	106 (22.04%)
	481	0.99	52 (10.81%)	98 (20.37%)	55 (11.43%)
QLD	481	0.90	192 (39.92%)	243 (50.52%)	190 (39.50%)
	481	0.95	171* (35.55%)	161 (33.47%)	151 (31.39%)
	481	0.99	123** (25.57%)	46 (9.56%)	102 (21.21%)
VIC	481	0.90	218 (45.32%)	311 (64.66%)	178 (37.01%)
	481	0.95	172 (35.76%)	270 (56.13%)	132 (27.44%)
	481	0.99	85 (17.67%)	162 (33.68%)	67 (13.93%)
SA	481	0.90	248 (51.56%)	267 (55.51%)	193 (40.12%)
	481	0.95	190 (39.50%)	214 (44.49%)	158 (32.85%)
	481	0.99	136** (28.27%)	94 (19.54%)	113 (23.49%)

Notes:

* - false alarm threshold is set to 0.9999.
** - false alarm threshold is set to 0.999999.

Table 6. Frame Test Results for Daily Load Demand (Returns) Data
Specific Details: Removed 'Weekly Mean' By Global AR(340) 'Prewhitening' Fit;
Applied Frame by Frame AR(5) 'Prewhitening' Fit to Remove Linear Dependence

Scenario / (State)	Total Num of Frames	False Alarm Threshold	Significant C Frames Num & (%)	Significant H Frames Num & (%)	Significant H4 Frames Num & (%)
NSW	3370	0.90	2 (0.06%)	858 (25.46%)	717 (21.28%)
	3370	0.95	2 (0.06%)	497 (14.75%)	397 (11.78%)
	3770	0.99	1 (0.03%)	151 (4.48%)	128 (3.80%)
QLD	3370	0.90	9 (0.27%)	744 (22.08%)	580 (17.21%)
	3370	0.95	8 (0.24%)	427 (12.67%)	351 (10.42%)
	3370	0.99	6 (0.18%)	137 (4.07%)	95 (2.82%)
VIC	3370	0.90	1 (0.03%)	977 (28.99%)	795 (23.59%)
	3370	0.95	0 (0.00%)	587 (17.42%)	479 (14.21%)
	3370	0.99	0 (0.00%)	188 (5.58%)	148 (4.39%)
SA	3370	0.90	0 (0.00%)	673 (19.97%)	602 (17.86%)
	3370	0.95	0 (0.00%)	360 (10.68%)	330 (9.79%)
	3370	0.99	0 (0.00%)	81 (2.40%)	91 (2.70%)

Table 7. Frame Test Results for Daily Spot Price (Returns) Data
Specific Details: No Global AR 'Prewhitening' Fit;
Applied Frame by Frame AR(5) 'Prewhitening' Fit to Remove Linear Dependence

Scenario / (State)	Total Num of Frames	False Alarm Threshold	Significant C Frames Num & (%)	Significant H Frames Num & (%)	Significant H4 Frames Num & (%)	Significant ARCH Frames Num & (%)
NSW	3370	0.90	1 (0.03%)	1373 (40.74%)	1394 (41.36%)	993 (29.47%)
	3370	0.95	1 (0.03%)	899 (26.68%)	966 (28.66%)	619 (18.37%)
	3370	0.99	0 (0.00%)	312 (9.26%)	375 (11.13%)	147 (4.36%)
QLD	3370	0.90	5 (0.15%)	1839 (54.57%)	1836 (54.48%)	1202 (35.67%)
	3370	0.95	2 (0.06%)	1168 (34.66%)	1338 (39.70%)	738 (21.90%)
	3370	0.99	0 (0.00%)	397 (11.78%)	518 (15.37%)	127 (3.77%)
VIC	3370	0.90	1 (0.03%)	1435 (42.58%)	1442 (42.79%)	978 (29.02%)
	3370	0.95	0 (0.00%)	922 (27.36%)	1031 (30.59%)	629 (18.66%)
	3370	0.99	0 (0.00%)	340 (10.09%)	379 (11.25%)	160 (4.75%)
SA	3370	0.90	6 (0.18%)	1543 (45.79%)	1529 (45.37%)	1028 (30.50%)
	3370	0.95	2 (0.06%)	985 (29.23%)	1082 (32.11%)	622 (18.46%)
	3370	0.99	0 (0.00%)	335 (9.94%)	421 (12.49%)	129 (3.83%)

Table 8. Frame Test Results for Daily Spot Price (Returns) Data

Specific Details: No Global AR 'Prewhitening' Fit;
Applied Frame by Frame AR(5) 'Prewhitening' Fit to Remove Linear Dependence
Frame by frame Hard Clipping of Residuals

Scenario / (State)	Total Num of Frames	False Alarm Threshold	Significant C Frames Num & (%)	Significant H Frames Num & (%)	Significant H4 Frames Num & (%)
NSW	3370	0.90	214 (6.35%)	329 (9.76%)	330 (9.79%)
	3370	0.95	114 (3.38%)	137 (4.07%)	192 (5.70%)
	3370	0.99	20 (0.59%)	21 (0.62%)	45 (1.34%)
QLD	3370	0.90	209 (6.20%)	233 (6.91%)	374 (11.10%)
	3370	0.95	84 (2.49%)	96 (2.85%)	124 (3.68%)
	3370	0.99	34** (1.01%)	13 (0.39%)	31 (0.92%)
VIC	3370	0.90	281 (8.34%)	331 (9.82%)	323 (9.58%)
	3370	0.95	151 (4.48%)	162 (4.81%)	181 (5.37%)
	3370	0.99	30 (0.89%)	30 (0.89%)	43 (1.28%)
SA	3370	0.90	243 (7.21%)	269 (7.98%)	361 (10.71%)
	3370	0.95	113 (3.35%)	108 (3.20%)	199 (5.91%)
	3370	0.99	20 (0.59%)	15 (0.45%)	32 (0.95%)

Notes:

** - false alarm threshold is set to 0.999999.